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

(19) World Intellectual Property Organization
International Bureau

(43) International Publication Date
26 May 2016 (26.05.2016)

WIPO | PCT

(10) International Publication Number
WO 2016/081118 A1

(51) International Patent Classification:
H04N 13/00 (2006.01) *H04N 5/232* (2006.01)

(21) International Application Number:
PCT/US2015/056156

(22) International Filing Date:
19 October 2015 (19.10.2015)

(25) Filing Language: English

(26) Publication Language: English

(30) Priority Data:
14/547,268 19 November 2014 (19.11.2014) US

(71) Applicant: **INTEL CORPORATION** [US/US]; 2200 Mission College Boulevard, Santa Clara, California 95054 (US).

(72) Inventors: **GROVER, Ginni**; 2200 Mission College Boulevard, RNB-6-61, Santa Clara, California 95054 (US). **NARAYANSWAMY, Ramkumar**; 845 Waite Drive, Boulder, Colorado 80303 (US).

(74) Agents: **TROP, Timothy N.** et al.; Trop, Pruner & Hu, P.C., 1616 S. Voss Rd., Ste. 750, Houston, Texas 77057-2631 (US).

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

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

Declarations under Rule 4.17:

— *as to the identity of the inventor (Rule 4.17(i))*

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

— *as to the applicant's entitlement to claim the priority of the earlier application (Rule 4.17(iii))*

Published:

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

(54) **Title**: MEASURING ACCURACY OF IMAGE BASED DEPTH SENSING SYSTEMS

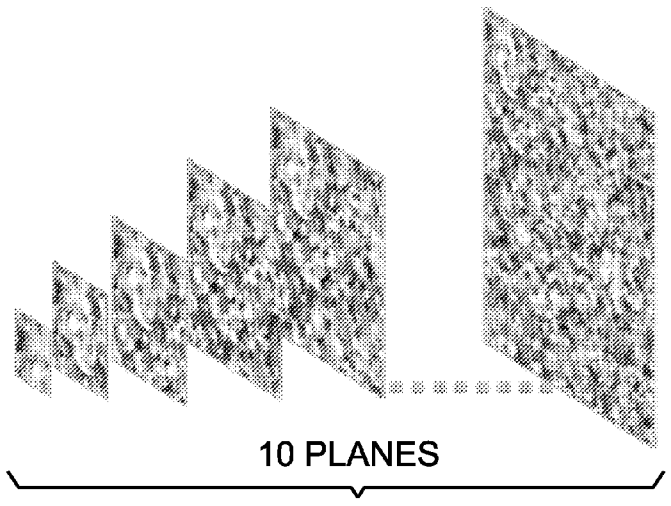


FIG. 1A

(57) **Abstract**: A special test target may enable standardized testing of performance of image based depth measuring systems. In addition, the error in measured depth with respect to the ground truth may be used as a metric of system performance. This test target may aid in identifying the limitations of the disparity estimation algorithms.

WO 2016/081118 A1

MEASURING ACCURACY OF IMAGE BASED DEPTH SENSING SYSTEMS

Background

[0001] Multi-camera imaging is an emerging field of computational photography. While the multi-camera imager is suitable for many applications, measurement of scene depth using parallax (disparity) is one of its fundamental advantages and leads to the most promising applications. Multiple camera platforms capture the same scene from different perspectives. The images are processed to determine the relative shift of the objects from one image to the next. The objects closer to the camera show more lateral shift, while objects farther from the camera show reduced lateral shift. This relative shift is referred to as disparity and is used to calculate depth. Various algorithms can be used for disparity (therefore depth) estimation.

[0002] Depth measurement performance in image based depth measuring systems depends on camera parameters, relative camera positions and disparity error. The disparity error further depends on scene characteristics such as object texture and color, noise characteristics of the cameras, optical aberrations in the lens, and the estimation algorithms.

[0003] Currently no standard test targets exist for determining depth measuring performance of the depth sensing systems. Test targets exist for measuring various camera properties such as Macbeth ColorChecker to measure color response, ISO-12233 for sharpness, etc. But no test targets exist for determining depth sensing performance of the cameras.

Brief Description Of The Drawings

[0004] Some embodiments are described with respect to the following figures:

Figure 1A is a schematic depiction of a depth target with boards placed at different scene depths according to one embodiment;

Figure 1B is a depth map of the scene with dark to white color representing farthest to closest distance according to one embodiment and a color bar in meters;

Figure 2A is a ground truth depth map for one embodiment;

Figure 2B is a measured depth map with color bar in meters for one embodiment;

Figure 3A shows the geometry of 6+1 and 2+1 camera systems with a maximum baseline of 78 mm and all cameras being equally separated according to one embodiment;

Figure 3B is a plot of errors for the systems of Figure 3A in depth measurement according to one embodiment where the dashed line shows the theoretical error calculated for 0.2 pixel errors in disparity measurement and the solid line is for 0.1 pixel errors;

Figure 4 is a camera image of a depth target with colored Dead-leaves chart at ten different depth positions according to one embodiment;

Figure 5 is a flow chart for one embodiment;

Figure 6 is a system depiction for one embodiment; and

Figure 7 is a front elevational view for one embodiment.

Detailed Description

[0005] A special test target may enable standardized testing of performance of image based depth measuring systems. In addition, the error in measured depth with respect to the ground truth may be used as a metric of system performance. This test target may also aid in identifying the limitations of the depth measuring cameras and disparity estimation algorithms.

[0006] A well characterized scene with two dimensional customized charts stacked at different but known depths may be assembled. The charts may be “customized” in that they can have different textures because image-based depth measuring algorithm results are texture dependent. In general, higher spatial frequency texture is required on charts to measure depth.

[0007] An example of the use of two dimensional charts is shown in Figure 1A where a scene is generated with 10 boards of progressively larger size. In one embodiment, the boards are rectangular with their centers aligned but other known

shapes may be used and non-aligned arrangements may also be used. The size of the boards is determined by field of view of the camera and the distance of the boards from the camera. An edge of each board may be less than $2*z*\tan(\theta/2)$ where z is the distance of the board from the camera and θ is the field of view of the camera along that direction. This size enables capturing the entire board in the camera.

[0008] Regarding the texture on these boards, a pattern on each board may be appropriately scaled with respect to its distance from the camera. This is because a camera magnifies objects based on their distance from it. Therefore, the same texture at shorter distances will exhibit lower spatial frequencies than at longer distances. So for the case when distance from the camera $z2$ is greater than distance from the camera $z1$, the displayed pattern on the board at distance $z2$ may be scaled by a factor of $z2/z1$ with respect to the displayed pattern on board at $z1$. This is done to make sure that the measured error in depth at different distances differs only due to the difference in depth and not due to difference in textures. The aim is to keep the spatial frequency content the same at different depths to make the error measurement independent of this parameter. Therefore, in practice if an image with a texture is displayed at a depth of 1 meter, the same image should be scaled up twice when placed at depth of 2 meters.

[0009] The ground truth depth map of this scene is shown in Figure 1B with dark to light colors representing farthest to closest distance. The “ground truth” refers to a set of measurements known to be more accurate or exact. This scene is then captured with a depth sensing system and depth is estimated. The measured depth is compared against the ground truth and measurement accuracy reported.

[0010] A real target can be used for physical testing and a virtual image of such a real target via computer simulation may be used for testing. The added advantage of testing via computer simulation is knowledge of the exact ground truth depth map. A test target is modelled and digitally rendered via three-dimensional (3D) rendering packages plus added optical and sensor effects, thereby giving exact ground truth. Another consideration for the test target is the pattern to be displayed on the stacked boards.

[0011] In one embodiment a 2+1-hybrid camera system may be compared with a 6+1-hybrid camera system. A 2+1 hybrid camera system has three cameras, two of which are two end cameras, as shown in Figure 3A. The end cameras may be 1 megapixel resolution cameras and the center camera may be an 8 megapixel resolution camera in one embodiment. This is referred to as a hybrid multi-camera array. A similar convention applies for the 6+1-hybrid system having a center camera and three cameras to both sides of the center camera.

