Innovation, Sciences et Développement économique Canada
Office de la Propriété Intellectuelle du Canada

Innovation, Science and Economic Development Canada
Canadian Intellectual Property Office

CA 3157194 C 2023/08/29

(11)(21) 3 157 194

(12) BREVET CANADIEN
CANADIAN PATENT

(13) C

(86) Date de dépôt PCT/PCT Filing Date: 2020/10/07
(87) Date publication PCT/PCT Publication Date: 2021/04/15
(45) Date de délivrance/Issue Date: 2023/08/29
(85) Entrée phase nationale/National Entry: 2022/04/06
(86) N° demande PCT/PCT Application No.: US 2020/054641
(87) N° publication PCT/PCT Publication No.: 2021/071992
(30) Priorités/Priorities: 2019/10/07 (US62/911,952); 2019/11/30 (US62/942,113); 2020/03/29 (US63/001,445)

(51) Cl.Int./Int.Cl. *G06T 5/00* (2006.01), *G06T 7/00* (2017.01)

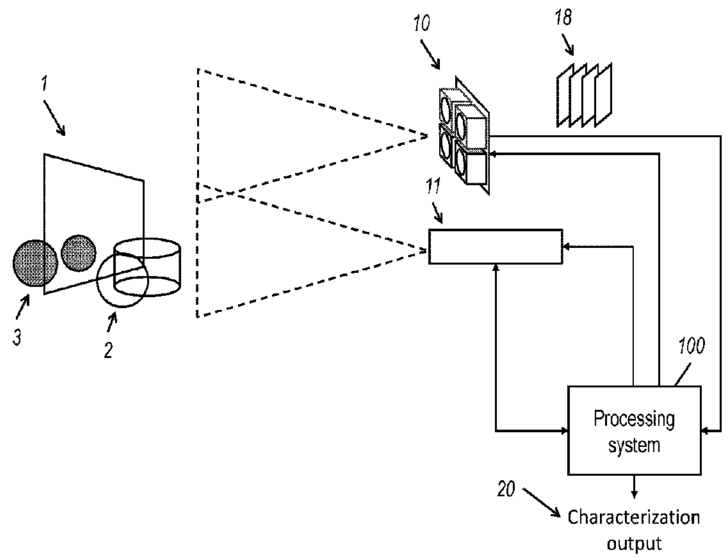
(72) Inventeurs/Inventors:
KADAMBI, ACHUTA, US;
RASKAR, RAMESH, US;
VENKATARAMAN, KARTIK, US;
RAO, SUPREETH KRISHNA, US;
KALRA, AGASTYA, US

(73) Propriétaire/Owner:
BOSTON POLARIMETRICS, INC., US

(74) Agent: AIRD & MCBURNEY LP

(54) Titre : SYSTEMES ET PROCEDES D'AUGMENTATION DE SYSTEMES DE CAPTEUR ET DE SYSTEMES D'IMAGERIE AVEC POLARISATION

(54) Title: SYSTEMS AND METHODS FOR AUGMENTATION OF SENSOR SYSTEMS AND IMAGING SYSTEMS WITH POLARIZATION



(57) Abrégé/Abstract:
A multi-modal sensor system includes: an underlying sensor system; a polarization camera system configured to capture polarization raw frames corresponding to a plurality of different polarization states; and a processing system including a processor and memory, the processing system being configured to control the underlying sensor system and the polarization camera system, the memory storing instructions that, when executed by the processor, cause the processor to: control the underlying sensor system to perform sensing on a scene and the polarization camera system to capture a plurality of polarization raw frames of the scene; extract first tensors in polarization representation spaces based on the plurality of polarization raw frames; and compute a characterization output based on an output of the underlying sensor system and the first tensors in polarization representation spaces.

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

(19) World Intellectual Property Organization
International Bureau

(43) International Publication Date
15 April 2021 (15.04.2021)

WIPO | PCT

(10) International Publication Number
WO 2021/071992 A1

(51) International Patent Classification:
G06T 5/00 (2006.01) *G06T 7/00* (2017.01)

(21) International Application Number:
PCT/US2020/054641

(22) International Filing Date:
07 October 2020 (07.10.2020)

(25) Filing Language: English

(26) Publication Language: English

(30) Priority Data:

62/911,952	07 October 2019 (07.10.2019)	US
62/942,113	30 November 2019 (30.11.2019)	US
63/001,445	29 March 2020 (29.03.2020)	US

(71) Applicant: BOSTON POLARIMETRICS, INC. [US/US]; 380 Portage Avenue, Palo Alto, CA 94306 (US).

(72) Inventors: **KADAMBI, Achuta**; c/o BOSTON POLARIMETRICS, INC., 380 Portage Avenue, Palo Alto, CA 94306 (US). **RASKAR, Ramesh**; c/o BOSTON POLARIMETRICS, INC., 380 Portage Avenue, Palo Alto, CA 94306 (US). **VENKATARAMAN, Kartik**; c/o BOSTON POLARIMETRICS, INC., 380 Portage Avenue, Palo Alto, CA 94306 (US). **RAO, Supreeth, Krishna**; c/o BOSTON POLARIMETRICS, INC., 380 Portage Avenue, Palo Alto, CA 94306 (US). **KALRA, Agastya**; c/o BOSTON POLARIMETRICS, INC., 380 Portage Avenue, Palo Alto, CA 94306 (US).

(74) Agent: **LEE, Shaun, P.** et al.; LEWIS ROCA ROTHGERBER CHRISTIE LLP, P.O. Box 29001, Glendale, CA 91209 (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,

(54) **Title:** SYSTEMS AND METHODS FOR AUGMENTATION OF SENSOR SYSTEMS AND IMAGING SYSTEMS WITH POLARIZATION

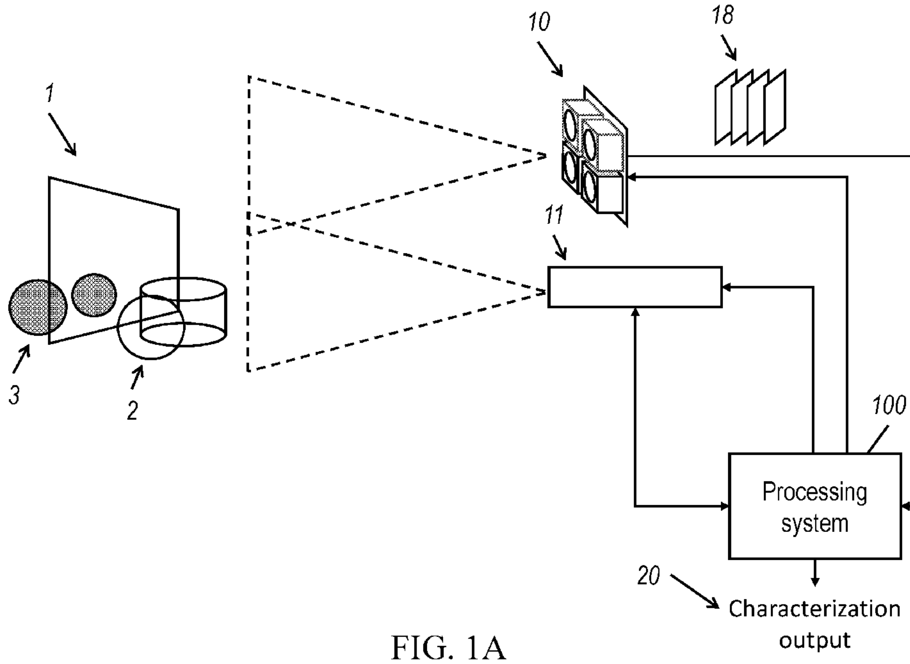


FIG. 1A

(57) **Abstract:** A multi-modal sensor system includes: an underlying sensor system; a polarization camera system configured to capture polarization raw frames corresponding to a plurality of different polarization states; and a processing system including a processor and memory, the processing system being configured to control the underlying sensor system and the polarization camera system, the memory storing instructions that, when executed by the processor, cause the processor to: control the underlying sensor system to perform sensing on a scene and the polarization camera system to capture a plurality of polarization raw frames of the scene; extract first tensors in polarization representation spaces based on the plurality of polarization raw frames; and compute a characterization output based on an output of the underlying sensor system and the first tensors in polarization representation spaces.

WO 2021/071992 A1

[Continued on next page]

WO 2021/071992 A1

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, IT, JO, 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, ST, SV, SY, TH, TJ, TM, TN, TR, TT, TZ, UA, UG, US, UZ, VC, VN, WS, 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).

Published:

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

— *before the expiration of the time limit for amending the claims and to be republished in the event of receipt of amendments (Rule 48.2(h))*

SYSTEMS AND METHODS FOR AUGMENTATION OF SENSOR SYSTEMS AND IMAGING SYSTEMS WITH POLARIZATION

CROSS-REFERENCE TO RELATED APPLICATION(S)

[0001] This application claims priority to and the benefit of U.S. Provisional Patent Application No. 62/911,952, filed in the United States Patent and Trademark Office on October 7, 2019, U.S. Provisional Patent Application No. 62/942,113, filed in the United States Patent and Trademark Office on November 30, 2019, and U.S. Provisional Patent Application No. 63/001,445, filed in the United States Patent and Trademark Office on March 29, 2020.

FIELD

[0002] Aspects of embodiments of the present disclosure relate to the field of sensor systems, including sensor systems augmented with polarization.

BACKGROUND

[0003] Sensor systems and imaging systems such as radar, lidar, cameras (e.g., visible light and/or infrared), and the like detect objects and features in the environment through the interactions of electromagnetic radiation with the environment. For example, camera systems and lidar systems detect light reflected off of objects in a scene or in an environment. Likewise, radar systems transmit lower frequency electromagnetic waves (e.g., radio frequency or microwave frequency) and determine properties of the objects based on the reflections of those signals. Other sensor systems may use other forms of radiation, such as pressure waves or sound waves in the case of ultrasound imaging.

SUMMARY

[0004] Aspects of embodiments of the present disclosure relate to systems and methods for augmentation of sensor systems and imaging systems using polarization. According to some aspects of embodiments of the present disclosure, sensors configured to detect the polarization of received electromagnetic radiation is used to augment the performance or behavior of other imaging modalities, such as cameras configured to detect the intensity of light without regard to the polarization of the light. In some aspects of embodiments of the present disclosure, sensors configured to detect the polarization of received electromagnetic radiation are used to form images that would otherwise be formed using comparative imaging systems

Date Reçue/Date Received 2022-07-05

such as digital cameras. Some aspects of embodiments of the present disclosure relate to camera systems configured to detect the polarization of light.

[0005] According to one embodiment of the present disclosure, a multi-modal sensor system includes: an underlying sensor system; a polarization camera system configured to capture polarization raw frames corresponding to a plurality of different polarization states; and a processing system including a processor and memory, the processing system being configured to control the underlying sensor system and the polarization camera system, the memory storing instructions that, when executed by the processor, cause the processor to: control the underlying sensor system to perform sensing on a scene and the polarization camera system to capture a plurality of polarization raw frames of the scene; extract first tensors in polarization representation spaces based on the plurality of polarization raw frames; and compute a characterization output based on an output of the underlying sensor system and the first tensors in polarization representation spaces.

[0006] The polarization camera system may include a polarization camera module including: a first polarization camera including a first polarizing filter at a first polarization orientation, the first polarization camera having a first optical axis; a second polarization camera including a second polarizing filter at a second polarization orientation, the second polarization camera having a second optical axis substantially parallel to the first optical axis; and a third polarization camera including a third polarizing filter at a third polarization orientation, the third polarization camera having a third optical axis substantially parallel to the first optical axis.

[0007] The polarization camera module may further include a fourth polarization camera including a fourth polarizing filter at a fourth polarization orientation, the fourth polarization camera having a fourth optical axis substantially parallel to the first optical axis.

[0008] The first tensors may include a degree of linear polarization (DOLP) and an angle of linear polarization (AOLP), and the memory may further store instructions that, when executed by the processor, cause the processor to compute the DOLP and the AOLP based on polarization raw frames captured by the first polarization camera, the second polarization camera, and the third polarization camera, the instructions including instructions to: initialize an estimated DOLP and an estimated AOLP based on stakes vectors; estimate a scene geometry based on parallax shifts in the polarization raw frames to generate a coarse model; and iteratively: refine the coarse model based on the estimated DOLP and the estimated AOLP to generate an estimated geometry; and update the estimated DOLP and the estimated AOLP based on the estimated geometry, until a change in the estimated

DOLP and a change in the estimated AOLP are both less than corresponding threshold values.

[0009] The polarization camera system may include a stereo polarization camera system including: a first polarization camera module having a first optical axis, the first polarization camera module being configured to capture a first plurality of polarization raw frames corresponding to a first plurality of different polarization states; and a second polarization camera module having a second optical axis and spaced apart from the first polarization camera module along a baseline, the second polarization camera module being configured to capture a second plurality of polarization raw frames corresponding to a second plurality of different polarization states, the first optical axis being substantially parallel to the second optical axis.

[0010] The first polarization camera module may include a first plurality of color filters configured to transmit light in three or more different first color spectra, and the second polarization camera module may include a second plurality of color filters configured to transmit light in three or more different second color spectra, wherein the three or more second color spectra may be different from the three or more first color spectra.

[0011] The memory may further store instructions that, when executed by the processor, cause the processor to: control the stereo polarization camera system to capture multi-spectral stereo polarization imaging data in the first color spectra and in the second color spectra; and extract first tensors in polarization representation spaces from the multi-spectral stereo polarization imaging data.

[0012] The underlying sensor system may include an active scanning system including an active emitter and a detector. The active scanning system may include a radar system. The active scanning system may include a lidar system. The active scanning system may include an active stereo depth camera system.

[0013] The multi-modal sensor system may be mounted on a vehicle.

[0014] The memory may further store instructions that, when executed by the processor of the processing system, cause the processor to: compute a sparse point cloud based on the output of the underlying sensor system; compute surface normals from the polarization raw frames; compute a 3-D surface based on the surface normals; and correct the 3-D surface based on the sparse point cloud to compute a 3-D model of the scene.

[0015] The memory may further store instructions that, when executed by the processor of the processing system, cause the processor to: compute a segmentation map based on the first tensors in the polarization representation spaces, the segmentation map identifying one or more regions of interest in the

1

scene; steer the active emitter to emit beams toward the one or more regions of interest; and detect a reflection of the beams emitted by the active emitter using the detector of the active scanning system.

[0016] The memory may further store instructions that cause the processor to implement a convolutional neural network trained to compute a segmentation map based on the first tensors.

[0017] The underlying sensor system may include a color camera system.

[0018] The color camera system may be a digital single lens reflex camera or a video camera.

[0019] The output of the color camera of the underlying sensor system may include a color image, and the memory may further store instructions that, when executed by the processor, cause the processor to: compute a plurality of surface normals of the scene based on the first tensors; and store the computed surface normals of the scene in a same file as the color image captured by the color camera.

[0020] According to one embodiment of the present disclosure, a polarization camera system includes: a polarization camera configured to capture polarization raw frames of a scene, the polarization raw frames corresponding to a plurality of different polarization states; and a processing system having a processor and memory, the processing system being configured to control the polarization camera, the memory storing instructions that, when executed by the processor, cause the processor to: control the polarization camera to capture a plurality of polarization raw frames; and synthesize a high dynamic range (HDR) image based on the polarization raw frames.

[0021] Each of the polarization raw frames may be captured based on a same set of exposure settings, at least one of the polarization raw frames may include saturated pixels in a portion of the polarization raw frame due to specular reflection from a surface in the scene, the memory may further stores instructions that, when executed by the processor, cause the processor to synthesize the HDR image without saturated pixels in a portion of the HDR image corresponding to the portion of the polarization raw frame based on data from corresponding portions of other ones of the polarization raw frames.

BRIEF DESCRIPTION OF THE DRAWINGS

[0022] The accompanying drawings, together with the specification, illustrate exemplary embodiments of the present invention, and, together with the description, serve to explain the principles of the present invention.

1

[0023] FIG. 1A is a schematic block diagram of an imaging system augmented with a polarization camera system according to embodiments of the present disclosure.

[0024] FIG. 1B is a schematic block diagram of a system using a polarization camera according to one embodiment of the present invention.

[0025] FIG. 1C is a perspective view of a polarization camera module according to one embodiment of the present disclosure.

[0026] FIG. 1D is a cross sectional view of a portion of a polarization camera module according to one embodiment of the present disclosure.

[0027] FIG. 1E is a perspective view of a stereo polarization camera system according to one embodiment of the present disclosure.

[0028] FIGS. 2A, 2B, 2C, and 2D provide background for illustrating the segmentation maps computed by a comparative approach and semantic segmentation or instance segmentation based on polarization raw frames according to aspects of embodiments of the present disclosure.

[0029] FIG. 3 is a high-level depiction of the interaction of light with transparent objects and non-transparent (e.g., diffuse and/or reflective) objects.

[0030] FIG. 4 is a flowchart of a method for estimating polarization cues under parallax ambiguities according to one embodiment of the present disclosure.

[0031] FIG. 5A is a perspective view of a multi-spectral stereo polarization camera system according to one embodiment of the present disclosure.

[0032] FIG. 5B is a view of a multi-spectral stereo polarization camera system according to one embodiment of the present disclosure, along a direction parallel to the optical axis of the multi-spectral stereo polarization camera system.

[0033] FIG. 5C depicts cut-away side views of example individual polarization cameras of a multi-spectral stereo polarization camera system according to one embodiment of the present disclosure.

[0034] FIG. 6A is a block diagram of processing circuit 100 for computing surface characterization outputs based on polarization data according to one embodiment of the present invention.

[0035] FIG. 6B is a flowchart of a method 600 for performing surface characterization based on input images to compute a surface characterization output according to one embodiment of the present invention.

[0036] FIG. 7A is a block diagram of a feature extractor 700 according to one embodiment of the present invention.

[0037] FIG. 7B is a flowchart depicting a method according to one embodiment of the present invention for extracting features from polarization raw frames.

1

[0038] FIG. 8A is an illustration of a Greek bust statue being scanned by an exemplary implementation of the imaging setup proposed in this invention.

[0039] FIG. 8B is a flowchart of a method for 3-D surface reconstruction using polarization according to one embodiment of the present disclosure.

[0040] FIG. 9A is an illustration of a flat surface of refractive index n, being scanned by an exemplary implementation of the imaging setup according to one embodiment of the present invention.

[0041] FIG. 9B is a flowchart of a method for 3-D surface reconstruction of flat or geometrically simple surfaces using polarization according to one embodiment of the present disclosure.

[0042] FIG. 10A is a schematic diagram of a system in which an active scanning system is augmented with a polarization camera system according to one embodiment of the present disclosure.

[0043] FIG. 10B is a flowchart depicting a method for fusing the 3-D model captured by an active 3-D scanning system with surface normals captured by a polarization camera according to one embodiment of the present disclosure.

[0044] FIG. 10C is a flowchart illustrating a method for augmenting an active sensing system using polarization according to one embodiment of the present disclosure.

[0045] FIG. 11 is a flowchart illustrating a method for synthesizing a high dynamic range (HDR) image from polarization raw frames.

DETAILED DESCRIPTION

[0046] In the following detailed description, only certain exemplary embodiments of the present invention are shown and described, by way of illustration. As those skilled in the art would recognize, the invention may be embodied in many different forms and should not be construed as being limited to the embodiments set forth herein.

[0047] Optically challenging objects such as transparent objects occur in many real-world applications of computer vision or machine vision systems, including automation and analysis for manufacturing, life sciences, and automotive industries. For example, in manufacturing, computer vision systems may be used to automate: sorting, selection, and placement of parts; verification of placement of components during manufacturing; and final inspection and defect detection. As additional examples, in life sciences, computer vision systems may be used to automate: measurement of reagents; preparation of samples; reading outputs of instruments; characterization of samples; and picking and placing container samples. Further examples in automotive industries include detecting transparent objects in street scenes for assisting drivers or for operating self-driving vehicles. Additional examples may include assistive technologies, such as self-navigating wheelchairs capable of detecting glass doors and other transparent barriers and devices for assisting people with vision impairment that are capable of detecting transparent drinking glasses and to distinguish between real objects and print-out spoofs.

[0048] In contrast to opaque objects, transparent objects lack texture of their own (e.g., surface color information, as the term is used in the field of computer graphics, such as in "texture mapping"). As a result, comparative systems generally fail to correctly identify instances of transparent objects that are present in scenes captured using standard imaging systems (e.g., cameras configured to capture monochrome intensity images or color intensity images such as red, green, and blue or RGB images). This may be because the transparent objects do not have a consistent texture (e.g., surface color) for the algorithms to latch on to or to learn to detect (e.g., during the training process of a machine learning algorithm). Similar issues may arise from partially transparent or translucent objects, as well as some types of reflective objects (e.g., shiny metal) and very dark objects (e.g., matte black objects).

[0049] Accordingly, aspects of embodiments of the present disclosure relate to using polarization imaging to provide additional information for augmenting sensor systems to detect transparent objects and other optically challenging objects and features in scenes. In addition, aspects of embodiments of the present disclosure also apply to detecting other optically challenging objects such as transparent, translucent, and reflective objects as well as dark objects.

[0050] As used herein, the term "optically challenging" refers to objects made of materials that satisfy one or more of the following four characteristics at a sufficient threshold level or degree: non-Lambertian (e.g., not matte); translucent; multipath inducing; and/or non-reflective. In some circumstances an object exhibiting only one of the four characteristics may be optically challenging to detect. In addition, objects or materials may exhibit multiple characteristics simultaneously. For example, a translucent object may have a surface reflection and background reflection, so it is challenging both because of translucency and the multipath. In some circumstances, an object may exhibit one or more of the four characteristics listed above, yet may not be optically challenging to detect because these conditions are not exhibited at a level or degree that would pose a problem to a comparative computer vision systems. For example, an object may be translucent, but still exhibit enough surface texture to be detectable and segmented from other instances of objects in a scene. As another example, a surface must be sufficiently non-Lambertian to introduce problems to other vision systems. In some embodiments, the degree or level to which an object is optically challenging is quantified using the full-width half max (FWHM) of the specular lobe of the bidirectional reflectance distribution function (BRDF) of the object. If this FWHM is below a threshold, the material is considered optically challenging.

[0051] FIG. 1A is a schematic block diagram of an imaging system augmented with a polarization camera system according to embodiments of the present disclosure. FIG. 1B is a schematic block diagram of a system using a polarization camera according to one embodiment of the present invention. In the arrangement shown in FIG. 1A and FIG. 1B, a scene 1 includes transparent objects 2 (e.g., depicted as a ball such as a glass marble, a cylinder such as a drinking glass or tumbler, and a plane such as a pane of transparent acrylic) that are placed in front of opaque matte objects 3 (e.g., a baseball and a tennis ball). A polarization camera 10 has a lens 12 with a field of view, where the lens 12 and the camera 10 are oriented such that the field of view encompasses the scene 1. The lens 12 is configured to direct light (e.g., focus light) from the scene 1 onto a light sensitive medium such as an image sensor 14 (e.g., a complementary metal oxide semiconductor (CMOS) image sensor or charge-coupled device (CCD) image sensor). As shown in FIG. 1A, the polarization camera 10 may be used to augment the sensor data captured by another imaging system 11.

[0052] The polarization camera 10 further includes a polarizer or polarizing filter or polarization mask 16 placed in the optical path between the scene 1 and the image sensor 14. According to various embodiments of the present disclosure, the polarizer or polarization mask 16 is configured to enable the polarization camera 10 to capture images of the scene 1 with the polarizer set at various specified angles (e.g., at 45° rotations or at 60° rotations or at non-uniformly spaced rotations).

[0053] As one example, FIG. 1B depicts an embodiment where the polarization mask 16 is a polarization mosaic aligned with the pixel grid of the image sensor 14 in a manner similar to a red-green-blue (RGB) color filter (e.g., a Bayer filter) of a color camera. In a manner similar to how a color filter mosaic filters incoming light based on wavelength such that each pixel in the image sensor 14 receives light in a particular portion of the spectrum (e.g., red, green, or blue) in accordance with the pattern of color filters of the mosaic, a polarization mask 16 using a polarization mosaic filters light based on linear polarization such that different pixels receive light at different angles of linear polarization (e.g., at 0°, 45°, 90°, and 135°, or at 0°, 60° degrees, and 120°). Accordingly, the polarization camera 10 using a polarization mask 16 such as that shown in FIG. 1B is capable of concurrently or simultaneously

1

capturing light at four different linear polarizations. One example of a polarization camera is the Blackfly® S Polarization Camera produced by FLIR® Systems, Inc. of Wilsonville, Oregon.

[0054] While the above description relates to some possible implementations of a polarization camera using a polarization mosaic, embodiments of the present disclosure are not limited thereto and encompass other types of polarization cameras that are capable of capturing images at multiple different polarizations. For example, the polarization mask 16 may have fewer than four polarizations or more than four different polarizations, or may have polarizations at different angles than those stated above (e.g., at angles of polarization of: 0°, 60°, and 120° or at angles of polarization of 0°, 30°, 60°, 90°, 120°, and 150°). As another example, the polarization mask 16 may be implemented using an electronically controlled polarization mask, such as an electro-optic modulator (e.g., may include a liquid crystal layer), where the polarization angles of the individual pixels of the mask may be independently controlled, such that different portions of the image sensor 14 receive light having different polarizations. As another example, the electro-optic modulator may be configured to transmit light of different linear polarizations when capturing different frames, e.g., so that the camera captures images with the entirety of the polarization mask set to, sequentially, to different linear polarizer angles (e.g., sequentially set to: 0 degrees; 45 degrees; 90 degrees; or 135 degrees). As another example, the polarization mask 16 may include a polarizing filter that rotates mechanically, such that different polarization raw frames are captured by the polarization camera 10 with the polarizing filter mechanically rotated with respect to the lens 12 to transmit light at different angles of polarization to image sensor 14. Furthermore, while the above examples relate to the use of a linear polarizing filter, embodiments of the present disclosure are not limited thereto and also include the use of polarization cameras that include circular polarizing filters (e.g., linear polarizing filters with a quarter wave plate). Accordingly, in various embodiments of the present disclosure, a polarization camera uses a polarizing filter to capture multiple polarization raw frames at different polarizations of light, such as different linear polarization angles and different circular polarizations (e.g., handedness).

[0055] As a result, the polarization camera 10 captures multiple input images 18 (or polarization raw frames) of the scene including the surface under inspection 2 of the object under inspection 1. In some embodiments, each of the polarization raw frames 18 corresponds to an image taken behind a polarization filter or polarizer at a different angle of polarization ϕ_{pol} (e.g., 0 degrees, 45 degrees, 90 degrees, or 135 degrees). Each of the polarization raw frames 18 is captured from substantially the same pose with respect to the scene 1 (e.g., the images captured with the polarization filter at 0 degrees, 45 degrees, 90 degrees, or 135 degrees are all captured by a same polarization camera 10 located at a same location and orientation), as opposed to capturing the polarization raw frames from disparate locations and orientations with respect to the scene. The polarization camera 10 may be configured to detect light in a variety of different portions of the electromagnetic spectrum, such as the human-visible portion of the electromagnetic spectrum, red, green, and blue portions of the human-visible spectrum, as well as invisible portions of the electromagnetic spectrum such as infrared and ultraviolet.

[0056] In some embodiments of the present disclosure, such as some of the embodiments described above, the different polarization raw frames are captured by a same polarization camera 10 and therefore may be captured from substantially the same pose (e.g., position and orientation) with respect to the scene 1. However, embodiments of the present disclosure are not limited thereto. For example, a polarization camera 10 may move with respect to the scene 1 between different polarization raw frames (e.g., when different raw polarization raw frames corresponding to different angles of polarization are captured at different times, such as in the case of a mechanically rotating polarizing filter), either because the polarization camera 10 has moved or because objects 3 have moved (e.g., if the object is on a moving conveyor system). In some embodiments, different polarization cameras capture images of the object at different times, but from substantially the same pose with respect to the object (e.g., different cameras capturing images of the same surface of the object at different points in the conveyor system). Accordingly, in some embodiments of the present disclosure different polarization raw frames are captured with the polarization camera 10 at different poses or the same relative pose with respect to the objects 2 and 3 being imaged in the scene 1.

[0057] The polarization raw frames 18 are supplied to a processing circuit 100, described in more detail below, which computes a processing output 20 based on the polarization raw frames 18. In the embodiment shown in FIG. 1B, the processing output 20 is an instance segmentation map identifying instances of different objects 2 and 3 that are present in the scene 1, but embodiments of the present disclosure are not limited thereto. Specific examples of processing outputs 20 that are computed based on polarization raw frames will be described in more detail below. In the embodiment shown in FIG. 1A, the processing circuit 100 is configured to control both the polarization camera 10 and the additional imaging system 11.

[0058] FIG. 1C is a perspective view of a polarization camera module according to one embodiment of the present disclosure. FIG. 1D is a cross sectional view of a

1

portion of a polarization camera module according to one embodiment of the present disclosure. Some aspects of embodiments of the present disclosure relate to a polarization camera module in which multiple polarization cameras (e.g., multiple cameras, where each camera has a polarizing filter in its optical path) are arranged adjacent to one another and in an array and may be controlled to capture images in a group (e.g., a single trigger may be used to control all of the cameras in the system to capture images concurrently or substantially simultaneously). The polarizing filters in the optical paths of each of the cameras in the array cause differently polarized light to reach the image sensors of the cameras. The individual polarization cameras in the camera system have optical axes that are substantially perpendicular to one another, are placed adjacent to one another, and have substantially the same field of view, such that the cameras in the camera system capture substantially the same view of a scene 1, but with different polarizations. In some embodiments, the individual polarization cameras are arranged such that parallax shift between cameras is substantially negligible based on the designed operating distance of the camera system to objects in the scene, where larger spacings between the cameras may be tolerated if the designed operating distance is large. In some embodiments of the present disclosure, the polarization camera module includes at least three polarization cameras, each having a polarizing filter with a different polarization state (e.g., each at a different angle of linear polarization, such as 0°, 60°, and 120°).