[0012] A depth target may use colored Dead-leaves charts at different depths in one embodiment. The Dead-leaves chart is known to have a wide range of spatial frequencies and to be scale-invariant i.e. scaling of the chart at different depths is not required to keep the spatial frequency content same. Therefore for the Dead-leaves chart and its variants, the scaling at different depths as mentioned above is not required. This makes the test scene setup easier. However if another texture, such as random noise, is used, the scaling may be required. A scene containing this depth target is digitally rendered. Images of this scene, captured by different cameras of the multi-camera array, are simulated. Image simulation includes all the optical, color and noise effects which are observed in real camera images. The resulting images are then fed into the depth (disparity) estimation algorithms.

[0013] The ground truth depth map and computed depth map are shown in Figures 2A and 2B, respectively. The bar shows distances in meters. The black patches (Figure 2B) in the resulting depth map are regions where the particular disparity algorithm used is unable to estimate depth. This is because these regions have lower spatial frequency content than required by this particular algorithm. This example demonstrates that the test target can also be used to characterize the limitations of the algorithms.

[0014] Figure 3A shows the geometry of the tested two systems. To compare the performances of the two systems, the errors in depth measurement (with respect to the ground truth) are shown in Figure 3B. The dashed line bars and dashed curve are for the 2+1 camera system. The theoretical curves, which depend on the widest baseline, are also shown. The results show that the 6+1 camera system is better both for depth accuracy and depth range than the 2+1 camera system. As known

from theory, the simulated results also indicate that the error in disparity measurement is also dependent on depth of the system. The simulation results reasonably follow the theoretical values. However they do not exactly match the theoretical values because the theoretical values are calculated with the same disparity error at each depth. But in practice, the disparity error could be slightly different at different depths because of variation in texture, color, optical aberration and lighting with depth.

[0015] The depth error is theoretically determined by the following formula:

$$\textit{depth error} = \frac{depth^2 \times disparity\ error}{focal\ length \times baseline}$$

showing that the error in depth measurement depends on disparity error, focal length of the lens, depth of the imaged object and maximum baseline (i.e. distance between end cameras) of the system. In the two systems, all parameters are the same except for the disparity error, which is different because of number of cameras in the systems and the algorithm used. The plot in 3B shows that error in depth or disparity is lower if more cameras are used.

[0016] Thus, the depth target and the measurement metric are useful for determining depth measurement performance of the whole system for comparison and also allow assessment of individual algorithms.

[0017] Depth sensors have different depth ranges depending upon the technology they use and the applications they are designed for. For example, some sensors are designed for gesture recognition and only work for the range of few centimeters to couple of meters, whereas, some sensors work in the range of a 1 meter to 10 meters. Given the different ranges of operation, the depth target needs to be modified to accommodate for testing within the required range. The boards in the test target have to be placed within depth sensing range of a particular system.

[0018] Moreover, the boards can be tilted to estimate depth resolution of the system. The boards can be tilted in vertical or horizontal or both directions to determine the resolution of the system along that direction. Resolution relates to

how continuous or fine is the measurement. Algorithms mostly quantize disparity measurement for speed and storage purposes.

[0019] Estimation algorithms for image based depth measurement systems depend on scene texture (in other words spatial frequency content) and on camera response to the scene texture. In these systems, generally, low texture in the final image gives poor performance whereas higher texture gives better performance. Therefore, the scene should have high spatial frequency content for depth estimation. The Dead-leaves chart has been found to represent texture for most natural scenes, has a wide range of spatial frequencies and most importantly is scale invariant to distance between the camera and the chart. Therefore the colored Dead-leaves chart may be used in the depth test target for image-based depth measuring systems.

[0020] However, for other depth measurement methods, such as structured illumination and time of flight, the pattern displayed doesn't matter since they are not image based. However, a material with higher reflection in infrared range is desired for these methods. So the boards may be coated with such materials.

[0021] In the following discussion, one disparity estimation algorithm useful in some embodiments is described. Other algorithms may also be used, including but not limited to disparity estimation by phase matching, graph cut or block matching. A hybrid camera array poses technical difficulties in matching across images from different types of sensors (sensors might have different resolutions, color characteristics, noise patterns, etc.). Given all the images from multiple cameras, images may be downsized to the same resolution as the lowest resolution camera. By doing that, all images have the same resolution and can efficiently perform pixel-to-pixel correspondence search in a disparity calculation. Images will then be transformed to new features representations. Since the color characteristics, noise patterns are quite different across different cameras, the variance may be reduced depending on what features are extracted to represent images. For example, if RGB color is used as features, "color histogram normalization" may be used to match images to the same color characteristics as the reference camera. If grayscale is used as features, "intensity histogram normalization" may be used to match images to the same intensity characteristics as the reference camera. Features such as

gradient, census, local binary pattern (LBP) are less sensitive to color variations, but sensitive to noise, hence "noise normalization" may be used to match images to the same noise characteristics as the reference camera.

[0022] Once features are extracted, a multi-baseline disparity algorithm can be implemented. First, an adaptive shape support region instead of a fixed size region is desired for accurate disparity estimates. Therefore only the pixels of the same depth are used for sum of absolute difference (SAD) calculation in one embodiment. To find the adaptive shape support region, each pixel (x, y) will extend to four directions (left, right, up and down) until it hits a pixel that the color, gradient or grayscale difference between this pixel and the pixel (x, y) is beyond certain thresholds. For each pair of cameras (Cr - reference camera versus Ci - the other camera), and each candidate disparity d=(dx, dy), where dx and dy are the candidate disparity on horizontal and vertical direction calculated using baseline ratio dx = d*bi_x/bi and dy = d*bi_y/bi, construct a support region for each pixel (x, y) in Cr and corresponding comparing pixel (x + dx, y + dy) in Ci. Repeat the same process to construct support regions for all pixels.

[0023] Second, for each pair of cameras Cr and Ci, initialize pixel-wise absolute difference (AD) using features at each candidate disparity d. Equation 1 shows an AD calculation for each pixel (x, y):

$$AD_i(d) = \left[I_r(x+k, y+t) - I_i(x+k+d_x, y+t+d_y) \right] \quad (1)$$

[0024] Third, for each pair of cameras Cr and Ci, aggregate the AD errors using a sum of absolute difference (SAD) in the support region S using integral image techniques for efficient calculation:

$$SAD_i(d) = \sum_{(k,t) \in S_{i,r}(d)} | I_r(x+k, y+t) - I_i(x+k+d_x, y+t+d_y) | \quad (2)$$

[0025] Fourth, resize all the pairwise SAD error costs between cameras Cr and Ci to the longest baseline based on baseline ratio using bilinear interpolation, and aggregate them together using an aggregate function. The aggregate function

could, for example, either be $SAD_i(d)$ with the minimum error or a subset of $\{SAD_i(d)\}$ with minimum error and take the average.

$$E(d) = aggregate(SAD \underset{1}{^{resized}}(d)) \qquad (3)$$

[0026] Finally, multi-baseline disparity value for a given pixel (x,y) in the reference camera along the longest baseline is calculated by finding the minimum d in the summarized error map from all camera pairs:

$$d(x,y) = \underset{d}{^{argmin}} E(d) \qquad (4)$$

[0027] Disparity from step 2 might contain noise. In order to get cleaner and smoother disparity output, a refinement step removes noise and low confident disparity values. Methods such as the uniqueness of global minimum cost, variance of the cost curve, etc. may be used. Then a median filter, joint bilateral filter, etc., may be used to fill holes that got removed in the previous step. Finally, if disparity map's resolution is lower than the original resolution of the reference camera image, the disparity map is upscaled to the same resolution as reference camera.

[0028] Figure 4 shows an image of a 3D depth target with colored Dead-leaves charts at different depths. One of these texture charts may be placed to determine depth measurement accuracy. If color, dynamic range and other responses of the cameras need to be characterized simultaneously with depth, test charts for these can also be placed on the stacked boards. One example of a use case is multiple cameras may have different color response, so a Macbeth color checker may be placed in the front. That will give color response of each camera and ability to compensate for differences.

[0029] A method to design and test with the depth target is described in the flow chart 10 shown in Figure 5. The flowchart shows both the simulation 16 and physical 28 testing paths. Either of these methods can be used and is independent of one another. Simulation is easier because the ground truth is exact.

[0030] The simulation testing path 16 including the steps set forth in blocks 12, 14, 18, 20, 22 and 26 may be implemented in software, firmware and/or hardware in

some embodiments. In software and firmware embodiments these steps may be implemented by computer executed instructions stored in one or more non-transitory computer readable media such as magnetic, optical or semiconductor storages.

[0031] Depending on the depth range of the three-dimensional sensor, the locations where boards need to be placed to be within that depth range is determined as indicated in block 12. In the simulated embodiment, an image of the boards is designed instead of using physical boards. Then the target scene with the physical or virtual boards is designed at the desired depths and with the desired charts as indicated in block 14. In the simulation testing path 16 the scene is digitally rendered and the associated ground truth depth map is digitally rendered (block 18). The image capture of the scene is simulated with the virtual multi-camera array as indicated in block 20.

[0032] In the actual physical capture depth testing sequence 28, the physical scene is set up as indicated at 30 and the image is captured with a multi-camera array as indicated in block 32.

[0033] In both the simulation and actual physical capture, the disparity estimation algorithm is run on the images and the result is converted to depth as indicated in block 22. Finally the estimated depths are compared with the ground truth in block 26.

[0034] A combination of the previously existing Dead-leaves chart and use of charts at different depth positions allows evaluation of image based depth sensing systems. Thus, this idea uses the Dead-leaves chart for depth testing which has not been proposed before.

[0035] A computer simulated test target may be used for system performance evaluation. The computer generated 3D test scenes may be used in camera simulation. The advantage, in some embodiments, is knowledge of exact ground truth depth map for evaluation.

[0036] Computer-graphics generated images may be used as the scenes to be captured by the camera. These scenes may be generated with much higher spatial and color resolution than the camera, without any distortions seen in the real

cameras. In multi-camera systems, the spatial separation between the cameras induces parallax in the captured images such that objects at different depths are seen shifted laterally between the images. For multi-camera systems, images for a single scene are generated from multiple viewpoint depending upon the geometry of the cameras. These images then serve as input scenes for the cameras.

[0037] Once the 2D projection of the scene is created, the scene image is converted to the optical image by light propagation methods. Light sources in 3D models determine the scene illumination and parameters like the type of light source (lambertian, point, diffuse, extended). The color temperature of the light and the luminance (brightness) can be controlled in the 3D model.

[0038] This high resolution version of the scene is propagated to the lens entrance-pupil. The scene is then processed such that the lens forms a de-magnified image of the scene on the sensor plane. Lens aberration and diffraction effects are faithfully imparted to the image generated on the sensor-plane. Various effects such as lens-to-lens misalignment, stray-light effects are also modelled.

[0039] The model includes opto-mechanical aspects of the camera module i.e. the alignment between the optical axis of the lens and the sensor center. It includes lens back-focal-length variations and error in the sensor position with respect to the lens focal plane. Thermal effects and other mechanical effects which manifest as one of the optical aberrations in the final image are also included in the model.

[0040] The sensor samples the scene and creates an image adding noise from sources such as shot-noise, read-noise, photo-response-non-uniformity, fixed pattern noise, pixel-cross-talk and other electronic noise sources. Other aspects of the sensor, such as photon to electron conversion, finite pixel size, color filters and efficiency of light conversion, etc are taken into account in the sensor model.

[0041] The “raw” image from the sensor is processed by a typical camera image-signal processing pipe to deliver an RGB image from each camera sub-system. All these RGB images are fed into a “multi-camera” image signal processor (ISP) to extract disparity, depth and similar multi-camera parameter. These RGB images and their “multi-camera” metadata are sent to the media-view which renders the special

effects as chosen by the end-user. The simulation may model all these aspects with high accuracy.

[0042] During the process of generating the computer-generated scenes at the onset, a “ground-truth depth-map” of the scene is generated with respect to each of the cameras. Having the “ground-truth” allows a comparison with the performance of our disparity and depth extraction algorithms as a function of various parameters such as texture, illumination, field-position, object distance and other characteristics. This analysis is highly desirable since passive-depth measurement is scene dependent and often has to be tested exhaustively over a range of scenes.

[0043] In the scene simulation part, high spectral resolution information about the scene is gathered enabling testing of the chromatic fidelity of the camera system if needed.

[0044] Figure 6 illustrates an embodiment of a system 700. In embodiments, system 700 may be a media system although system 700 is not limited to this context. For example, system 700 may be incorporated into a personal computer (PC), laptop computer, ultra-laptop computer, tablet, touch pad, portable computer, handheld computer, palmtop computer, personal digital assistant (PDA), cellular telephone, combination cellular telephone/PDA, television, smart device (e.g., smart phone, smart tablet or smart television), mobile internet device (MID), messaging device, data communication device, and so forth.

[0045] In embodiments, system 700 comprises a platform 702 coupled to a display 720. Platform 702 may receive content from a content device such as content services device(s) 730 or content delivery device(s) 740 or other similar content sources. A navigation controller 750 comprising one or more navigation features may be used to interact with, for example, platform 702 and/or display 720. Each of these components is described in more detail below.

[0046] In embodiments, platform 702 may comprise any combination of a chipset 705, processor 710, memory 712, storage 714, graphics subsystem 715, applications 716 and/or radio 718. Chipset 705 may provide intercommunication among processor 710, memory 712, storage 714, graphics subsystem 715,

applications 716 and/or radio 718. For example, chipset 705 may include a storage adapter (not depicted) capable of providing intercommunication with storage 714.

[0047] Processor 710 may be implemented as Complex Instruction Set Computer (CISC) or Reduced Instruction Set Computer (RISC) processors, x86 instruction set compatible processors, multi-core, or any other microprocessor or central processing unit (CPU). In embodiments, processor 710 may comprise dual-core processor(s), dual-core mobile processor(s), and so forth. The processor may implement the sequence of Figure 5 together with memory 712.

[0048] Memory 712 may be implemented as a volatile memory device such as, but not limited to, a Random Access Memory (RAM), Dynamic Random Access Memory (DRAM), or Static RAM (SRAM).

[0049] Storage 714 may be implemented as a non-volatile storage device such as, but not limited to, a magnetic disk drive, optical disk drive, tape drive, an internal storage device, an attached storage device, flash memory, battery backed-up SDRAM (synchronous DRAM), and/or a network accessible storage device. In embodiments, storage 714 may comprise technology to increase the storage performance enhanced protection for valuable digital media when multiple hard drives are included, for example.

[0050] Graphics subsystem 715 may perform processing of images such as still or video for display. Graphics subsystem 715 may be a graphics processing unit (GPU) or a visual processing unit (VPU), for example. An analog or digital interface may be used to communicatively couple graphics subsystem 715 and display 720. For example, the interface may be any of a High-Definition Multimedia Interface, DisplayPort, wireless HDMI, and/or wireless HD compliant techniques. Graphics subsystem 715 could be integrated into processor 710 or chipset 705. Graphics subsystem 715 could be a stand-alone card communicatively coupled to chipset 705.

[0051] The graphics and/or video processing techniques described herein may be implemented in various hardware architectures. For example, graphics and/or video functionality may be integrated within a chipset. Alternatively, a discrete

graphics and/or video processor may be used. As still another embodiment, the graphics and/or video functions may be implemented by a general purpose processor, including a multi-core processor. In a further embodiment, the functions may be implemented in a consumer electronics device.

[0052] Radio 718 may include one or more radios capable of transmitting and receiving signals using various suitable wireless communications techniques. Such techniques may involve communications across one or more wireless networks. Exemplary wireless networks include (but are not limited to) wireless local area networks (WLANs), wireless personal area networks (WPANs), wireless metropolitan area network (WMANs), cellular networks, and satellite networks. In communicating across such networks, radio 718 may operate in accordance with one or more applicable standards in any version.

[0053] In embodiments, display 720 may comprise any television type monitor or display. Display 720 may comprise, for example, a computer display screen, touch screen display, video monitor, television-like device, and/or a television. Display 720 may be digital and/or analog. In embodiments, display 720 may be a holographic display. Also, display 720 may be a transparent surface that may receive a visual projection. Such projections may convey various forms of information, images, and/or objects. For example, such projections may be a visual overlay for a mobile augmented reality (MAR) application. Under the control of one or more software applications 716, platform 702 may display user interface 722 on display 720.