[0059] For example, in the embodiment of the polarization camera module 10′ shown in FIG. 1C, four cameras 10A′, 10B′, 10C′, and 10D′ are arranged in a 2×2 grid to form a camera array, where the four cameras have substantially parallel optical axes. The four cameras may be controlled together such that they capture images substantially simultaneously and using the same exposure settings (e.g., same aperture, length of exposure, and gain or "ISO" settings). In various embodiments of the present disclosure, each of the separate cameras 10A′, 10B′, 10C′, and 10D′ includes a different polarizing filter.

[0060] FIG. 1D shows a cross sectional view of two of the polarization cameras 10A′ and 10B′ shown in FIG. 1C. As seen in FIG. 1D, each a polarization camera (10A′ and 10B′) system includes a corresponding lens, a corresponding image sensor, and a corresponding polarizing filter. In particular, polarization camera 10A′ includes lens 12A′, image sensor 14A′, and polarizing filter 16A′. Likewise, polarization camera 10B′ includes lens 12B′, image sensor 14B′, and polarizing filter 16B′. In some embodiments of the present disclosure, the image sensors four cameras 10A′, 10B′, 10C′, and 10D′ are monolithically formed on a same semiconductor die, and the four cameras are located in a same housing with separate apertures for the lenses 12 corresponding to the different image sensors. Similarly, the polarizing filters 16 may correspond to different portions of a single physical layer that has different polarizing filters (e.g., different linear polarizing angles) in different regions of the layer (corresponding to the different cameras).

[0061] In some embodiments of the present disclosure, each of the cameras in the camera system 10′ has a corresponding polarizing filter that is configured to filter differently polarized light. For example, in the embodiment shown in FIGS. 1C and 1D, polarizing filter 16A′ of camera 10A′ may be a linear polarizing filter oriented at an angle of 0°, polarizing filter 16B′ of camera 10B′ may be a linear polarizing filter oriented at an angle of 45°, polarizing filter 16C′ of camera 10C′ may be a linear polarizing filter oriented at an angle of 90°, and polarizing filter 16D′ of camera 10D′ may be a linear polarizing filter oriented at an angle of 135°. In some embodiments, one or more of the cameras may include a circular polarizer. In some embodiments of the present disclosure, the camera system 10′ includes polarizing filters configured to filter light in at least two different polarizations. In some embodiments of the present disclosure, the camera system 10′ includes polarizing filters configured to filter light in at least three different polarizations. In the embodiment shown in FIG. 1D, the polarizing filter 16 is located behind the lens 12 (e.g., between the lens 12 and the image sensor 14), but embodiments of the present disclosure are not limited thereto. In some embodiments, the polarizing filter is located in front of the lens 12.

[0062] In some embodiments, the various individual cameras of the camera array are registered with one another by determining their relative poses (or relative positions and orientations) by capturing multiple images of a calibration target, such as a checkerboard pattern, an ArUco target (see, e.g., Garrido-Jurado, Sergio, et al. "Automatic generation and detection of highly reliable fiducial markers under occlusion." *Pattern Recognition* 47.6 (2014): 2280-2292.) or a ChArUco target (see, e.g., An, Gwon Hwan, et al. "Charuco board-based omnidirectional camera calibration method." *Electronics* 7.12 (2018): 421.). In particular, the process of calibrating the targets may include computing intrinsic matrices characterizing the internal parameters of each camera (e.g., matrices characterizing the focal length, image sensor format, and principal point of the camera) and extrinsic matrices characterizing the pose of each camera with respect to world coordinates (e.g., matrices for performing transformations between camera coordinate space and world or scene coordinate space).

[0063] While not shown in FIG. 1D, in some embodiments of the present disclosure, each polarization camera may also include a color filter having in a mosaic pattern such as a Bayer filter, such that individual pixels of the image sensors 14 receive light corresponding to, for example, red (R), green (G), and blue (B) portions of the spectrum, such that each camera captures light in a visible portion of the electromagnetic spectrum in accordance with a mosaic pattern. In some embodiments, a demosaicing process is used to compute separate red, green, and blue channels from the raw data. In some embodiments of the present disclosure, each polarization camera may be used without a color filter or with filters used to transmit or selectively transmit various other portions of the electromagnetic spectrum, such as infrared light.

[0064] FIG. 1E is a perspective view of a stereo polarization camera system according to one embodiment of the present disclosure. In some applications, stereo vision techniques are used to capture multiple images of scene from different perspectives. As noted above, in some embodiments of the present disclosure, individual polarization cameras within a camera system are placed adjacent to one another such that parallax shifts between the cameras is substantially negligible based on the designed operating distance of the camera system to the subjects being imaged. In stereo polarization camera systems, some of the individual polarization cameras are spaced apart such that parallax shifts are significant and detectable for objects in the designed operating distance of the camera system. This enables the distances to various surfaces in a scene (the “depth”) to be detected in accordance with a magnitude of a parallax shift (e.g., larger parallax shifts in the locations of corresponding portions of the images indicate that those corresponding portions are on surfaces that are closer to the camera system and smaller parallax shifts indicate that the corresponding portions are on surfaces that are farther away from the camera system). These techniques for computing depth based on parallax shifts are sometimes referred to as Depth from Stereo

[0065] Accordingly, FIG. 1E depicts a stereo polarization camera system 10″ having a first polarization camera module 10-1″ and a second polarization camera module 10-2″ having substantially parallel optical axes and spaced apart along a baseline 10-B. In the embodiment shown in FIG. 1E, the first polarization camera module 10-1″ and includes polarization cameras 10A″, 10B″, 10C″, and 10D″ arranged in a 2×2 array similar to that shown in FIG. 1C and 1D. Likewise, the second polarization camera module 10-2″ and includes polarization cameras 10E″, 10F″, 10G″, and 10H″ arranged in a 2×2 array, and the overall stereo polarization camera module 10″ includes eight individual polarization cameras (e.g., eight separate image sensors behind eight separate lenses). In some embodiments of the present disclosure, corresponding polarization cameras of polarization camera modules 10-1″ and 10-2″ are configured to capture polarization raw frames with

1

substantially the same polarizations. For example, cameras 10A″ and 10E″ may both have linear polarizing filters at a same angle of 0°, cameras 10B″ and 10F″ may both have linear polarizing filters at a same angle of 45°, cameras 10C″ and 10G″ may both have linear polarizing filters at a same angle of 90°, and cameras 10D″ and 10H″ may both have linear polarizing filters at a same angle of 135°.

[0066] Embodiments of the present disclosure are not limited to the particular embodiment shown in FIG. 1E. In some embodiments, a stereo polarization camera system includes three or more polarization camera modules, where each polarization camera module includes multiple polarization cameras arranged in array and configured, using polarizing filters, to capture polarization raw frames of different polarizations. As noted above, in some embodiments of the present disclosure, one or more of the individual polarization cameras of a polarization camera module may include a color filter and, as such, one or more of the polarization cameras in a stereo polarization camera module may also include a color filter.

[0067] In a manner similar to that described for calibrating or registering cameras within a camera module, the various polarization camera modules of a stereo polarization camera system may also be registered with one another by capturing multiple images of calibration targets and computing intrinsic and extrinsic matrices for the various camera modules.

[0068] While the embodiment of a stereo polarization camera system shown in FIG. 1E includes two polarization camera modules, each having four polarization cameras, embodiments of the present disclosure are not limited thereto.

[0069] For example, in some embodiments of the present disclosure, a stereo polarization camera system includes a plurality of polarization camera modules, where each of the polarization camera modules includes three or more individual polarization cameras, each of the individual polarization cameras of a polarization camera module having polarizing filters with different polarization states (e.g., different angles of linear polarization).

[0070] In some embodiments of the present disclosure, a stereo polarization camera system includes a plurality of polarization camera modules that are spaced apart along one or more baselines, where each of the polarization camera modules includes a single polarization camera configured to capture polarization raw frames with different polarizations, in accordance with embodiments such as that described above with respect to FIG. 1B. For example, in some embodiments of the present disclosure, the polarization camera of each module may include a polarization mask (e.g., similar to the polarization mask shown in FIG. 1B) such that each individual polarization camera captures images where the pixels detect light in accordance with

1

a mosaic pattern of different polarizing filters (e.g., polarizing filters at different angles). For example, in the embodiment shown in FIG. 1B, each 2×2 block of pixels of the polarization mask includes linear polarizers at linear polarization angles of 0°, 45°, 90°, and 135°. In other embodiments of the present disclosure, the individual polarization cameras may include mechanically or electronically controllable polarizing filters, as discussed above with respect to FIG. 1B, to enable the polarization cameras to capture polarization raw frames of different polarizations.

[0071] While the above embodiments specified that the individual polarization camera modules or the polarization cameras that are spaced apart along one or more baselines in the stereo polarization camera system have substantially parallel optical axes, embodiments of the present disclosure are not limited thereto. For example, in some embodiment of the present disclosure, the optical axes of the polarization camera modules are angled toward each other such that the polarization camera modules provide differently angled views of objects in the designed working distance (e.g., where the optical axes cross or intersect in the neighborhood of the designed working distance from the stereo camera system).

[0072] According to various embodiments of the present disclosure, the processing circuit 100 is implemented using one or more electronic circuits configured to perform various operations as described in more detail below. Types of electronic circuits may include a central processing unit (CPU), a graphics processing unit (GPU), an artificial intelligence (AI) accelerator (e.g., a vector processor, which may include vector arithmetic logic units configured efficiently perform operations common to neural networks, such dot products and softmax), a field programmable gate array (FPGA), an application specific integrated circuit (ASIC), a digital signal processor (DSP), or the like. For example, in some circumstances, aspects of embodiments of the present disclosure are implemented in program instructions that are stored in a non-volatile computer readable memory where, when executed by the electronic circuit (e.g., a CPU, a GPU, an AI accelerator, or combinations thereof), perform the operations described herein to compute a processing output 20, such as an instance segmentation map, from input polarization raw frames 18. The operations performed by the processing circuit 100 may be performed by a single electronic circuit (e.g., a single CPU, a single GPU, or the like) or may be allocated between multiple electronic circuits (e.g., multiple GPUs or a CPU in conjunction with a GPU). The multiple electronic circuits may be local to one another (e.g., located on a same die, located within a same package, or located within a same embedded device or computer system) and/or may be remote from one other (e.g., in communication over a network such as a local personal area network such as Bluetooth®, over a local area network such as a local wired and/or wireless network, and/or over wide area network such as the internet, such a case where some operations are performed locally and other operations are performed on a server hosted by a cloud computing service). One or more electronic circuits operating to implement the processing circuit 100 may be referred to herein as a computer or a computer system, which may include memory storing instructions that, when executed by the one or more electronic circuits, implement the systems and methods described herein.

[0073] FIGS. 2A, 2B, 2C, and 2D provide background for illustrating the segmentation maps computed by a comparative approach and semantic segmentation or instance segmentation based on polarization raw frames according to aspects of embodiments of the present disclosure. In more detail, FIG. 2A is an image or intensity image of a scene with one real transparent ball placed on top of a printout of photograph depicting another scene containing two transparent balls ("spoofs") and some background clutter. FIG. 2B depicts a segmentation mask as computed by a Mask Region-based Convolutional Neural Network (Mask R-CNN) identifying instances of transparent balls overlaid on the intensity image of FIG. 2A using different patterns of lines, where the real transparent ball is correctly identified as an instance, and the two spoofs are incorrectly identified as instances. In other words, the Mask R-CNN algorithm has been fooled into labeling the two spoof transparent balls as instances of actual transparent balls in the scene.

[0074] FIG. 2C is an angle of linear polarization (AOLP) image computed from polarization raw frames captured of the scene according to one embodiment of the present invention. As shown in FIG. 2C, transparent objects have a very unique texture in polarization space such as the AOLP domain, where there is a geometry-dependent signature on edges and a distinct or unique or particular pattern that arises on the surfaces of transparent objects in the angle of linear polarization. In other words, the *intrinsic texture* of the transparent object (e.g., as opposed to extrinsic texture adopted from the background surfaces visible through the transparent object) is more visible in the angle of polarization image of FIG. 2C than it is in the intensity image of FIG. 2A.

[0075] FIG. 2D depicts the intensity image of FIG. 2A with an overlaid segmentation mask as computed using polarization data in accordance with an embodiment of the present invention, where the real transparent ball is correctly identified as an instance using an overlaid pattern of lines and the two spoofs are correctly excluded as instances (e.g., in contrast to FIG. 2B, FIG. 2D does not include overlaid patterns of lines over the two spoofs). While FIGS. 2A, 2B, 2C, and

1

2D illustrate an example relating to detecting a real transparent object in the presence of spoof transparent objects, embodiments of the present disclosure are not limited thereto and may also be applied to other optically challenging objects, such as transparent, translucent, and non-matte or non-Lambertian objects, as well as non-reflective (e.g., matte black objects) and multipath inducing objects.

[0076] Accordingly, some aspects of embodiments of the present disclosure relate to extracting, from the polarization raw frames, tensors in representation space (or first tensors in first representation spaces, such as polarization feature maps) to be supplied as input to surface characterization algorithms or other computer vision algorithms. These first tensors in first representation space may include polarization feature maps that encode information relating to the polarization of light received from the scene such as the AOLP image shown in FIG. 2C, degree of linear polarization (DOLP) feature maps, and the like (e.g., other combinations from Stokes vectors or transformations of individual ones of the polarization raw frames). For example, in some embodiments of the present disclosure, the feature extractor 700 further computes surface normals in accordance with equations (2), (3), (4), and (5), discussed in more detail below, based on the polarization raw frames. In some embodiments, these polarization feature maps are used together with non-polarization feature maps (e.g., intensity images such as the image shown in FIG. 2A) to provide additional channels of information for use by semantic segmentation algorithms.

[0077] While embodiments of the present invention are not limited to use with particular computer vision algorithms for analyzing images, some aspects of embodiments of the present invention relate to deep learning frameworks for polarization-based detection of optically challenging objects (e.g., transparent, translucent, non-Lambertian, multipath inducing objects, and non-reflective or very dark objects), where these frameworks may be referred to as Polarized Convolutional Neural Networks (Polarized CNNs). This Polarized CNN framework includes a backbone that is suitable for processing the particular texture of polarization and can be coupled with other computer vision architectures such as Mask R-CNN (e.g., to form a Polarized Mask R-CNN architecture) to produce a solution for accurate and robust characterization of transparent objects and other optically challenging objects. Furthermore, this approach may be applied to scenes with a mix of transparent and non-transparent (e.g., opaque objects) and can be used to characterize transparent, translucent, non-Lambertian, multipath inducing, dark, and opaque surfaces of the object or objects under inspection.

[0078] *Polarization feature representation spaces*

[0079] Some aspects of embodiments of the present disclosure relate to systems and methods for extracting features from polarization raw frames in operation 650, where these extracted features are used in operation 690 in the robust detection of optically challenging characteristics in the surfaces of objects. In contrast, comparative techniques relying on intensity images alone may fail to detect these optically challenging features or surfaces (e.g., comparing the intensity image of FIG. 2A with the AOLP image of FIG. 2C, discussed above). The term “first tensors” in “first representation spaces” will be used herein to refer to features computed from (e.g., extracted from) polarization raw frames 18 captured by a polarization camera, where these first representation spaces include at least polarization feature spaces (e.g., feature spaces such as AOLP and DOLP that contain information about the polarization of the light detected by the image sensor) and may also include non-polarization feature spaces (e.g., feature spaces that do not require information regarding the polarization of light reaching the image sensor, such as images computed based solely on intensity images captured without any polarizing filters).

[0080] The interaction between light and transparent objects is rich and complex, but the material of an object determines its transparency under visible light. For many transparent household objects, the majority of visible light passes straight through and a small portion (~4% to ~8%, depending on the refractive index) is reflected. This is because light in the visible portion of the spectrum has insufficient energy to excite atoms in the transparent object. As a result, the texture (e.g., appearance) of objects behind the transparent object (or visible through the transparent object) dominate the appearance of the transparent object. For example, when looking at a transparent glass cup or tumbler on a table, the appearance of the objects on the other side of the tumbler (e.g., the surface of the table) generally dominate what is seen through the cup. This property leads to some difficulties when attempting to detect surface characteristics of transparent objects such as glass windows and glossy, transparent layers of paint, based on intensity images alone:

[0081] FIG. 3 is a high-level depiction of the interaction of light with transparent objects and non-transparent (e.g., diffuse and/or reflective) objects. As shown in FIG. 3, a polarization camera 10 captures polarization raw frames of a scene that includes a transparent object 302 in front of an opaque background object 303. A light ray 310 hitting the image sensor 14 of the polarization camera 10 contains polarization information from both the transparent object 302 and the background object 303. The small fraction of reflected light 312 from the transparent object 302 is heavily polarized, and thus has a large impact on the polarization measurement, in contrast

1

to the light 313 reflected off the background object 303 and passing through the transparent object 302.

[0082] Similarly, a light ray hitting the surface of an object may interact with the shape of the surface in various ways. For example, a surface with a glossy paint may behave substantially similarly to a transparent object in front of an opaque object as shown in FIG. 3, where interactions between the light ray and a transparent or translucent layer (or clear coat layer) of the glossy paint causes the light reflecting off of the surface to be polarized based on the characteristics of the transparent or translucent layer (e.g., based on the thickness and surface normals of the layer), which are encoded in the light ray hitting the image sensor. Similarly, as discussed in more detail below with respect to shape from polarization (SfP) theory, variations in the shape of the surface (e.g., direction of the surface normals) may cause significant changes in the polarization of light reflected by the surface of the object. For example, smooth surfaces may generally exhibit the same polarization characteristics throughout, but a scratch or a dent in the surface changes the direction of the surface normals in those areas, and light hitting scratches or dents may be polarized, attenuated, or reflected in ways different than in other portions of the surface of the object. Models of the interactions between light and matter generally consider three fundamentals: geometry, lighting, and material. Geometry is based on the shape of the material. Lighting includes the direction and color of the lighting. Material can be parameterized by the refractive index or angular reflection/transmission of light. This angular reflection is known as a bi-directional reflectance distribution function (BRDF), although other functional forms may more accurately represent certain scenarios. For example, the bidirectional subsurface scattering distribution function (BSSRDF) would be more accurate in the context of materials that exhibit subsurface scattering (e.g. marble or wax).

[0083] A light ray 310 hitting the image sensor 16 of a polarization camera 10 has three measurable components: the intensity of light (intensity image/I), the percentage or proportion of light that is linearly polarized (degree of linear polarization/DOLP/ρ), and the direction of that linear polarization (angle of linear polarization/AOLP/ϕ). These properties encode information about the surface curvature and material of the object being imaged, which can be used by the predictor 750 to detect transparent objects, as described in more detail below. In some embodiments, the predictor 750 can detect other optically challenging objects based on similar polarization properties of light passing through translucent objects and/or light interacting with multipath inducing objects or by non-reflective objects (e.g., matte black objects).

[0084] Therefore, some aspects of embodiments of the present invention relate to using a feature extractor 700 to compute first tensors in one or more first representation spaces, which may include derived feature maps based on the intensity I, the DOLP ρ, and the AOLP ϕ. The feature extractor 700 may generally extract information into first representation spaces (or first feature spaces) which include polarization representation spaces (or polarization feature spaces) such as "polarization images," in other words, images that are extracted based on the polarization raw frames that would not otherwise be computable from intensity images (e.g., images captured by a camera that did not include a polarizing filter or other mechanism for detecting the polarization of light reaching its image sensor), where these polarization images may include DOLP ρ images (in DOLP representation space or feature space), AOLP ϕ images (in AOLP representation space or feature space), other combinations of the polarization raw frames as computed from Stokes vectors, as well as other images (or more generally first tensors or first feature tensors) of information computed from polarization raw frames. The first representation spaces may include non-polarization representation spaces such as the intensity I representation space.

[0085] Measuring intensity I, DOLP ρ, and AOLP ϕ at each pixel requires 3 or more polarization raw frames of a scene taken behind polarizing filters (or polarizers) at different angles, ϕ_{pol} (e.g., because there are three unknown values to be determined: intensity I, DOLP ρ, and AOLP ϕ. For example, the FLIR® Blackfly® S Polarization Camera described above captures polarization raw frames with polarization angles ϕ_{pol} at 0 degrees, 45 degrees, 90 degrees, or 135 degrees, thereby producing four polarization raw frames $I_{\phi_{pol}}$, denoted herein as I_0, I_{45}, I_{90}, and I_{135}.

[0086] The relationship between $I_{\phi_{pol}}$ and intensity I, DOLP ρ, and AOLP ϕ at each pixel can be expressed as:

$$I_{\phi_{pol}} = I\left(1 + \rho\cos\left(2(\phi - \phi_{pol})\right)\right) \qquad (1)$$

[0087] Accordingly, with four different polarization raw frames $I_{\phi_{pol}}$ (I_0, I_{45}, I_{90}, and I_{135}), a system of four equations can be used to solve for the intensity I, DOLP ρ, and AOLP ϕ.

[0088] Shape from Polarization (SfP) theory (see, e.g., Gary A Atkinson and Edwin R Hancock. Recovery of surface orientation from diffuse polarization. IEEE transactions on image processing, 15(6):1653–1664, 2006.) states that the relationship between the refractive index (n), azimuth angle (θ_a) and zenith angle (θ_z) of the surface normal of an object and the ϕ and ρ components of the light ray coming from that object follow the following characteristics when diffuse reflection is dominant:

$$\rho = \frac{\left(n - \frac{1}{n}\right)^2 \sin^2(\theta_z)}{2 + 2n^2 - \left(n + \frac{1}{n}\right)^2 \sin^2 \theta_z + 4 \cos \theta_z \sqrt{n^2 - \sin^2 \theta_z}} \quad (2)$$

$$\phi = \theta_a \quad (3)$$

and when the specular reflection is dominant:

$$\rho = \frac{2 \sin^2 \theta_z \cos \theta_z \sqrt{n^2 - \sin^2 \theta_z}}{n^2 - \sin^2 \theta_z - n^2 \sin^2 \theta_z + 2 \sin^4 \theta_z} \quad (4)$$

$$\phi = \theta_a - \frac{\pi}{2} \quad (5)$$

Note that in both cases ρ increases exponentially as θ_z increases and if the refractive index is the same, specular reflection is much more polarized than diffuse reflection.

[0089] Accordingly, some aspects of embodiments of the present disclosure relate to applying SfP theory to detect the shapes of surfaces (e.g., the orientation of surfaces) based on the raw polarization frames 18 of the surfaces. This approach enables the shapes of objects to be characterized without the use of other computer vision techniques for determining the shapes of objects, such as time-of-flight (ToF) depth sensing and/or stereo vision techniques, although embodiments of the present disclosure may be used in conjunction with such techniques.

[0090] More formally, aspects of embodiments of the present disclosure relate to computing first tensors 50 in first representation spaces, including extracting first tensors in polarization representation spaces such as forming polarization images (or extracting derived polarization feature maps) in operation 650 based on polarization raw frames captured by a polarization camera 10.

[0091] Light rays coming from a transparent objects have two components: a reflected portion including reflected intensity I_r, reflected DOLP ρ_r, and reflected AOLP ϕ_r and the refracted portion including refracted intensity I_t, refracted DOLP ρ_t, and refracted AOLP ϕ_t. The intensity of a single pixel in the resulting image can be written as:

$$I = I_r + I_t \quad (6)$$

[0092] When a polarizing filter having a linear polarization angle of ϕ_{pol} is placed in front of the camera, the value at a given pixel is:

$$I_{\phi_{pol}} = I_r \left(1 + \rho_r \cos\left(2(\phi_r - \phi_{pol})\right)\right) + I_t \left(1 + \rho_t \cos\left(2(\phi_t - \phi_{pol})\right)\right) \quad (7)$$

[0093] Solving the above expression for the values of a pixel in a DOLP ρ image and a pixel in an AOLP ϕ image in terms of I_r, ρ_r, ϕ_r, I_t, ρ_t, and ϕ_t:

$$\rho = \frac{\sqrt{(I_r\rho_r)^2 + (I_t\rho_t)^2 + 2I_t\rho_t I_r\rho_r \cos(2(\phi_r - \phi_t))}}{I_r + I_t} \tag{8}$$

$$\phi = \arctan\left(\frac{I_r\rho_r \sin(2(\phi_r - \phi_t))}{I_t\rho_t + I_r\rho_r \cos(2(\phi_r - \phi_t))}\right) + \phi_r \tag{9}$$

[0094] Accordingly, equations (7), (8), and (9), above, provide a model for forming first tensors 50 in first representation spaces that include an intensity image I, a DOLP image ρ, and an AOLP image ϕ according to one embodiment of the present disclosure, where the use of polarization images or tensor in polarization representation spaces (including DOLP image ρ and an AOLP image ϕ based on equations (8) and (9)) enables the reliable detection of optically challenging surface characteristics of objects that are generally not detectable by comparative systems that use only intensity I images as input.

[0095] Equations (8) and (9), above, can be represented more generally in accordance with Stokes parameters:

$$\rho = \frac{\sqrt{(S_1)^2 + (S_2)^2}}{S_0} \tag{10}$$

$$\phi = 0.5 * \arctan\frac{S_2}{S_1} \tag{11}$$

where S_0, S_1, and S_2 are the Stokes parameters.

[0096] In more detail, first tensors in polarization representation spaces (among the derived feature maps 50) such as the polarization images DOLP ρ and AOLP ϕ can reveal surface characteristics of objects that might otherwise appear textureless in an intensity I domain. A transparent object may have a texture that is invisible in the intensity domain I because this intensity is strictly dependent on the ratio of I_r / I_t (see equation (6)). Unlike opaque objects where $I_t = 0$, transparent objects transmit most of the incident light and only reflect a small portion of this incident light. As another example, thin or small deviations in the shape of an otherwise smooth surface (or smooth portions in an otherwise rough surface) may be substantially invisible or have low contrast in the intensity I domain (e.g., a domain that does not encode polarization of light), but may be very visible or may have high contrast in a polarization representation space such as DOLP ρ or AOLP ϕ.

[0097] As such, one exemplary method to acquire surface topography is to use polarization cues in conjunction with geometric regularization. The Fresnel equations relate the AOLP ϕ and the DOLP ρ with surface normals. These equations can be useful for detecting optically challenging objects by exploiting what is known as *polarization patterns* of the surfaces of these optically challenging objects. A polarization pattern is a tensor of size $[M, N, K]$ where M and N are horizontal and vertical pixel dimensions, respectively, and where K is the polarization data channel, which can vary in size. For example, if circular polarization is ignored and only linear polarization is considered, then K would be equal to two, because linear polarization has both an angle and a degree of polarization (AOLP ϕ and DOLP ρ). Analogous to a Moire pattern, in some embodiments of the present disclosure, the feature extraction module 700 extracts a polarization pattern in polarization representation spaces (e.g., AOLP space and DOLP space).

[0098] While the preceding discussion provides specific examples of polarization representation spaces based on linear polarization in the case of using a polarization camera having one or more linear polarizing filters to capture polarization raw frames corresponding to different angles of linear polarization and to compute tensors in linear polarization representation spaces such as DOLP and AOLP, embodiments of the present disclosure are not limited thereto. For example, in some embodiments of the present disclosure, a polarization camera includes one or more circular polarizing filters configured to pass only circularly polarized light, and where polarization patterns or first tensors in circular polarization representation space are further extracted from the polarization raw frames. In some embodiments, these additional tensors in circular polarization representation space are used alone, and in other embodiments they are used together with the tensors in linear polarization representation spaces such as AOLP and DOLP. For example, a polarization pattern including tensors in polarization representation spaces may include tensors in circular polarization space, AOLP, and DOLP, where the polarization pattern may have dimensions $[M, N, K]$, where K is three to further include the tensor in circular polarization representation space.

[0099] Accordingly, some aspects of embodiments of the present disclosure relate to supplying first tensors in the first representation spaces (e.g., including feature maps in polarization representation spaces) extracted from polarization raw frames as inputs to a predictor for computing or detecting surface characteristics of transparent objects and/or other optically challenging surface characteristics of objects under inspection. These first tensors may include derived feature maps which may include an intensity feature map I, a degree of linear polarization (DOLP) ρ feature map, and an angle of linear polarization (AOLP) ϕ feature map, and where the DOLP ρ feature map and the AOLP ϕ feature map are examples of polarization feature maps or tensors in polarization representation spaces, in reference to feature maps that encode information regarding the polarization of light detected by a polarization camera.

[00100] In some embodiments, the feature maps or tensors in polarization representation spaces are supplied as input to, for example, detection algorithms that make use of SfP theory to characterize the shape of surfaces of objects imaged by the polarization cameras 10. For example, in some embodiments, in the case of diffuse reflection, equations (2) and (3) are used to compute the zenith angle (θ_z) and the azimuth angle (θ_a) of the surface normal of a surface in the scene based on the DOLP ρ and the index of refraction n. Likewise, in the case of specular reflection, equations (3) and (5) are used to compute the zenith angle (θ_z) and the azimuth angle (θ_a) of the surface normal of a surface in the scene based on the DOLP ρ and the index of refraction n. As one example, a closed form solution for computing the zenith angle (θ_z) based on Equation (2) according to one embodiment of the present disclosure in accordance with the following steps:

$$aa = \left(n - \frac{1}{n}\right)^2 + \rho\left(n + \frac{1}{n}\right)^2$$

$$bb = 4\rho(n^2 + 1)(aa - 4\rho)$$

$$cc = bb^2 + 16\rho^2(16\rho^2 - aa^2)(n^2 - 1)^2$$

$$dd = \sqrt{\frac{-bb - \sqrt{cc}}{2(16\rho^2 - aa^2)}}$$

$$\theta_z = aa\sin dd$$

[00101] Additional details on computing surface normal directions based on polarization raw frames can be found, for example, in U.S. Pat. Nos. 10,260,866 and 10,557,705 and Kadambi, Achuta, et al. "Polarized 3D: High-quality depth sensing with polarization cues." *Proceedings of the IEEE International Conference on Computer Vision*. 2015.