[0054] In embodiments, content services device(s) 730 may be hosted by any national, international and/or independent service and thus accessible to platform 702 via the Internet, for example. Content services device(s) 730 may be coupled to platform 702 and/or to display 720. Platform 702 and/or content services device(s) 730 may be coupled to a network 760 to communicate (e.g., send and/or receive) media information to and from network 760. Content delivery device(s) 740 also may be coupled to platform 702 and/or to display 720.

[0055] In embodiments, content services device(s) 730 may comprise a cable television box, personal computer, network, telephone, Internet enabled devices or appliance capable of delivering digital information and/or content, and any other similar device capable of unidirectionally or bidirectionally communicating content between content providers and platform 702 and/display 720, via network 760 or directly. It will be appreciated that the content may be communicated unidirectionally and/or bidirectionally to and from any one of the components in system 700 and a content provider via network 760. Examples of content may include any media information including, for example, video, music, medical and gaming information, and so forth.

[0056] Content services device(s) 730 receives content such as cable television programming including media information, digital information, and/or other content. Examples of content providers may include any cable or satellite television or radio or Internet content providers. The provided examples are not meant to limit the applicable embodiments.

[0057] In embodiments, platform 702 may receive control signals from navigation controller 750 having one or more navigation features. The navigation features of controller 750 may be used to interact with user interface 722, for example. In embodiments, navigation controller 750 may be a pointing device that may be a computer hardware component (specifically human interface device) that allows a user to input spatial (e.g., continuous and multi-dimensional) data into a computer. Many systems such as graphical user interfaces (GUI), and televisions and monitors allow the user to control and provide data to the computer or television using physical gestures.

[0058] Movements of the navigation features of controller 750 may be echoed on a display (e.g., display 720) by movements of a pointer, cursor, focus ring, or other visual indicators displayed on the display. For example, under the control of software applications 716, the navigation features located on navigation controller 750 may be mapped to virtual navigation features displayed on user interface 722, for example. In embodiments, controller 750 may not be a separate component but

integrated into platform 702 and/or display 720. Embodiments, however, are not limited to the elements or in the context shown or described herein.

[0059] In embodiments, drivers (not shown) may comprise technology to enable users to instantly turn on and off platform 702 like a television with the touch of a button after initial boot-up, when enabled, for example. Program logic may allow platform 702 to stream content to media adaptors or other content services device(s) 730 or content delivery device(s) 740 when the platform is turned "off." In addition, chip set 705 may comprise hardware and/or software support for 5.1 surround sound audio and/or high definition 7.1 surround sound audio, for example. Drivers may include a graphics driver for integrated graphics platforms. In embodiments, the graphics driver may comprise a peripheral component interconnect (PCI) Express graphics card.

[0060] In various embodiments, any one or more of the components shown in system 700 may be integrated. For example, platform 702 and content services device(s) 730 may be integrated, or platform 702 and content delivery device(s) 740 may be integrated, or platform 702, content services device(s) 730, and content delivery device(s) 740 may be integrated, for example. In various embodiments, platform 702 and display 720 may be an integrated unit. Display 720 and content service device(s) 730 may be integrated, or display 720 and content delivery device(s) 740 may be integrated, for example. These examples are not meant to be scope limiting.

[0061] In various embodiments, system 700 may be implemented as a wireless system, a wired system, or a combination of both. When implemented as a wireless system, system 700 may include components and interfaces suitable for communicating over a wireless shared media, such as one or more antennas, transmitters, receivers, transceivers, amplifiers, filters, control logic, and so forth. An example of wireless shared media may include portions of a wireless spectrum, such as the RF spectrum and so forth. When implemented as a wired system, system 700 may include components and interfaces suitable for communicating over wired communications media, such as input/output (I/O) adapters, physical connectors to connect the I/O adapter with a corresponding wired communications

medium, a network interface card (NIC), disc controller, video controller, audio controller, and so forth. Examples of wired communications media may include a wire, cable, metal leads, printed circuit board (PCB), backplane, switch fabric, semiconductor material, twisted-pair wire, co-axial cable, fiber optics, and so forth.

[0062] Platform 702 may establish one or more logical or physical channels to communicate information. The information may include media information and control information. Media information may refer to any data representing content meant for a user. Examples of content may include, for example, data from a voice conversation, videoconference, streaming video, electronic mail ("email") message, voice mail message, alphanumeric symbols, graphics, image, video, text and so forth. Data from a voice conversation may be, for example, speech information, silence periods, background noise, comfort noise, tones and so forth. Control information may refer to any data representing commands, instructions or control words meant for an automated system. For example, control information may be used to route media information through a system, or instruct a node to process the media information in a predetermined manner. The embodiments, however, are not limited to the elements or in the context shown or described in Figure 6.

[0063] As described above, system 700 may be embodied in varying physical styles or form factors. Figure 7 illustrates embodiments of a small form factor device 800 in which system 700 may be embodied. In embodiments, for example, device 800 may be implemented as a mobile computing device having wireless capabilities. A mobile computing device may refer to any device having a processing system and a mobile power source or supply, such as one or more batteries, for example.

[0064] As described above, examples of a mobile computing device may include a personal computer (PC), laptop computer, ultra-laptop computer, tablet, touch pad, portable computer, handheld computer, palmtop computer, personal digital assistant (PDA), cellular telephone, combination cellular telephone/PDA, television, smart device (e.g., smart phone, smart tablet or smart television), mobile internet device (MID), messaging device, data communication device, and so forth.

[0065] Examples of a mobile computing device also may include computers that are arranged to be worn by a person, such as a wrist computer, finger computer, ring computer, eyeglass computer, belt-clip computer, arm-band computer, shoe computers, clothing computers, and other wearable computers. In embodiments, for example, a mobile computing device may be implemented as a smart phone capable of executing computer applications, as well as voice communications and/or data communications. Although some embodiments may be described with a mobile computing device implemented as a smart phone by way of example, it may be appreciated that other embodiments may be implemented using other wireless mobile computing devices as well. The embodiments are not limited in this context.

[0066] The processing techniques described herein may be implemented in various hardware architectures. For example, the functionality may be integrated within a chipset. Alternatively, a discrete processor may be used. As still another embodiment, the functions may be implemented by a general purpose processor, including a multicore processor.

[0067] The following clauses and/or examples pertain to further embodiments:

One example embodiment may be a test target for testing depth measuring ability comprising an array of planar boards of successively larger size placed in different but known depths, and said boards depicting colored Dead leaves. The target may also include where boards are progressively scaled up by an amount that depends on the distance from a camera. The target may also include wherein an edge of each board is less than $2*z*\tan(\theta/2)$, where z is distance of the board from the camera and θ is field of view of the camera. The target may also include wherein the distance from the camera is z2 is greater than a distance z1, the displayed pattern on the board at distance z2 is scaled by a factor of z2/z1 with respect to the pattern on the board at distance z1. The target may also include wherein the boards are within the depth sensing range of a camera system to be tested.

[0068] Another example embodiment may be a method comprising arranging a test target in the form of a plurality of real or simulated spaced apart planar boards of

successively larger size, said boards depicting Dead colored leaves, within the depth range of a camera system to be tested, and designing a target scene. The method may also include digitally rendering the scene and the associated ground truth map. The method may also include simulating image capture of the scene with a virtual multi-camera array. The method may also include running a disparity estimation algorithm on the images. The method may also include converting the images to depths. The method may also include comparing an estimated depth to ground truth.

[0069] In another example embodiment one or more non-transitory computer readable media storing instructions executed to perform a sequence comprising arranging a test target in the form of a plurality of images of spaced apart planar boards of successively larger size, said boards depicting Dead colored leaves, within the depth range of a camera system to be tested, and designing a target scene. The media may include said sequence including digitally rendering the scene and the associated ground truth map. The media may include said sequence including simulating image capture of the scene with a virtual multi-camera array. The media may include said sequence said sequence including running a disparity estimation algorithm on the images. The media may include said sequence including converting the images to depths. The media may include said sequence including comparing an estimated depth to ground truth.