[00102] *Computing Polarization Cues from Multi-camera Arrays*

[00103] Ordinarily, multipolar cues are obtained from a monocular viewpoint. Existing methods use multipolar filters (e.g., a polarization mask as shown in FIG. 1B) or multiple CCD or CMOS sensors to multiplex different polarization channels in a single view (e.g., multiple sensors behind a single lens system) or time multiplexed systems (e.g., where different polarization raw frames are captured at different times, such as sequentially captured, which may require that the scene 1 remain substantially or constant from one capture to the next in order for the views to be the same). In particular, the techniques described above for calculating polarization cues such as the angle of linear polarization (AOLP) ϕ and the degree of linear

Date Reçue/Date Received 2022-07-05

1

polarization (DOLP) ρ generally assume that the polarization raw frames are captured from the same viewpoint.

[00104] However, there are some circumstances in which the above assumption of a single viewpoint may not hold. For example, polarization raw frames corresponding to different polarization states may be captured from different viewpoints when using a polarization camera array that includes multiple polarization cameras at different locations, such as the embodiments shown in FIGS. 1C, 1D, and 1E. While placing the individual polarization cameras closer together may reduce error, physical constraints (e.g., the size of the individual polarization cameras, such as the size and shape of their corresponding packaging as well as lenses and image sensors contained therein) may limit the placement of the polarization cameras.

[00105] Accordingly, some aspects of embodiments of the present disclosure relate to systems and methods for computing polarization cues such as AOLP ϕ and DOLP ρ from polarization raw frames captured from different viewpoints, such as by using an array of polarization cameras. Generally, this involves a technique for decoupling parallax cues due to the different positions of the separate polarization cameras and the desired polarization cues. This is challenging because parallax cues and polarization cues are linked in that both the parallax between two views and the sensed polarization are related to the geometry of the relationship between the polarization cameras and the imaged surface. The comparative approaches to obtaining AOLP and DOLP assume that the polarization channels are acquired from the same viewpoint and therefore applying comparative techniques to the data captured by the array of polarization cameras likely results in errors or ambiguity.

[00106] FIG. 4 is a flowchart of a method for estimating polarization cues under parallax ambiguities according to one embodiment of the present disclosure. In the embodiment shown in FIG. 4, polarization raw frames captured from a plurality of different viewpoints, such as by an array of polarization cameras such as that shown in FIGS. 1C, 1D, and 1E are supplied as input to the process. In some embodiments of the present disclosure, estimates of the DOLP ρ and AOLP ϕ in accordance with embodiments of the present disclosure are computed by a processing circuit 100 through an iterative process. Note that the estimated DOLP ρ and estimated AOLP ϕ correspond to tensors (e.g., two dimensional tensors) having aspect ratios corresponding to the polarization raw frames, e.g., where the values of the DOLP ρ tensor and AOLP ϕ tensor correspond to the estimated degree of linear polarization and the angle of linear polarization in various portions of the captured polarization raw frames.

[00107] In operation 410, the processing circuit computes an initial estimated DOLP ρ_0 and an initial estimated AOLP ϕ_0 using the Stokes vectors (e.g., in accordance with equations (10) and (11), above or, more specifically, in accordance with equations (8) and (9). These initial estimated DOLP ρ_0 and AOLP ϕ_0 will likely be incorrect due to the parallax shift between the different individual polarization cameras of the polarization camera array.

[00108] In operation 430, the processing circuit 100 estimates the geometry of the surfaces of the scene depicted in the polarization raw frames. In some embodiments of the present disclosure, the processing circuit 100 uses a view correspondence-based approach to generate a coarse model of the scene using parallax from the stereo view of the scene, due to the offset between the locations of the cameras in the array (e.g., using depth from stereo techniques, as discussed, for example, in Kadambi, A. et al. (2015)). In operation 450, this coarse geometry may then be refined using the current calculated DOLP ρ_i and AOLP ϕ_i values (initially, $i = 0$) (see, e.g., U.S. Pat. Nos. 10,260,866 and 10,557,705 and Kadambi, A. et al. (2015)).

[00109] The estimated geometry computed in operation 450 is then used to update the estimated values of the DOLP ρ and the AOLP ϕ. For example, in an i-th iteration, a previously calculated DOLP ρ_{i-1} and a previously calculated AOLP ϕ_{i-1} may be used to compute the estimated geometry in operation 450 and, in operation 470, the processing system 100 refines the DOLP and AOLP calculations based on the new estimated geometry (in accordance with the Fresnel equations that relate AOLP and DOLP to slope) to compute new estimates DOLP ρ_i and AOLP ϕ_i.

[00110] In operation 490, the processing system 100 determines whether to continue with another iteration of the process of estimating the DOLP ρ and AOLP ϕ. In more detail, in some embodiments, a change in the DOLP $\Delta\rho$ is computed based on the difference between the updated DOLP ρ_i and the previously calculated DOLP ρ_{i-1} (e.g., $\Delta\rho = |\rho_i - \rho_{i-1}|$). Likewise, a change in the AOLP $\Delta\phi$ is computed based on the difference between the updated AOLP ϕ_i and the previously calculated AOLP ϕ_{i-1} (e.g., $\Delta\phi = |\phi_i - \phi_{i-1}|$). If either of these changes in polarization cues (e.g., both $\Delta\rho$ and $\Delta\phi$) is greater than corresponding threshold values (e.g., ρ_{th} and ϕ_{th}) across the computed tensors, then the process continues by using the updated DOLP ρ_i and AOLP ϕ_i to refine the coarse model in operation 450, and then updating the DOLP and AOLP values based on this new estimated geometry. If both of the changes in the polarization cues are less than their corresponding thresholds, then the estimation process is complete and the estimated DOLP ρ_i and AOLP ϕ_i are output from the estimation process, and may be used in computing further processing outputs, such as instance segmentation maps.

[00111] *Multi-spectral stereo with Polarization Imaging*

[00112] In many circumstances, such as in remote sensing, multi-spectral images of scenes are capable of capturing information that would otherwise be hidden from view. For example, multi-spectral or hyper-spectral imaging is capable of detecting surface properties of scenes, such as detecting soil properties like moisture, organic content, and salinity, oil impacted soils, which may be useful in agriculture. As another example, multi-spectral imaging may enable the detection of camouflaged targets, such as military vehicles under partial vegetation cover or small military objects within relatively larger pixels. As a further example, multi-spectral imaging enables material identification and mapping, such as detecting the presence or absence of materials in relief geography, mapping of heavy metals and other toxic wastes in mining areas. Multi-spectral imaging also enables the detection of the presence of particular materials, such as water/oil spills (this is of particular importance to indoor robots so they can avoid or perform path planning around these spills and for robotic vacuum cleaners to detect, locate, and clean up spills and other small, dark, and/or specular dirt). Multi-spectral imaging may also be used for material inspection, such as detecting cracks and rust in industrial equipment such as industrial boilers and railway tracks, in which failure can be extremely hazardous and where recovery can be expensive.

[00113] In these above examples, computer vision techniques that use comparative and standard color images (e.g., red, green, and blue images) as input, may not be able to detect these types of objects, but the use of multi-spectral or hyper-spectral imaging, combined with polarization information, may provide additional cues that can be detected and recognized by computer vision algorithms and instance detection techniques (e.g., using trained convolutional neural networks).

[00114] Generally, the spectral radiance of a surface measures the rate of photons reflected from a surface as a function of surface area, slope, and incident wavelength. The spectral radiance function of most natural images are regular functions of wavelengths which makes it possible to represent these using a low-dimensional linear model. In other words, the spectral representation of light reflected from the surface can be represented as a linear combination of spectral basis functions:

$$s \approx \sum_{i=0}^{n} w_i B_i \tag{12}$$

where w_i are the linear weights, B_i represents the spectral basis function, and n is the dimensionality of the system. Related work in the area of spectral radiance profiles of natural objects show that, for the most part, the spectral radiance of natural objects can be represented accurately by five or six linear basis functions.

[00115] Accordingly, some aspects embodiments of the present disclosure, relate to collecting spectral information simultaneously with polarization information using a stereo imaging pair wherein each camera system (or camera module) of the stereo pair includes a camera array that allows for capturing both the spectral and polarization information.

[00116] FIG. 5A is a perspective view of a multi-spectral stereo polarization camera system according to one embodiment of the present disclosure. Embodiments of a multi-spectral stereo polarization camera system as shown in FIG. 5A are substantially similar to the stereo polarization camera system shown in FIG. 1E in that FIG. 5A depicts a multi-spectral stereo polarization camera system 510 having a first polarization camera module 510-1″ and a second polarization camera module 510-2″ having substantially parallel optical axes and spaced apart along a baseline 510-B. In the embodiment shown in FIG. 5A, the first polarization camera module 510-1″ and includes polarization cameras 510A″, 510B″, 510C″, and 510D″ arranged in a 2×2 array similar to that shown in FIG. 1C and 1D. Likewise, the second polarization camera module 510-2″ and includes polarization cameras 510E″, 510F″, 510G″, and 510H″ arranged in a 2×2 array, and the overall multi-spectral stereo polarization camera module 510 includes eight individual polarization cameras (e.g., eight separate image sensors behind eight separate lenses). In some embodiments of the present disclosure, corresponding polarization cameras of polarization camera modules 510-1″ and 510-2″ are configured to capture polarization raw frames with substantially the same polarizations. For example, cameras 510A″ and 510E″ may both have linear polarizing filters at a same angle of 0°, cameras 510B″ and 510F″ may both have linear polarizing filters at a same angle of 45°, cameras 510C″ and 510G″ may both have linear polarizing filters at a same angle of 90°, and cameras 510D″ and 510H″ may both have linear polarizing filters at a same angle of 135°.

[00117] FIG. 5B is a view of a multi-spectral stereo polarization camera system according to one embodiment of the present disclosure, along a direction parallel to the optical axis of the multi-spectral stereo polarization camera system. FIG. 5C depicts cut-away side views of example individual polarization cameras of a multi-spectral stereo polarization camera system according to one embodiment of the present disclosure. As shown in FIG. 5C, each of the individual polarization cameras (e.g., 510A″, 510B″, 510E″, and 510F″) includes a corresponding color filter 518 (e.g., 518A″, 518B″, 518E″, and 518F″) in the optical path of the individual polarization camera, in addition to a corresponding lens 512, a corresponding image sensors 514, and a corresponding polarizing filter 516. While FIG. 5C depicts the color filter 518 as being within a housing and behind the lens 512, embodiments of the present disclosure are not limited thereto. For example, in some embodiments, the color filter 518 is located in front of the lens 512. Likewise, in some embodiments, the polarizing filter is located in front of the lens 512.

[00118] In the embodiment shown in FIG. 5B, each of the individual polarization cameras includes a color filter that transmits light in only one corresponding portion of the visible spectrum (as opposed to a Bayer filter, which has a mosaic pattern and that typically transmits light in three different portions of the spectrum corresponding to red, green, and blue light). In the example embodiment shown in FIG. 5B, first polarization camera 510A" has a color filter 518A" that is configured to transmit light in a red (R) portion of the spectrum, second polarization camera 510B" has a color filter 518B" that is configured to transmit light in a first green (G1) portion of the spectrum, third polarization camera 510C" has a color filter 518C" that is configured to transmit light in a second green (G2) portion of the spectrum (which may be different from the first green portion G1 of the spectrum, e.g., with a peak shifted by 15 to 20 nm), and fourth polarization camera 510D" has a color filter 518D" that is configured to transmit light in a blue (B) portion of the spectrum. Together, the four polarization cameras of the first polarization camera module 510-1" capture light at four different polarization states (e.g., four different linear polarizations of 0°, 45°, 90°, and 135°) and four different colors (e.g., R, G1, G2, and B). In the particular embodiment shown in FIG. 5B, for example, the first polarization camera 510A" captures red light polarized at 0°, the second polarization camera 510B" captures first green light polarized at 45°, the third polarization camera 510C" captures second green light polarized at 90°, and the fourth polarization camera 510D" captures blue light polarized at 135°. However, embodiments of the present disclosure are not limited thereto. For example, in various embodiments the color filters may be associated with different polarizing filters.

[00119] In a similar manner, the individual polarization cameras (e.g., cameras 510E", 510F", 510G", and 510BH") of the second polarization camera module 510-2" includes a separate color filter 518 that are configured to transmit light in different portions of the electromagnetic spectrum and different from one another. In some embodiment of the present invention, each of the color filters of the second polarization camera module 510-2" transmits light in a portion of the spectrum that is shifted by some amount (e.g., where the peak of the spectral profile of the color filter is shifted, either toward the longer wavelengths or toward shorter wavelengths, by about 10 nanometers to about 20 nanometers) from the corresponding color filter in the first polarization camera module 510-1″.

[00120] In the example embodiment shown in FIG. 5B, fifth polarization camera 510E″ has a color filter 518E″ that is configured to transmit light in a red (R′) portion of the spectrum that is shifted by about 10 to 20 nanometers from the spectrum R transmitted by corresponding color filter 518A″ of the corresponding polarization camera 510A″ of the first polarization camera module 510-1″. Likewise, sixth polarization camera 510F″ has a color filter 518F″ that is configured to transmit light in a first green (G1′) portion of the spectrum that is shifted by about 10 to 20 nanometers from the spectrum G1 transmitted by corresponding color filter 518B″ of the corresponding polarization camera 510B″ of the first polarization camera module 510-1″ (and, in some embodiments, also a different spectrum from spectrum G2). The seventh polarization camera 510G″ has a color filter 518G″ that is configured to transmit light in a second green (G2′) portion of the spectrum that is shifted by about 10 to 20 nanometers from the spectrum G2 transmitted by corresponding color filter 518C″ of the corresponding polarization camera 510C″ of the first polarization camera module 510-1″ (and, in some embodiments, also a different spectrum for spectrum G1). The eighth polarization camera 510H″ has a color filter 518H″ that is configured to transmit light in a blue (B′) portion of the spectrum that is shifted by about 10 to 20 nanometers from the spectrum B transmitted by corresponding color filter 518D″ of the corresponding polarization camera 510D″ of the first polarization camera module 510-1″.

[00121] Together, the four polarization cameras of the second polarization camera module 510-2″ capture light at four different polarization states (e.g., four different linear polarizations of 0°, 45°, 90°, and 135°) and four different colors (e.g., R′, G1′, G2′, and B′) that are also different from the four colors captured by the first polarization camera module 510-1″. As a result, the multi-spectral stereo polarization camera system 510 shown in FIGS. 5A, 5B, and 5C is configured to detect light of eight different colors and at four different polarization angles.