[0070] Another example embodiment may be an apparatus comprising a hardware processor to arrange a test target in the form of images of a plurality of spaced apart planar boards of successively larger size, said boards depicting Dead colored leaves, within the depth range of a camera system to be tested and design a target scene, and a storage coupled to said processor. The apparatus may include said sequence including digitally rendering the scene and the associated ground truth map. The apparatus may include said sequence including simulating image capture of the scene with a virtual multi-camera array. The apparatus may include said sequence including running a disparity estimation algorithm on the images. The apparatus may include said sequence including converting the images to depths. The apparatus may include said sequence including comparing an estimated depth

to ground truth. The apparatus may include boards of successively larger size placed in different but known depths and said boards depicting colored Dead leaves. The apparatus may include a battery coupled to the processor.

[0071] References throughout this specification to “one embodiment” or “an embodiment” mean that a particular feature, structure, or characteristic described in connection with the embodiment is included in at least one implementation encompassed within the present disclosure. Thus, appearances of the phrase “one embodiment” or “in an embodiment” are not necessarily referring to the same embodiment. Furthermore, the particular features, structures, or characteristics may be instituted in other suitable forms other than the particular embodiment illustrated and all such forms may be encompassed within the claims of the present application.

[0072] While a limited number of embodiments have been described, those skilled in the art will appreciate numerous modifications and variations therefrom. It is intended that the appended claims cover all such modifications and variations as fall within the true spirit and scope of this disclosure.

What is claimed is:

1. A test target for testing depth measuring ability comprising:
an array of planar boards of successively larger size placed in different but known depths; and
said boards depicting colored Dead leaves.

2. The target of claim 1 where boards are progressively scaled up by an amount that depends on the distance from a camera.

3. The target of claim 2 wherein an edge of each board is less than $2*z*\tan(\theta/2)$, where z is distance of the board from the camera and θ is field of view of the camera.

4. The target of claim 2 wherein the distance from the camera is z2 is greater than a distance z1, the displayed pattern on the board at distance z2 is scaled by a factor of z2/z1 with respect to the pattern on the board at distance z1.

5. The target of claim 1 wherein the boards are within the depth sensing range of a camera system to be tested.

6. A method comprising:
arranging a test target in the form of a plurality of real or simulated spaced apart planar boards of successively larger size, said boards depicting Dead colored leaves, within the depth range of a camera system to be tested; and
running a disparity estimation algorithm.

7. The method of claim 6 including digitally rendering the scene and the associated ground truth map.

8. The method of claim 7 including simulating image capture of the scene with a virtual multi-camera array.

9. The method of claim 6 including converting the result of said algorithm to depth.

10. The method of claim 9 including comparing an estimated depth to ground truth.

11. One or more non-transitory computer readable media storing instructions executed to perform a sequence comprising:

developing a test target in the form of a plurality of images of spaced apart planar boards of successively larger size, said boards depicting Dead colored leaves, within the depth range of a camera system to be tested; and

running a disparity estimation algorithm.

12. The media of claim 11, said sequence including digitally rendering the scene and the associated ground truth map.

13. The media of claim 12, said sequence including simulating image capture of the scene with a virtual multi-camera array.

14. The media of claim 11, said sequence including converting the result of said algorithm to depth.

15. The media of claim 14, said sequence including comparing an estimated depth to ground truth.

16. An apparatus comprising:

a hardware processor to simulate a test target in the form of images of a plurality of spaced apart planar boards of successively larger size, said boards depicting Dead colored leaves, within the depth range of a camera system to be tested and design a target scene; and

a storage coupled to said processor.

17. The apparatus of claim 16, said sequence including digitally rendering the scene and the associated ground truth map.

18. The apparatus of claim 17, said sequence including simulating image capture of the scene with a virtual multi-camera array.

19. The apparatus of claim 16, said sequence including running a disparity estimation algorithm on the images.

20. The apparatus of claim 19, said sequence including converting the images to depths.

21. The apparatus of claim 20, said sequence including comparing an estimated depth to ground truth.

22. The apparatus of claim 16 including boards of successively larger size placed in different but known depths and said boards depicting colored Dead leaves.