[00122] While some embodiments of the present disclosure are described in detail above with respect to FIGS. 5A, 5B, and 5C, embodiments of the present disclosure are not limited thereto. For example, as noted above, in some embodiments of the present disclosure, each polarization camera module may include only three polarization cameras. In some embodiments, the three individual polarization cameras may include corresponding linear polarizers with linear polarization filters at 0°, 45°, and 90° or at 0°, 60°, and 120°. In some embodiments, the three individual polarization cameras of the first polarization camera module have corresponding color filters to transmit red (R), green (G), and blue (B) light having corresponding first spectral profiles, and the three individual polarization cameras of the second polarization camera module may have corresponding color filters to transmit red (R′), green (G′), and blue (B′) light having second spectral profiles that are different from the first spectral profile (e.g., where each of the second spectral profiles is shifted from corresponding first spectral profiles by 10 to 20 nm).

[00123] In addition, while some embodiments of the present disclosure are described above with respect to color filters that transmit different portions of the visible electromagnetic spectrum, embodiments of the present disclosure are not limited thereto, and may also include the use of color filters that selectively transmit light in other portions of the electromagnetic spectrum, such as infrared light or ultraviolet light.

[00124] In some embodiments of the present disclosure, the two different polarization camera modules of the multi-spectral stereo polarization camera system include polarization cameras that are configured to capture polarization raw frames of different polarization states (e.g., different polarization angles), such as using a polarization mask as shown in FIG. 1B or a mechanically or electronically controllable polarizing filter. According to some embodiments of the present disclosure, each polarization camera further includes a color filter configured to filter light in a plurality of colors in accordance to a mosaic pattern such as a Bayer pattern, where each polarization camera may have a different color filter to enable multi-spectral or hyper-spectral capture. For example, in some embodiments, a first polarization camera of a stereo pair includes a first color filter configured to capture light in the R, G1, G2, B spectra (or R, G, B spectra), as described above, and a second polarization camera of the stereo pair includes a second color filter configured to capture light in the R′, G1′, G2′, B′ spectra (or R′, G′, B′), as described above.

[00125] Some aspects of embodiments of the present disclosure relate to capturing multi-spectral scenes using hardware arrangements such as those discussed above by determining the spectral basis functions for representation. By estimating the spectral power distribution of scene illumination and using the spectral reflectivity function of the Macbeth color chart, it is possible to simulate a set of basis functions B representing that illumination. This becomes especially feasible when estimating the spectral profile of natural sunlight for outdoor use as is typically the case with multispectral imaging for geo-spatial applications. Once the spectral basis functions are determined, it is straightforward to determine the spectral coefficients for each scene by simply solving for w (weights) in the following equation

1

$$p = TS = TBw \tag{13}$$

where, p represents the pixel values in the different spectral (color) channels (e.g., eight different color channels R, G1, G2, B, R′, G1′, G2′, and B′), T represents the spectral responsivities of the various spectral channels (e.g., the captured values), S is the illumination source, which can be decomposed into a spectral basis, B represents the spectral basis functions, and w represents the coefficients for the basis functions.

[00126] Accordingly, applying equation (13) above enables computation of per-pixel polarization information as well as spectral information.

[00127] The multi-spectral or hyper-spectral information computed from multi-spectral hardware, such as that described above, maybe supplied as inputs to other object detection or instance segmentation algorithms (e.g., using convolutional neural networks that are trained or retrained based on labeled multi-spectral polarization image training data), or may be supplied as inputs to classical computer vision algorithms (e.g., such as for detecting the depth of surfaces based on parallax shift of multi-spectral and polarization cues) for detecting the presence of objects in the scenes imaged by stereo multi-spectral polarization camera systems according to embodiments of the present disclosure.

[00128] While some embodiments of the present disclosure as described above relate to multi-viewpoint multi-spectral polarization imaging using a stereo camera system (e.g., a stereo pair), embodiments of the present disclosure are not limited thereto. For example, in some embodiments of the present disclosure, a multi-spectral camera system (e.g., using a camera system configured to capture six or more different spectra, such as R, G, B, R′, G′, and B′, as discussed above) sweeps across multiple viewpoints over time, such as when an object of interest is located on a conveyor belt that passes through the field of view of the camera system, or where the camera system moves across the field of view of the object of interest.

[00129] As one example, for applications in satellite imaging one has the added advantage of viewing the scene from multiple angles that are highly correlated. The systematic way in which satellites move in straight lines above a given point on the ground allows satellites to obtain highly correlated multi-spectral and polarization data of the surfaces of the ground for each viewing angle across a wide range of viewing angles. Accordingly, in some embodiments of the present disclosure, a processing system 100 determines, for each point on the ground, the optimal angle at which the degree of polarization (DOLP) signal is strongest, thereby providing a strong correlation as to its surface orientation. See, e.g., equations (2) and (4). In addition, because specularity is generally highly viewpoint dependent, most of the views of a given surface will be non-specular, such that equation (2) may be sufficient to compute the orientation of the surface being imaged, without needing to select between the non-specular (or diffuse) equation versus the specular equation (4).

[00130] In addition, satellite imaging enables the capture of images of objects captured from very different viewpoints. This large baseline enables the estimation of coarse distances of ground-based objects by leveraging multispectral imaging with polarization and parallax shifts due to the large changes in position. Detecting these coarse distances provides information for disaster management, power transmission line monitoring, and security. For example, utility companies are concerned with the uncontrolled growth of vegetation in and around power transmission and distribution lines due to risks of fire or damage to the transmission lines. By imaging the areas around the power lines from different viewpoints, detecting the parallax shift of the objects when viewed from different viewpoints enables estimations of the surface height of the vegetation and the height of the transmission and distribution lines. Accordingly, this enables the automatic detection of when ground vegetation reaches critical thresholds with respect to proximity of said lines with respect to vegetation growth. To monitor such data both at day and night, some embodiments of the present disclosure relate to fusing polarization data with thermal sensors (e.g., infrared sensors) to provide clear heat signatures irrespective of illumination conditions.

[00131] *Image Segmentation using Polarimetric Cues*

[00132] Some aspects of embodiments of the present disclosure relate to performing instance segmentation using polarimetric cues captured in accordance with embodiments of the present disclosure. Some techniques for performing instance segmentation using polarimetric cues are described in more detail in U.S. Provisional Patent Application No. 62/942,113, filed in the United States Patent and Trademark Office on November 30, 2019 and U.S. Provisional Patent Application No. 63/001,445, filed in the United States Patent and Trademark Office on March 29, 2020.

[00133] FIG. 6A is a block diagram of processing circuit 100 for computing surface characterization outputs based on polarization data according to one embodiment of the present invention. FIG. 6B is a flowchart of a method 600 for performing surface characterization based on input images to compute a surface characterization output according to one embodiment of the present invention.

[00134] According to various embodiments of the present disclosure, the processing circuit 100 is implemented using one or more electronic circuits

Date Reçue/Date Received 2022-07-05

configured to perform various operations as described in more detail below. Types of electronic circuits may include a central processing unit (CPU), a graphics processing unit (GPU), an artificial intelligence (AI) accelerator (e.g., a vector processor, which may include vector arithmetic logic units configured efficiently perform operations common to neural networks, such dot products and softmax), a field programmable gate array (FPGA), an application specific integrated circuit (ASIC), a digital signal processor (DSP), or the like. For example, in some circumstances, aspects of embodiments of the present disclosure are implemented in program instructions that are stored in a non-volatile computer readable memory where, when executed by the electronic circuit (e.g., a CPU, a GPU, an AI accelerator, or combinations thereof), perform the operations described herein to compute a characterization output 20 from input polarization raw frames 18. The operations performed by the processing circuit 100 may be performed by a single electronic circuit (e.g., a single CPU, a single GPU, or the like) or may be allocated between multiple electronic circuits (e.g., multiple GPUs or a CPU in conjunction with a GPU). The multiple electronic circuits may be local to one another (e.g., located on a same die, located within a same package, or located within a same embedded device or computer system) and/or may be remote from one other (e.g., in communication over a network such as a local personal area network such as Bluetooth®, over a local area network such as a local wired and/or wireless network, and/or over wide area network such as the internet, such a case where some operations are performed locally and other operations are performed on a server hosted by a cloud computing service). One or more electronic circuits operating to implement the processing circuit 100 may be referred to herein as a computer or a computer system, which may include memory storing instructions that, when executed by the one or more electronic circuits, implement the systems and methods described herein.

[00135] As shown in FIG. 6A, in some embodiments, a processing circuit 100 includes a feature extractor or feature extraction system 700 and a predictor 750 (e.g., a classical computer vision prediction algorithm and/or a trained statistical model such as a trained neural network) configured to compute a prediction output 20 (e.g., a statistical prediction) regarding surface characteristics of objects based on the output of the feature extraction system 700. Various embodiments of the present disclosure are described herein in the context of surface characterization in circumstances where surface features may be optically challenging to detect, and/or where polarization-based imaging techniques provide information on surface normal that may otherwise be difficult to obtain, embodiments of the present disclosure are not limited thereto. For example, some aspects of embodiments of the present disclosure may be applied to techniques for characterizing the surfaces of objects made of materials or have surface characteristics that are optically challenging to detect, such as surfaces of translucent objects, multipath inducing objects, objects that are not entirely or substantially matte or Lambertian, and/or very dark objects. These optically challenging objects include objects and surface characteristics thereof that are difficult to resolve or detect through the use of images that are capture by camera systems that are not sensitive to the polarization of light (e.g., based on images captured by cameras without a polarizing filter in the optical path or where different images do not capture images based on different polarization angles). For example, these surface characteristics may have surface appearances or colors that are very similar to the surfaces on which the characteristics appear (e.g., dents have the same color as the underlying material and scratches on transparent materials such as glass may also be substantially transparent). In addition, while embodiments of the present disclosure are described herein in the context of detecting optically challenging surface characteristics, embodiments of the present disclosure are not limited to detecting only optically challenging surface defects. For example, in some embodiments, a predictor 750 is configured (e.g., a statistical model is trained using training data) to detect both surface characteristics that are optically challenging as well as surface characteristics that are robustly detectable without using polarization information.

[00136] Polarization may be used to detect surface characteristics or features that would otherwise be optically challenging when using intensity information (e.g., color intensity information) alone. For example, polarization information can detect changes in geometry and changes in material in the surfaces of objects. The changes in material (or material changes), such as boundaries between different types of materials (e.g., a black metallic object on a black road or a colorless liquid on a surface may both be substantially invisible in color space, but would both have corresponding polarization signatures in polarization space), may be more visible in polarization space because differences in the refractive indexes of the different materials cause changes in the polarization of the light. Likewise, differences in the specularity of various materials cause different changes in the polarization phase angle of rotation, also leading to detectable features in polarization space that might otherwise be optically challenging to detect without using a polarizing filter. Accordingly, this causes contrast to appear in images or tensors in polarization representation spaces, where corresponding regions of tensors computed in intensity space (e.g., color representation spaces that do not account for the polarization of light) may fail to capture these surface characteristics (e.g., where these surface characteristics have low contrast or may be invisible in these spaces). Examples of optically challenging surface characteristics include: the particular shapes of the surfaces (e.g., degree of smoothness and deviations from ideal or acceptable physical design tolerances for the surfaces); surface roughness and shapes of the surface roughness patterns (e.g., intentional etchings, scratches, and edges in the surfaces of transparent objects and machined parts), burrs and flash at the edges of machined parts and molded parts; and the like. Polarization would also be useful to detect objects with identical colors, but differing material properties, such as scattering or refractive index.

[00137] In addition, as discussed above, polarization may be used to obtain the surface normals of objects based on the degree of linear polarization (DOLP) ρ and the angle of linear polarization (AOLP) ϕ computed from the polarization raw frames based on, for example, equations (2), (3), (4), and (5). These surface normal, in turn, provide information about the shapes of the surfaces.

[00138] As shown in FIG. 6B and referring, for example, to FIG. 1B, in operation 610 the processing circuit 100 captures polarization raw frames 18 of surfaces in a scene 1. For example, in some embodiments, the processing circuit 100 controls one or more polarization cameras 10 (e.g., one or more individual polarization cameras, which may be organized into polarization camera arrays and/or stereo polarization camera systems that include multiple polarization camera modules) to capture polarization raw frames 18 depicting a surfaces of object in a scene 1.

[00139] FIG. 7A is a block diagram of a feature extractor 700 according to one embodiment of the present invention. FIG. 7B is a flowchart depicting a method according to one embodiment of the present invention for extracting features from polarization raw frames. In the embodiment shown in FIG. 7A, the feature extractor 700 includes an intensity extractor 720 configured to extract an intensity image I 52 in an intensity representation space (e.g., in accordance with equation (7), as one example of a non-polarization representation space) and polarization feature extractors 730 configured to extract features in one or more polarization representation spaces. In some embodiments of the present disclosure, the intensity extractor 720 is omitted and the feature extractor does not extract an intensity image I 52. In the embodiment shown in FIG. 7A, the features extracted in polarization representation spaces (e.g., DOLP ρ and AOLP ϕ) are supplied to a surface normals calculator 780 to compute surface normals 58 of objects in the scene

[00140] As shown in FIG. 7B, the extraction of polarization images in operation 650 may include extracting, in operation 651, a first tensor in a first polarization

1

representation space from the polarization raw frames from a first Stokes vector. In operation 652, the feature extractor 700 further extracts a second tensor in a second polarization representation space from the polarization raw frames. For example, the polarization feature extractors 730 may include a DOLP extractor 740 configured to extract a DOLP ρ image 54 (e.g., a first polarization image or a first tensor in accordance with equation (8) with DOLP as the first polarization representation space) and an AOLP extractor 760 configured to extract an AOLP ϕ image 56 (e.g., a second polarization image or a second tensor in accordance with equation (9), with AOLP as the second polarization representation space) from the supplied polarization raw frames 18. In addition, in various embodiments, the feature extraction system 700 extracts two or more different tensors (e.g., n different tensors) in two or more representation spaces (e.g., n representation spaces), where the n-th tensor is extracted in operation 614. As discussed above, in some embodiments of the present disclosure, the polarization feature extractors 730 extract polarization features in polarization representation spaces including both linear polarization representation spaces (e.g., tensors in the aforementioned AOLP and DOLP representation spaces extracted from polarization raw frames captured with a linear polarizing filter) and circular polarization representation spaces (e.g., tensors extracted from polarization raw frames captured with a circular polarizing filter). In various embodiments, the representation spaces include, but are not limited to, polarization representation spaces.

[00141] The polarization representation spaces may include combinations of polarization raw frames in accordance with Stokes vectors. As further examples, the polarization representations may include modifications or transformations of polarization raw frames in accordance with one or more image processing filters (e.g., a filter to increase image contrast or a denoising filter). The feature maps 52, 54, and 56 in first polarization representation spaces may then be supplied to a predictor 750 for detecting surface characteristics based on the feature maps 50.