23. The apparatus of claim 16 including a battery coupled to the processor.

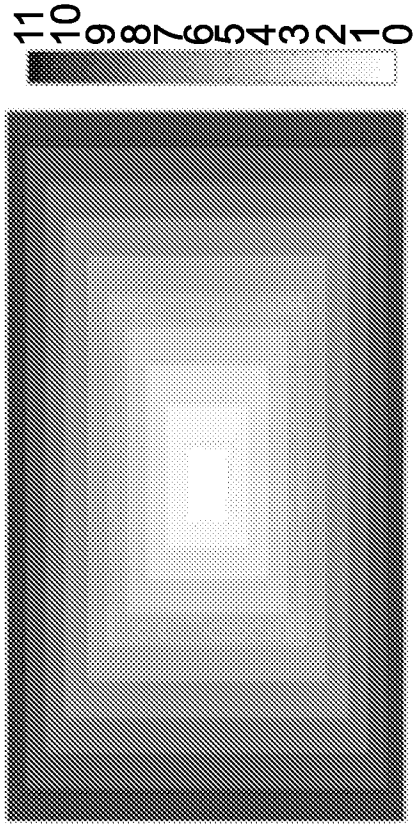


FIG. 1B

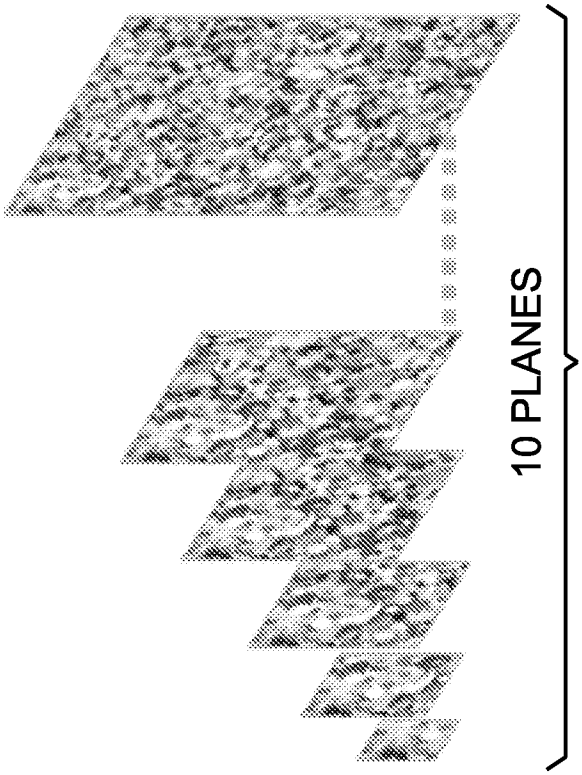


FIG. 1A



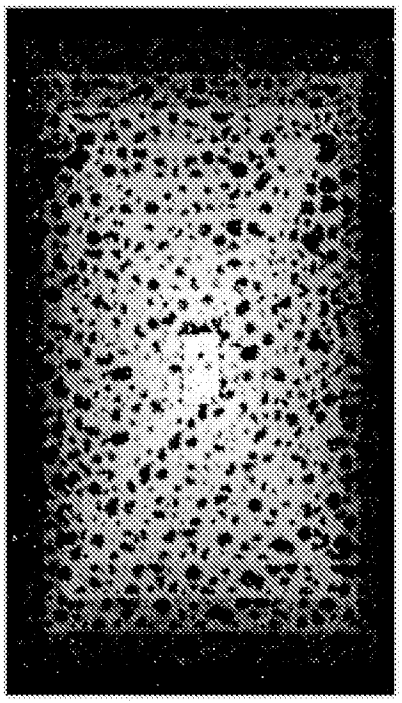


FIG. 2B

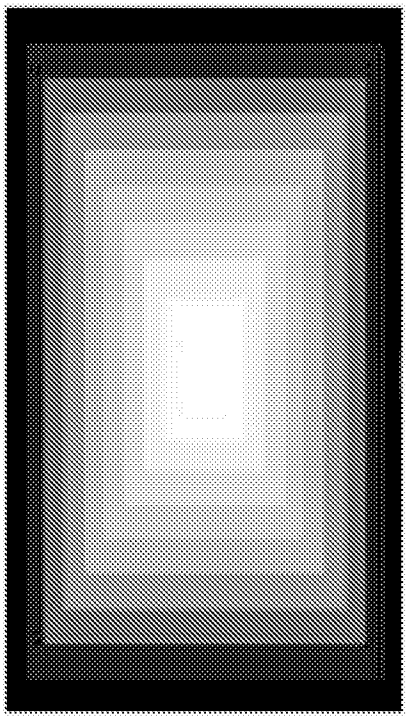


FIG. 2A

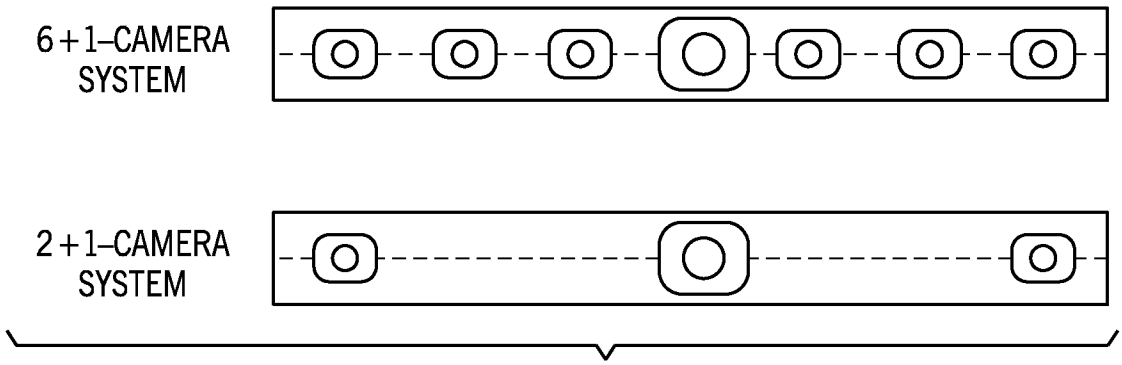

6+1–CAMERA SYSTEM
2+1–CAMERA SYSTEM

FIG. 3A

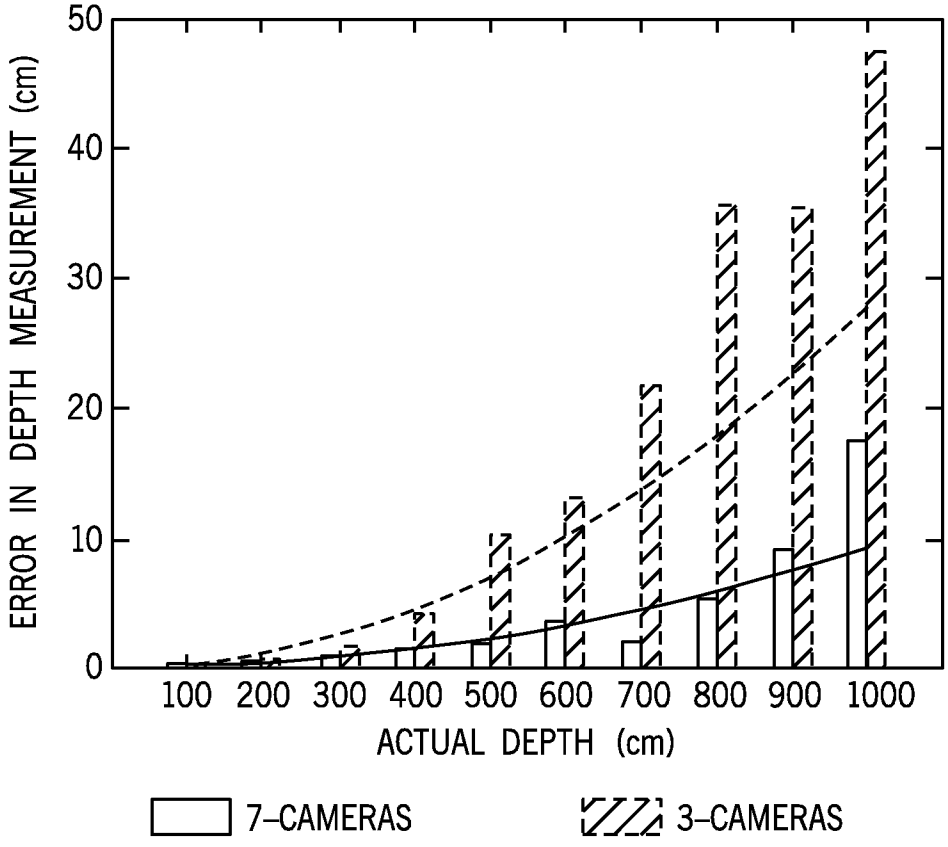

ERROR IN DEPTH MEASUREMENT (cm)
0
10
20
30
40
50
100 200 300 400 500 600 700 800 900 1000
ACTUAL DEPTH (cm)
7–CAMERAS
3–CAMERAS