[00142] While FIG. 7B illustrates a case where two or more different tensors are extracted from the polarization raw frames 18 in more than two different representation spaces, embodiments of the present disclosure are not limited thereto. For example, in some embodiments of the present disclosure, exactly one tensor in a polarization representation space is extracted from the polarization raw frames 18. For example, one polarization representation space of raw frames is AOLP ϕ and another is DOLP ρ (e.g., in some applications, AOLP may be sufficient for detecting surface characteristics of transparent objects or surface characteristics of other optically challenging objects such as translucent, non-Lambertian, multipath inducing, and/or non-reflective objects).

[00143] Furthermore, as discussed above with respect to FIG. 7A, in some embodiments of the present disclosure, one or more feature vectors are computed based on features computed from other representation spaces. In the particular example shown in FIG. 7A, the surface normals calculator 780 computes surface normals of surfaces in the imaged scene 1 in surface normals space (e.g., azimuth angle θ_a and zenith angle θ_z) based on the computed AOLP ϕ and DOLP ρ tensors. In some embodiments, the surface normal are encoded using Cartesian coordinates (e.g., a three-dimensional vector indicating a direction of the surface normal). the computed surface normals 58 may be included among the features 50 extracted by the feature extractor 700.

[00144] Accordingly, extracting features such as polarization feature maps, polarization images, and/or surface normals from polarization raw frames 18 produces first tensors 50 from which optically challenging surface characteristics may be detected from images of surfaces of objects under inspection. In some embodiments, the first tensors extracted by the feature extractor 700 may be explicitly derived features (e.g., hand crafted by a human designer) that relate to underlying physical phenomena that may be exhibited in the polarization raw frames (e.g., the calculation of AOLP and DOLP images in linear polarization spaces and the calculation of tensors in circular polarization spaces, as discussed above). In some additional embodiments of the present disclosure, the feature extractor 700 extracts other non-polarization feature maps or non-polarization images, such as intensity maps for different colors of light (e.g., red, green, and blue light) and transformations of the intensity maps (e.g., applying image processing filters to the intensity maps). In some embodiments of the present disclosure, the feature extractor 700 further computes surface normals of surfaces depicted by the polarization raw frames, in accordance with shape from polarization techniques, as described above. In some embodiments of the present disclosure the feature extractor 700 may be configured to extract one or more features that are automatically learned (e.g., features that are not manually specified by a human) through an end-to-end supervised training process based on labeled training data. In some embodiments, these learned feature extractors may include deep convolutional neural networks, which may be used in conjunction with traditional computer vision filters (e.g., a Haar wavelet transform, a Canny edge detector, a depth-from-stereo calculator through block matching, and the like).

1

[00145] In some embodiments of the present disclosure, the predictor 750 implements one or more classical computer vision algorithms (e.g., depth from stereo using block matching) based on the first tensors 50.

[00146] In some embodiments of the present disclosure, the predictor 750 implements an image segmentation algorithm, such as by including a trained convolutional neural network. Image segmentation refers to labeling pixels based on the class of object contained within the pixel. Traditional algorithms are adapted to conventional red-green-blue (RGB) channel or gray channel (hereafter RGB/gray) information, using RGB/gray image gradients to enforce decision boundaries. For many industrial applications, RGB/gray image gradients might not be present. One example is a photograph of a black car on a black road. It is difficult to segment this scene to separate the car from the road. However, by capturing both RGB/gray gradients and polarimetric gradients using polarization camera systems in accordance with embodiments of the present disclosure, computer vision systems can be trained, based on the combined RGB/gray and polarization cues (e.g., including surface normals computed from the polarization cues), to perform semantic segmentation of images in conditions that would be optically challenging when using only RGB/gray images. As used herein, convolutional neural networks that are configured to take polarization cues as input (e.g., features in polarization feature spaces, as extracted from polarization raw frames by the feature extractor 700, such as AOLP ϕ, DOLP ρ, and/or surface normals computed based on AOLP and DOLP) will be referred to herein as Polarized CNNs, of which the Polarized Mask R-CNN described above and in U.S. Provisional Patent Application No. 62/942,113, filed in the United States Patent and Trademark Office on November 30, 2019 and U.S. Provisional Patent Application No. 63/001,445, filed in the United States Patent and Trademark Office on March 29, 2020.

[00147] *Augmenting 3-D Surface Reconstruction with Polarization Imaging*

[00148] Some aspects of embodiments of the present disclosure relate to recover high quality reconstructions of closed objects. In some embodiments of the present surface reconstruction is used in conjunction with high quality three-dimensional (3-D) models of the objects, such as computer-aided-design (CAD) models of the objects to be scanned to resolve ambiguities arising from a polarization-based imaging process. Previous attempts have devised methods for unknown geometry without having access to CAD models.

[00149] Capturing a high quality 3-D reconstruction of a physical object for which a high-quality 3-D computer model already exists is important in a variety of contexts, such as quality control in the fabrication and/or manufacturing of objects.

For example, in the case of additive manufacturing or 3-D printing, a designer may create a 3-D model of an object and supply the 3-D model to a 3-D printer, which fabricates a physical object based on the 3-D model. During or after the 3-D printing process, the physical object fabricated by the 3-D printer may be scanned using a stereo polarization camera system according to some embodiments of the present disclosure, and the captured polarization data may be used to assist in the 3-D reconstruction of the surfaces of the physical object. This 3-D reconstruction can then be compared, in software, to the designed 3-D model to detect defects in the 3-D printing process. Similar techniques may be applied to other manufacturing processes, such as for creating 3-D reconstructions of the shapes of objects created through other manufacturing processes such as injection molding, die-casting, bending, and the like.

[00150] As one example, a stereo polarization camera system, such as that described above with respect to FIG. 1E, is used to image an object that is intended to be reconstructed in 3-D, e.g., to create a 3-D model of the object automatically from the captured polarization raw frames. Due to practical manufacturing constraints and/or defects in the manufacturing process, the surface of the object may have sparse irregularities, and may not be ideally smooth. These irregularities may appear as high frequency variations on the surface. High frequency variations (HFV) appear due to 3 scenarios:

[00151] First, there could be regions on the object surface that have valid high-frequency variations (e.g., designed and intended to be present). For example, when creating a replica of a Greek bust or statue, details near the eyes and hair of the scanned 3-D model may also be present in the high-quality 3-D model that was used to guide the fabrication of the physical object.

[00152] Second, there may be regions on the object surface that have high-frequency variations due to blemishes, defects, or other damage on the surface. For example, in the case of 3-D printing or additive manufacturing, high frequency patterns may arise due to the layer-wise manufacturing process, causing a “steeped” appearance to surfaces of the object. As another example, an injection molding process may leave seams or flashing in the produced object where the two parts of the mold meet. These details are not reflected in the high-quality 3-D model.

[00153] Third, combinations of the first and second forms of high frequency variations may occur physically close to one another (e.g., flashing may appear near the hair of the replica of the bust, thereby causing additional lines to appear in the hair).

1

[00154] High-frequency variations due to details are desirable on the real object, while the HFVs due to irregularities are not. However, it is important to be able to recover both of these kinds of HFVs in the 3D reconstruction for the purposes of inspection and profilometry. While some of these HFV details as well as irregularities may not be recovered by a commercially available 3D scanner (due to poor resolution arising from quantization error & other noise sources), embodiments of the present disclosure are able to handle these cases, as discussed in more detail below. Some exemplary implementations may make use of an additional structured lighting projector device to illuminate the object if the object has no visual features. Some embodiments of the present disclosure relate to the use of passive illumination (e.g., based on ambient lighting in the scene).

[00155] FIG. 8A is an illustration of a Greek bust statue being scanned by an exemplary implementation of the imaging setup proposed in this invention. Three types of High-Frequency-Variation (HFV) details are annotated (801A: desirable details such as hair and eyes; 801B: undesirable blemishes & defects near the cheek & nose; and 801C: a combination of cases A & B in close proximity with each other). These HFVs may not be recovered using standard 3D imaging techniques. Aspects of embodiments of the present invention relate to handling all of these cases. FIG. 8B is a flowchart of a method for 3-D surface reconstruction using polarization according to one embodiment of the present disclosure.

[00156] In some embodiments of the present disclosure, in operation 810, polarization raw frames 18 are captured of an object from multiple viewpoints using, for example, a stereo polarization camera system as describe above with respect to FIG. 1E. A set of four separate polar-angle images (0, 45, 90, 135) can be extracted from each of the raw images acquired. These may be denoted as P_{C1} and P_{C2}. In exemplary implementations of this setup, the cameras may be in housed in standard stereo configurations (optical axes parallel to each other), or other configurations (e.g., where the optical axes intersect with each other).

[00157] In operation 820, degree and angle of linear polarization (DOLP ρ and AOLP ϕ) may be computed from Stokes vector formulation for both cameras using P_{C1} and P_{C2} as described above. These may be denoted as ρ_{C1}, ϕ_{C1}, ρ_{C2}, and ϕ_{C2}. In operation 830, surface normals (e.g., Zenith θ_z and Azimuth θ_a) from polarization are computed using shape from polarization (SFP) using DOLP ρ and AOLP ϕ as discussed above with respect to equations (2), (3), (4), and (5) for both cameras C1 and C2 (e.g., based on polarization raw frames P_{C1} and P_{C2}). These surface normal from the two viewpoints may be denoted as $N_{Pol_{C1}}$ and $N_{Pol_{C2}}$.

[00158] However, these surface normals suffer from Azimuthal θ_a ambiguity by an angle of π, which can be disambiguated and corrected by using the CAD reference model as a constraint (e.g., by selecting the azimuthal angle θ_a that results in a surface that has the smaller distance or error with respect to the reference model). Accordingly, low-frequency noise (e.g., ambiguity by an angle of π) can be resolved using the reference model.

[00159] Depending on whether the object is dielectric or non-dielectric (taking cues from the strength of DOLP), an appropriate DOLP computation model may be employed to estimate the zenith angle as discussed above. In some embodiments, the material may be assumed to be dielectric with a refractive index of 1.5 because the refractive index of dielectrics is typically in the range [1.3, 1.6], and that this variation causes negligible change in DOLP ρ. In cases where the material is non-dielectric, the accuracy of the estimated zenith angle would suffer from refractive distortion. Refractive error in zenith is a low-frequency phenomenon and therefore may also be corrected by leveraging the reference model to use as a prior for resolving the refractive error.

[00160] Normals $N_{Pol_{C1}}$ and $N_{Pol_{C2}}$ may both independently be integrated over a sample space (Ω) to recover the entire surface off the object or a part of the surface of the object (e.g., the surface normals indicate the slope of the surfaces of the object and therefore integrating over the slopes, after accounting for the direction of the normal versus the orthogonal direction of the slope, recovers the underlying shape of the object). The surface recovered from such integration should match the shape constrained by the CAD reference model. Differences between the surface recovered from integration and the reference model may indicate defective portions of the physical object.

[00161] In addition to only relying on the CAD model for resolving ambiguities and errors in 3-D reconstruction based on polarization data from one polarization camera (or one polarization camera array), some aspects of embodiments of the present disclosure relate to further improving the quality of the 3-D reconstruction by enforcing view-point consistency between the cameras of the stereo polarization camera system.

[00162] Accordingly, while some embodiments of the present disclosure relate to computing estimated surface normal as described above through operation 830 shown in FIG. 8B, some embodiments of the present disclosure relate to further refining the estimated surface normals. Still referring to FIG. 8B, in operation 840, the high-quality CAD reference model is aligned to orientation of the physical object based on visual keypoints that are estimated on the object in the polarization raw frames captured by the two cameras P_{C1} and P_{C2}. These keypoints are correlated with the same set of keypoints in the CAD reference model to obtain the six degree of freedom (6DoF) pose of the object with respect to the cameras using Perspective-N-Point (PnP) (see, e.g., Fischler, M. A.; Bolles, R. C. (1981). "Random Sample Consensus: A Paradigm for Model Fitting with Applications to Image Analysis and Automated Cartography". *Communications of the ACM*. **24** (6): 381–395.) and/or random sample consensus (RANSAC). Use of multiple registered cameras at different viewpoints enables a more accurate pose reconstruction than using a single camera having a single viewpoint, although embodiments of the present disclosure are not limited thereto and single view PnP may also be used. The CAD reference model may be transformed into the camera spaces corresponding to the different camera modules of the stereo camera system (e.g., transforming the pose of the CAD reference model to the pose of the actual physical object with respect to the camera system), thereby aligning the reference model to the physical object. In the case of two cameras, this may be denoted as CAD_{C1} and CAD_{C2}. Surface normals are then extracted from CAD_{C1} and CAD_{C2} (e.g., based on the orientations of surfaces with respect to the virtual cameras corresponding to the camera modules of the stereo camera system). These reference surface normals may be denoted as $N_{CAD_{C1}}$ and $N_{CAD_{C2}}$.

[00163] The transformed CAD reference model can then be used as a guidance constraint to correct high frequency azimuthal π ambiguity as well as the low frequency scaling error in zenith due to refractive distortion. Corrected normals will have consistency between the 2 cameras due to Multiview PnP, making this approach more robust. In more detail, in operation 850, the estimated normals $N_{Pol_{C1}}$ and $N_{Pol_{C2}}$ computed (in operation 830) from the polarization raw frames P_{C1} and P_{C2} from the two cameras are then corrected to compute corrected normals $Corrected\_N_{Pol_{C1}}$ and $Corrected\_N_{Pol_{C2}}$. The relative poses between the corrected normals should be consistent with the relative pose between the cameras ($N_{CAD_{C1}}$ and $N_{CAD_{C2}}$). This imposes additional pose-consistency constraints, thereby making SFP-normal correction more robust in general, and specifically in the case of fronto-parallel facets where Zenith θ_z is close to 0° (or 0 radians), which tend to have noisy estimated normal due to poor strength of the DOLP ρ along the viewing direction. However, any given facet will be less likely to be fronto-parallel to both the camera modules of a stereo polarization camera system, given the spacing (or baseline) between the camera modules. Accordingly, regions with higher DOLP may be voted and selected from across the multiple cameras to recover more robust surface normal for surfaces that are fronto-parallel to a subset of the camera modules.

[00164] In some circumstances, specularity causes problems in surface reconstruction because the surface texture information is lost due to oversaturation in the intensity of the image. This causes estimated normals on a specular patch to be highly noisy. According to some embodiments of the present disclosure, the polarization camera system includes multiple cameras (e.g., two or more) that are viewing overlapping regions of the scene from multiple viewpoints (e.g., a stereo polarization camera system) spaced apart by a baseline . Specularity is generally a highly viewpoint dependent issue. That is, specularity is less likely to be observed by all the cameras in a setup such as the arrangement shown in FIG. 1E, where different cameras have different viewpoints of surfaces of an object.

[00165] In more detail, some aspects of embodiments of the present disclosure relate to automatically recovering robust surface normals, even in highly specular materials, by imaging the surfaces from multiple viewpoints. Under most lighting conditions, it is highly unlikely that any given patch of a surface will appear specular to all of the cameras in a stereo multi-view camera system.