FIG. 3B

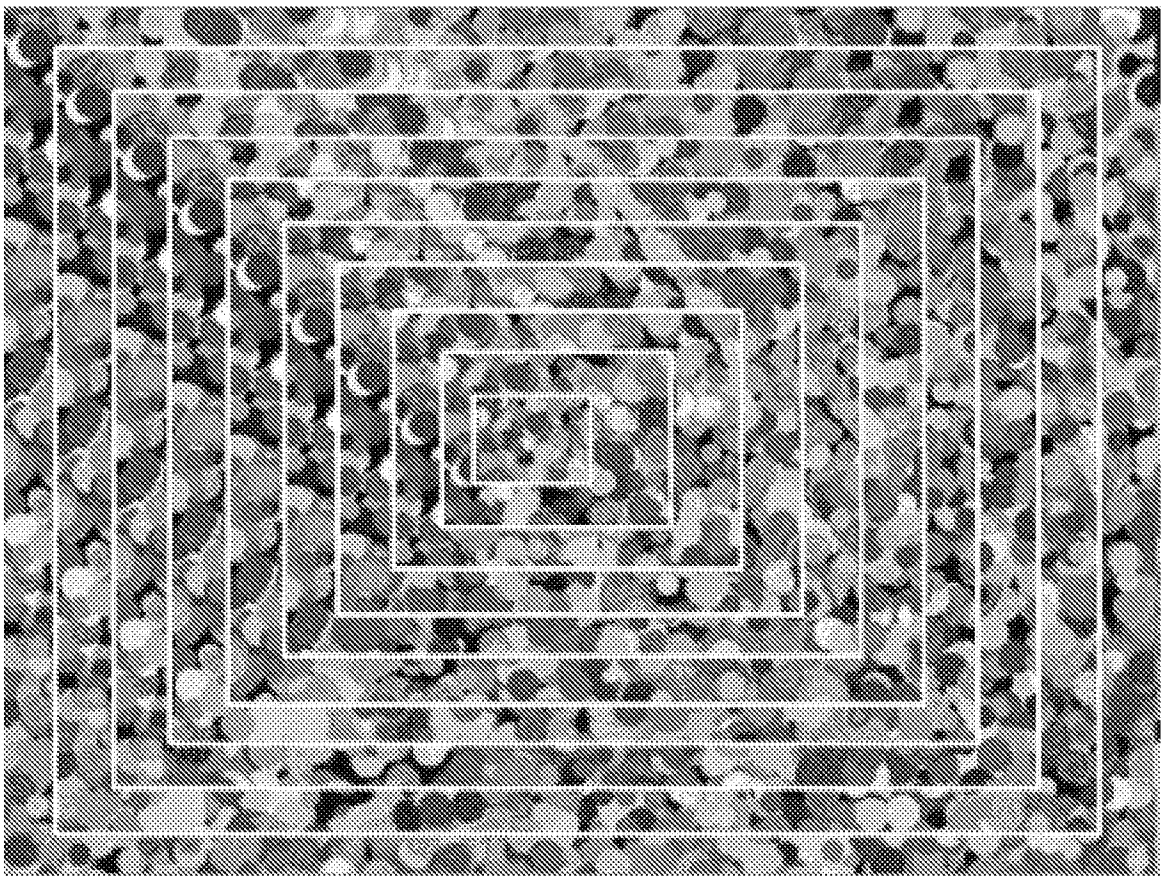


FIG. 4

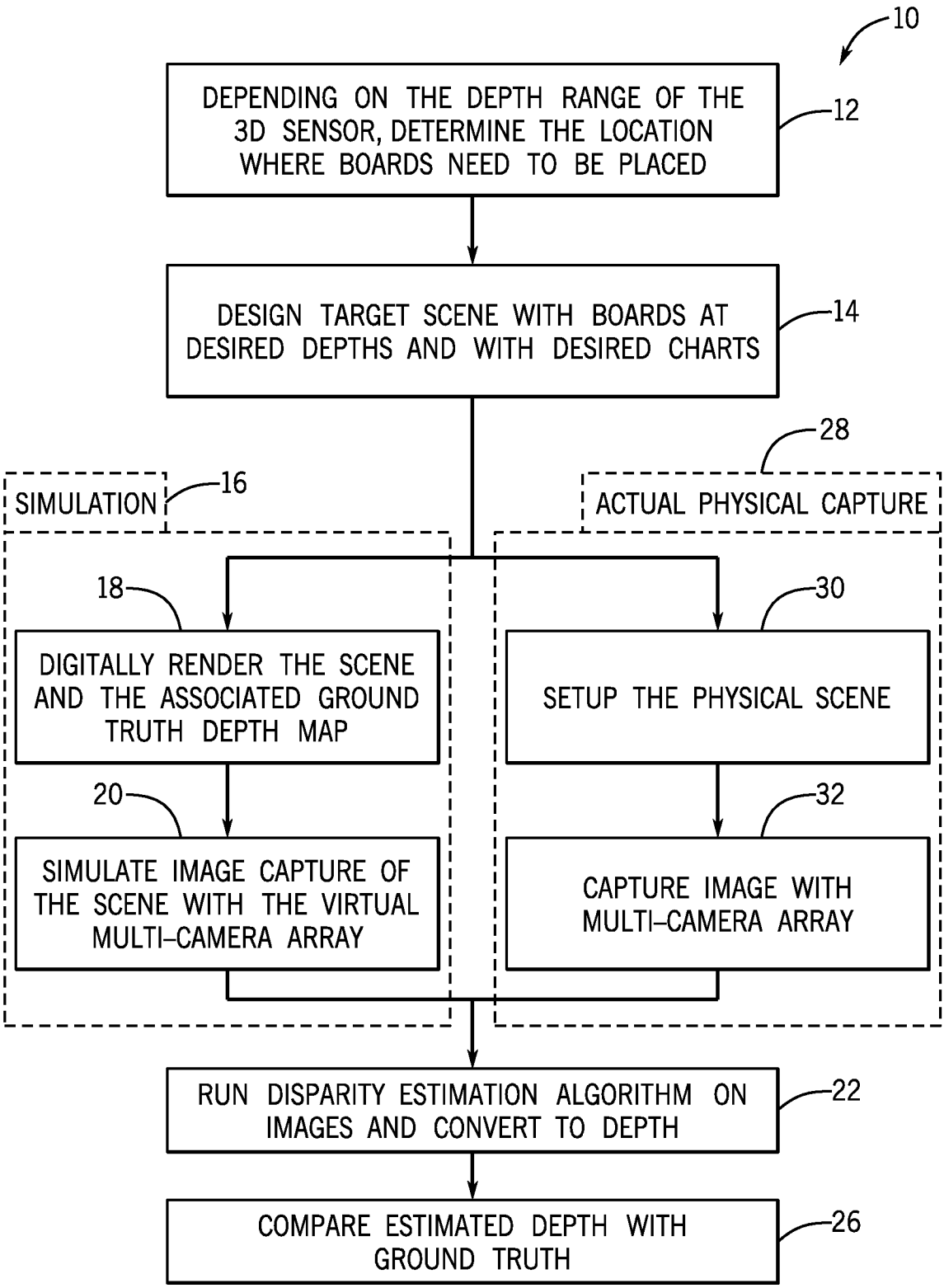

10
DEPENDING ON THE DEPTH RANGE OF THE 3D SENSOR, DETERMINE THE LOCATION WHERE BOARDS NEED TO BE PLACED
12
DESIGN TARGET SCENE WITH BOARDS AT DESIRED DEPTHS AND WITH DESIRED CHARTS
14
28
SIMULATION
16
ACTUAL PHYSICAL CAPTURE
18
DIGITALLY RENDER THE SCENE AND THE ASSOCIATED GROUND TRUTH DEPTH MAP
30
SETUP THE PHYSICAL SCENE
20
SIMULATE IMAGE CAPTURE OF THE SCENE WITH THE VIRTUAL MULTI-CAMERA ARRAY
32
CAPTURE IMAGE WITH MULTI-CAMERA ARRAY
RUN DISPARITY ESTIMATION ALGORITHM ON IMAGES AND CONVERT TO DEPTH
22
COMPARE ESTIMATED DEPTH WITH GROUND TRUTH
26

FIG. 5

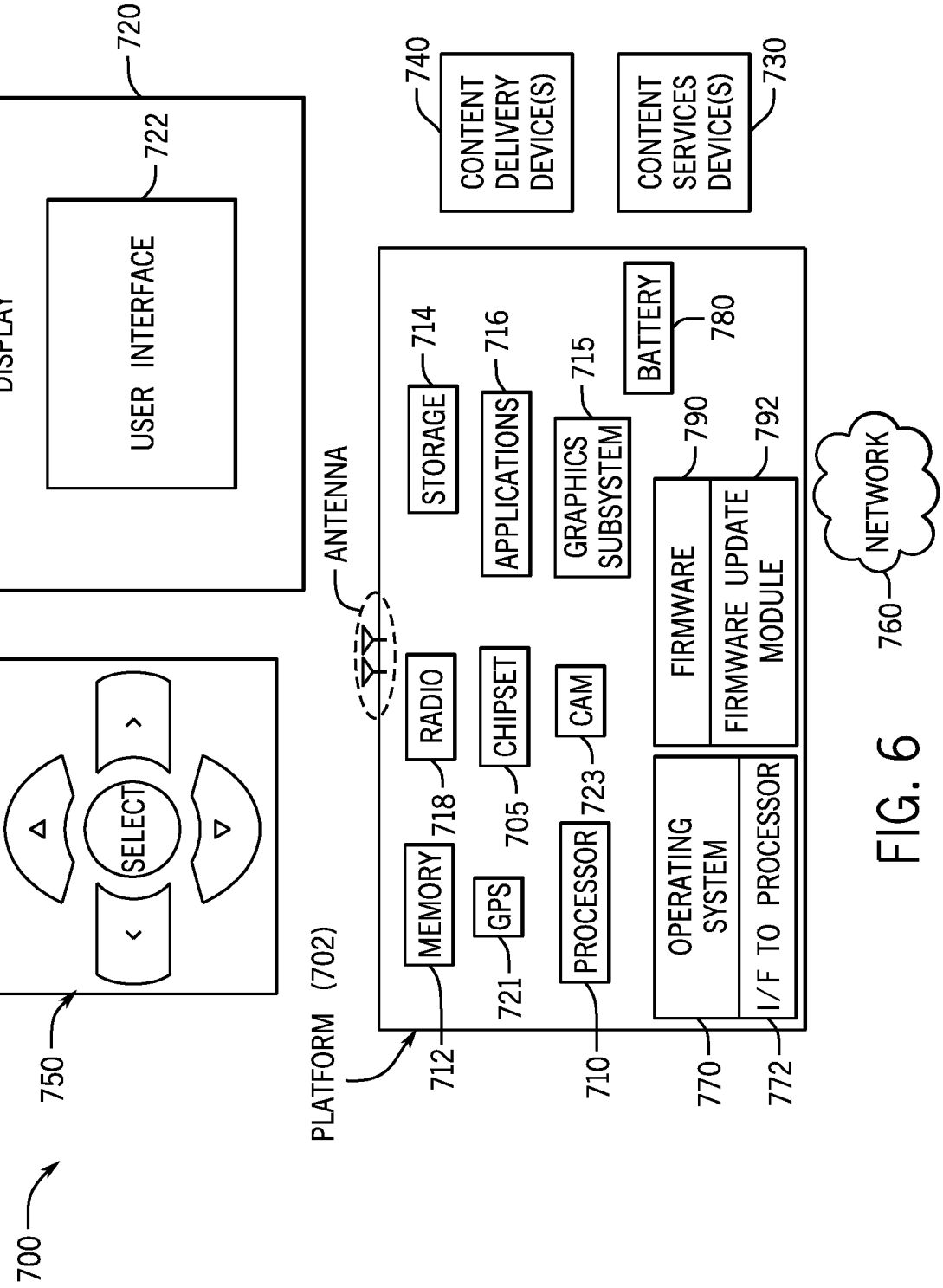

700
750
DISPLAY
720
USER INTERFACE
722
SELECT
PLATFORM (702)
712
ANTENNA
RADIO
MEMORY
718
GPS
721
CHIPSET
705
PROCESSOR
723
CAM
STORAGE
714
APPLICATIONS
716
GRAPHICS SUBSYSTEM
715
BATTERY
780
FIRMWARE
790
FIRMWARE UPDATE MODULE
792
OPERATING SYSTEM
770
I/F TO PROCESSOR
772
CONTENT DELIVERY DEVICE(S)
740
CONTENT SERVICES DEVICE(S)
730
NETWORK
760

FIG. 6

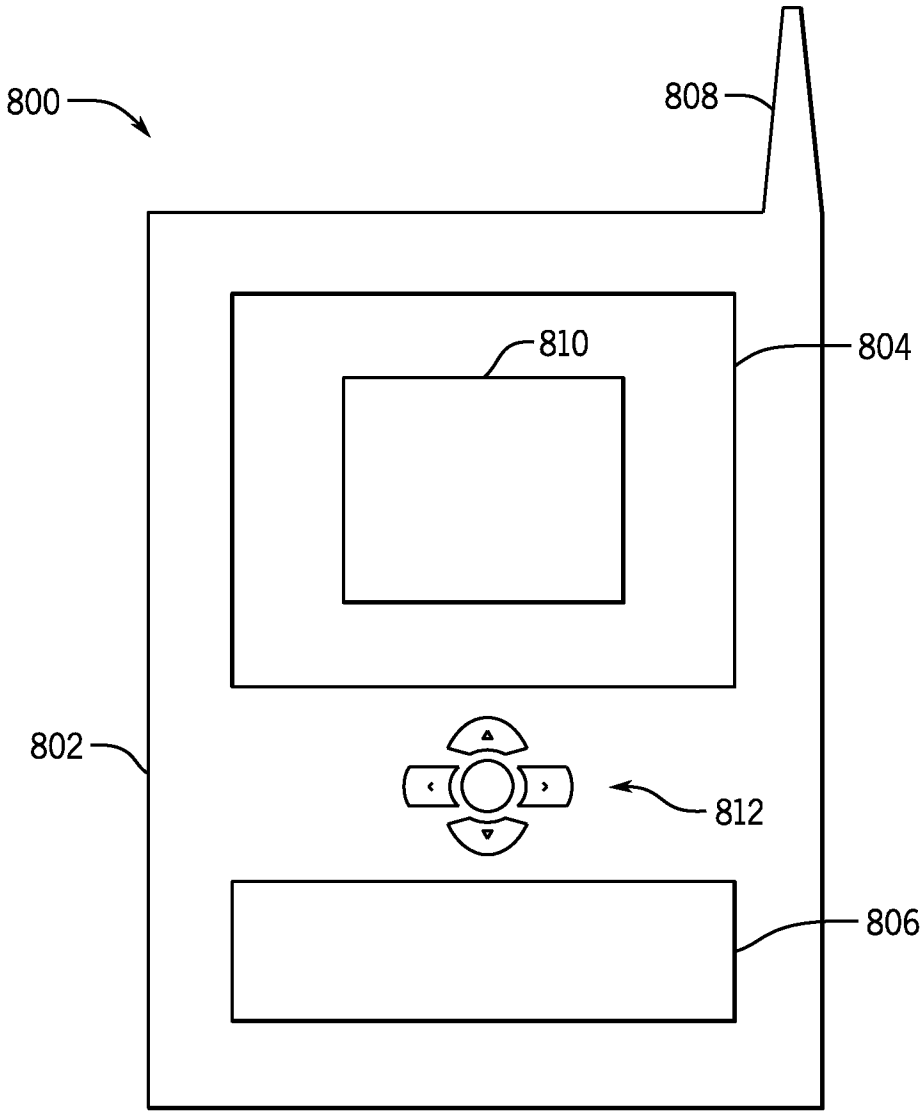


FIG. 7

INTERNATIONAL SEARCH REPORT

International application No.
PCT/US2015/056156

A. CLASSIFICATION OF SUBJECT MATTER

H04N 13/00(2006.01)i, H04N 5/232(2006.01)i

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

B. FIELDS SEARCHED

Minimum documentation searched (classification system followed by classification symbols)
H04N 13/00; G01M 11/02; G06T 15/00; G01B 11/00; G01C 3/06; G06T 7/00; H04N 5/232

Documentation searched other than minimum documentation to the extent that such documents are included in the fields searched
Korean utility models and applications for utility models
Japanese utility models and applications for utility models

Electronic data base consulted during the international search (name of data base and, where practicable, search terms used)
eKOMPASS(KIPO internal) & Keywords: camera, depth, estimation, Dead leaves, disparity

C. DOCUMENTS CONSIDERED TO BE RELEVANT

Category*	Citation of document, with indication, where appropriate, of the relevant passages	Relevant to claim No.
A	LEONIE KIRK et al., "Description of texture loss using the dead leaves target: Current issues and a new intrinsic approach", Proc. SPIE 9023, Digital Photography X, 90230C, 07 March 2014, Retrieved from the Internet: <URL: http://proceedings.spiedigitallibrary.org/proceeding.aspx?articleid=1844586> See page 1, line 27 - page 2, line 23.	1-23
A	US 2013-0251204 A1 (MICROSOFT CORPORATION) 26 September 2013 See paragraphs [0054]-[0055]; claim 1; and figure 1B.	1-23
A	JP 2005-207969 A (OLYMPUS CORP.) 04 August 2005 See paragraphs [0024]-[0026]; claim 1; and figures 1, 7.	1-23
A	US 2013-0201177 A1 (NGAI-MAN CHEUNG et al.) 08 August 2013 See paragraph [0068] claim 1; and figure 4.	1-23
A	JP 2009-210286 A (TOSHIBA TELI CORP.) 17 September 2009 See paragraphs [0031]-[0032]; claims 1-2; and figure 1.	1-23

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

☒ See patent family annex.

* Special categories of cited documents:
"A" document defining the general state of the art which is not considered to be of particular relevance
"E" earlier application or patent but published on or after the international filing date
"L" document which may throw doubts on priority claim(s) or which is cited to establish the publication date of another citation or other special reason (as specified)
"O" document referring to an oral disclosure, use, exhibition or other means
"P" document published prior to the international filing date but later than the priority date claimed
"T" later document published after the international filing date or priority date and not in conflict with the application but cited to understand the principle or theory underlying the invention
"X" document of particular relevance; the claimed invention cannot be considered novel or cannot be considered to involve an inventive step when the document is taken alone
"Y" document of particular relevance; the claimed invention cannot be considered to involve an inventive step when the document is combined with one or more other such documents,such combination being obvious to a person skilled in the art
"&" document member of the same patent family

Date of the actual completion of the international search	Date of mailing of the international search report
25 February 2016 (25.02.2016)	**25 February 2016 (25.02.2016)**

Name and mailing address of the ISA/KR
International Application Division
Korean Intellectual Property Office
189 Cheongsa-ro, Seo-gu, Daejeon, 35208, Republic of Korea
Facsimile No. +82-42-472-7140

Authorized officer
COMMISSIONER
Telephone No. +82-42-481-3578

INTERNATIONAL SEARCH REPORT

Information on patent family members

International application No.
PCT/US2015/056156

Patent document cited in search report	Publication date	Patent family member(s)	Publication date
US 2013-0251204 A1	26/09/2013	CN 102609954 A	25/07/2012
		US 2012-0159290 A1	21/06/2012
		US 8448056 B2	21/05/2013
		US 8775916 B2	08/07/2014
JP 2005-207969 A	04/08/2005	JP 4328224 B2	09/09/2009
		US 2005-0167499 A1	04/08/2005
		US 8259180 B2	04/09/2012
US 2013-0201177 A1	08/08/2013	US 8692840 B2	08/04/2014
JP 2009-210286 A	17/09/2009	JP 4839330 B2	21/12/2011