[00166] Accordingly, in some embodiments of the present disclosure, a voting mechanism may be employed to reject normals from a specular patch observed in a particular camera, while selecting the normals from the other cameras for the particular patch, that are more likely to be consistent with each other as well as the CAD model. For example, surface normals may be computed based on the polarization raw frames captured from each of the polarization camera modules in the stereo polarization camera array. If the surface normals computed based on the polarization raw frames are highly inconsistent with one another (e.g., more than a threshold angular distance apart), then the computed surface normals that are closest to the surface normals of the reference model are assumed to be the correct values.

[00167] In other embodiments of the present disclosure, specular patches may be detected automatically by identifying saturated pixels in the polarization raw frames. The saturation of the pixels is used to suggest that the particular patch may be observing specularity and therefore information in that region may be inaccurate.

[00168] In still other embodiments of the present disclosure, the stereo camera system includes more than two polarization camera modules (e.g., three or more polarization camera modules) which image the surfaces of the objects from different viewpoints. Accordingly, a voting mechanism may be employed, in which the surface normals computed based on the polarization raw frames captured by the various cameras are clustered based on similarity (after transforming the surface normals to correspond to a same frame of reference, such as one of the polarization camera modules). Because most of the polarization camera modules are unlikely to observe specularity, most of the calculated normals should be consistent, within an error range. Accordingly, the clustering process may identify outliers in the calculated surface normals, as caused by the specular artifacts.

[00169] A pseudocode description of an algorithm for normals correction based on voting with a CAD reference model prior is presented in more detail as follows. As notation:

[00170] **N_P_C1** - shape-from-polarization (SFP) normals in Camera1

[00171] **N_P_C2** - SFP normals in Camera2

[00172] **N_CAD_C1** - Normals in CAD reference model aligned with the object pose with respect to Camera1

[00173] **N_CAD_C2** - Normals in CAD reference model aligned with the object pose with respect to Camera2

[00174] **Trans_C2_C1** - Camera2's pose with respect to Camera1 obtained through extrinsic and intrinsic camera calibration (e.g., determined by imaging calibration targets visible to both Camera1 and Camera2)

[00175] **Trans_CAD_C1** - Transform used to align CAD reference model with the object in Camera1 image space obtained through multiview P*n*P

[00176] **Trans_CAD_C2** - Transform used to align CAD reference model with the object in Camera2 image space obtained through multiview P*n*P

[00177] **(~)** - Consistency operator

[00178] The consistency operator (~) may be modeled as a distance metric (e.g., a cosine similarity based angular distance metric) computed between the normals being compared for consistency. If the angular distance is less than a threshold, the normals being compared are consistent with each other, else not (!~). The normals being compared are transformed into the same coordinate frame (master-camera or Camera1 image space in this case) using the transforms listed above before applying the consistency operator (~).

[00179] **Pseudocode implementation of voting based on CAD reference model prior:**

For each pixel normal in N_P_C1 (master camera image space, in this case Camera1):

 # Case 1: SFP normals in both cameras are consistent with CAD (No ambiguity)

if (N_P_C1 ~ N_P_C2) && (N_P_C1 ~ N_CAD_C1) && (N_P_C2 ~ N_CAD_C2) then:

retain N_P_C1 or Trans_C2_C1*(N_P_C2) depending on which of the 2 camera normals are more aligned (consistent) with the CAD model

 # Case 2: SFP normals in the 2 cameras are consistent with each other, but both are inconsistent with CAD normal (HFV Blemish/defect on surface)

else if (N_P_C1 ~ N_P_C2) && (N_P_C1 !~ N_CAD_C1) && (N_P_C2 !~ N_CAD_C2) then:

retain N_P_C1

 # Case 3: SFP normals in only one of the cameras are consistent with CAD (specularity / fronto parallel issue in the other camera)

else if (N_P_C1 ~ N_CAD_C1) && (N_P_C2 !~ N_CAD_C2) then:

retain N_P_C1 *#specularity / fronto parallel facet issue in Camera2*

else if (N_P_C2 ~ N_CAD_C2) && (N_P_C1 !~ N_CAD_C1) then:

retain Trans_C2_C1*(N_P_C2) *#specularity / fronto parallel facet issue in Camera1*

 # Case 4: SFP normals in both cameras are inconsistent with each other, as well as with CAD

else if (N_P_C1 !~ N_CAD_C1) && (N_P_C2 !~ N_CAD_C2) && (N_P_C1 !~ N_P_C2) then:

retain N_CAD_C1

[00180] In some embodiments of the present disclosure, the corrected surface normals $Corrected\_N_{Pol_{C1}}$ and $Corrected\_N_{Pol_{C2}}$ are integrated over sample space (Ω) to synthesize a 3-D reconstruction of the object imaged by the stereo polarization camera system.

[00181] While the embodiments discussed above relate to the 3-D reconstruction of 3-D objects based on a high-quality 3-D model such as a CAD design model, some aspects of embodiments of the present disclosure further relate to 3-D reconstruction of generally flat surfaces or surfaces having known, simple geometry, using multi-view polarized camera system such as that shown in FIG. 1E. The simple geometry case may apply to circumstances where the objects to be analyzed are unknown, but can be approximated using, for example, flat planes, spheres, and other simple parametric curves of known geometry. For example, flat planes may apply to many surfaces in an environment for a self-driving vehicle, such as the surface of a road, as well as walls and signage. In addition, depending on the resolution and/or accuracy demanded by a particular application, many surfaces may be approximated as being locally flat.

[00182] FIG. 9A is an illustration of a flat surface of refractive index n, being scanned by an exemplary implementation of the imaging setup according to one embodiment of the present invention. For particular applications in profilometry and inspection, this surface is examined for its smoothness. Ideally, this surface is desired to be smooth. Practically, due to defects/wear and tear, there may be sparse irregularities 901 at random locations on this otherwise flat surface 902. These irregularities manifest as High-Frequency-Variation (HFV) details that may not be recovered using standard 3D imaging techniques due to noise and poor resolution. However, embodiments of the present invention are able to recover these HFV irregularities leveraging polarization in conjunction with flatness and multi-view constraints.

[00183] Accordingly, for the sake of discussion, some embodiments of the present disclosure relate to detecting random, sparse irregularities on an otherwise substantially smooth surface (e.g., a substantially flat surface). As a motivating example, embodiments of the present disclosure may be used to detect potholes in a road using a stereo polarization camera system, such that a self-driving vehicle can avoid those potholes, as practical based on traffic conditions. As another motivating example, embodiments of the present disclosure may be used to detect surface defects in surfaces with generally simple geometries, such as detecting surface irregularities in the smoothness of a pane of glass or in a sheet of metal.

[00184] In some embodiments of the present disclosure, a multi-view polarization camera system may further include a structured light projector 903 configured to project patterned light onto a scene to provide additional detectable surface texture for the depth from stereo processes to match between views (e.g., using block matching) for measuring parallax shifts. In some circumstances, the structured light projector is configured to project infrared light and the camera system includes cameras configured to detect infrared light along with light in other spectral bands. Any following analysis of the surfaces may then be performed based on the data collected in the other spectral bands such that the projected pattern is not inadvertently detected as defects in the surface of the material.

[00185] FIG. 9B is a flowchart of a method for 3-D surface reconstruction of flat or geometrically simple surfaces using polarization according to one embodiment of the present disclosure.

[00186] In a manner similar to that described above, in some embodiments of the present disclosure, in operation 910, polarization raw frames 18 are captured of a

1

scene (e.g., including substantially flat or smooth surfaces) from multiple viewpoints using, for example, a stereo polarization camera system as describe above with respect to FIG. 1E. A set of four separate polar-angle images (0, 45, 90, 135) can be extracted from each of the raw images acquired. These may be denoted as P_{C1} and P_{C2}. In exemplary implementations of this setup, the cameras may be in housed in standard stereo configurations (optical axes parallel to each other), or other configurations (e.g., where the optical axes intersect with each other).

[00187] In operation 920, degree and angle of linear polarization (DOLP ρ and AOLP ϕ) are computed from Stokes vector formulation for both cameras using P_{C1} and P_{C2} as described above. These may be denoted as ρ_{C1}, ϕ_{C1}, ρ_{C2}, and ϕ_{C2}.

[00188] In operation 930, surface normals (e.g., Zenith θ_z and Azimuth θ_a) from polarization can be obtained using shape from polarization (SFP) using DOLP ρ and AOLP ϕ as discussed above with respect to equations (2), (3), (4), and (5) for both cameras C1 and C2 (e.g., based on polarization raw frames P_{C1} and P_{C2}). Depending on whether the object is dielectric or non-dielectric (taking cues from the strength of DOLP), an appropriate DOLP computation model may be employed to estimate the zenith angle as discussed above. In some embodiments, the material may be assumed to be dielectric with a refractive index of 1.5 because the refractive index of dielectrics is typically in the range [1.3, 1.6], and that this variation causes negligible change in DOLP ρ. In cases where the material is non-dielectric, the accuracy of the estimated zenith angle would suffer from refractive distortion.

[00189] These surface normal from the two viewpoints may be denoted as $N_{Pol_{C1}}$ and $N_{Pol_{C2}}$. However, these surface normals suffer from Azimuthal θ_a ambiguity by an angle of π, which can be disambiguated and corrected by using the coarse depth map as a constraint (e.g., by selecting the azimuthal angle θ_a that results in a surface that has the smaller distance or error with respect to the reference model). Accordingly, low-frequency noise (e.g., ambiguity by an angle of π) can be resolved using the coarse depth map created from the stereo view of the scene.

[00190] In addition, in operation 940, a coarse depth map (CDM) is computed based on the parallax shift between pairs of cameras in the stereo polarization camera system, based on depth-from-stereo approaches (e.g., where larger parallax shifts indicate surfaces that are closer to the camera system and smaller parallax shifts indicate that surfaces are farther away). As noted above, in some embodiments, the stereo polarization camera system includes a structured light illumination system, which may improve the matching of corresponding portions of the images when the surfaces do not have intrinsic texture or other visual features. In operation 940, the computed coarse depth map is also aligned to the image spaces corresponding the viewpoints $C1$ and $C2$ (e.g., using the relative pose and the extrinsic matrices from the camera calibration), where the coarse depth maps corresponding to these image spaces are denoted CDM_{C1} and CDM_{C2}.

[00191] In operation 950, the estimated normals as $N_{Pol_{C1}}$ and $N_{Pol_{C2}}$ are corrected based on normals are obtained from the CDM $N_{CDM_{C1}}$ and $N_{CDM_{C2}}$ to compute corrected surface normals $Corrected\_N_{Pol_{C1}}$ and $Corrected\_N_{Pol_{C2}}$. In some embodiments of the present disclosure, these normals are computed from the CDM using the Plane Principal Component method described in Kadambi et al. 2015, cited above. In more detail, in some embodiments, the normals computed from the CDM, $N_{CDM_{C1}}$ and $N_{CDM_{C2}}$ are used as guidance to correct high frequency azimuthal ambiguity as well as refractive error zenith distortion in $N_{Pol_{C1}}$, jointly taking into account multi-view consistency with camera P_{C2}. These corrected normals are also more robust than otherwise noisy SFP normals in case of fronto-parallel facets as well as specularity, as described above. In some embodiments, the flatness prior of the surface (or other simple geometric shape of the surface) is also used to further refine the zenith distortion. In particular, estimated normals $N_{Pol_{C1}}$ and $N_{Pol_{C2}}$ should generally be flat, and therefore the normals from the recovered surface (apart from areas with local surface irregularities) should be approximately 90 degrees and parallel to each other in each of the cameras. In some embodiments of the present disclosure, a voting scheme is used to perform the normals correction.

[00192] A pseudocode description of an algorithm for normals correction based on voting with a flat surface prior is presented in more detail as follows. As notation:

[00193] **N_P_C1** - shape-from-polarization (SFP) normals in Camera1

[00194] **N_P_C2** - SFP normals in Camera2

[00195] **CDM** - Coarse Depth Map

[00196] **N_CDM_C1** - Normals in CDM in Camera1 image space

[00197] **N_CDM_C2** - Normals in CDM in Camera2 image space

[00198] **Trans_C2_C1** - Relative pose of Camera2 with respect to Camera1 obtained through extrinsic and intrinsic camera calibration

[00199] **Trans_CDM_C1** - Transform used to align CDM with the object in Camera1 image space

[00200] **Trans_CDM_C2** - Transform used to align CDM with the object in Camera2 image space

[00201] **(~)** - Consistency operator

[00202] **obeys_flatness()** - operator that checks if the normals being selected obey a flatness constraint

[00203] The consistency operator (~) may be modeled as a distance metric (e.g., a cosine similarity based angular distance metric) computed between the normals being compared for consistency. If the angular distance is less than a threshold, the normals being compared are consistent with each other, else not (!~). The normals being compared are transformed into the same coordinate frame (master-camera or Camera1 image space in this case) using the transforms listed above before applying the consistency operator (~).

[00204] Pseudocode implementation of voting based on flatness prior:

```
For each pixel normal in N_P_C1:
        # Case 1: SFP normals in both cameras are consistent with CDM & Obey Flatness Constraints (No ambiguity)
        if (N_P_C1 ~ N_P_C2) && (N_P_C1 ~ N_CDM_C1) && (N_P_C2 ~ N_CDM_C2) && obeys_flatness(N_P_C1)==True && obeys_flatness(N_P_C2)==True then:
                retain N_P_C1 or Trans_C2_C1 (N_P_C2) depending on which of the 2 camera normals are more aligned (consistent) with the CDM+Flatness Constraint

        # Case 2: SFP normals in the 2 cameras are consistent with each other, but both are inconsistent with CDM normal (HFV Blemish/defect on surface)
        else if (N_P_C1 ~ N_P_C2) && (N_P_C1 !~ N_CDM_C1) && (N_P_C2 !~ N_CDM_C2) then:
                retain N_P_C1

        # Case 3: SFP normals in only one of the cameras are consistent with CDM+Flatness Constraints (specularity / fronto parallel issue in the other camera)
        else if (N_P_C1 ~ N_CDM_C1) && (N_P_C2 !~ N_CDM_C2) && obeys_flatness(N_P_C1)==True then:
                retain N_P_C1 #specularity / fronto parallel facet issue in Camera2
        else if (N_P_C2 ~ N_CDM_C2) && (N_P_C1 !~ N_CDM_C1) && obeys_flatness(N_P_C2)==True then:
                retain Trans_C2_C1 (N_P_C2) #specularity / fronto parallel facet issue in Camera1

        # Case 4: SFP normals in both cameras are inconsistent with each other, as well as with CDM
```

1

else if (N_P_C1 !~ N_CDM_C1) && (N_P_C2 !~ N_CDM_C2) && (N_P_C1 !~ N_P_C2) then:

retain N_CDM_C1

[00205] In some embodiments, the corrected surface normals $Corrected\_N_{Pol_{C1}}$ and $Corrected\_N_{Pol_{C2}}$ are used to reconstruct the shape of the surface object. For example, in some embodiments, a sparse matrix inverse algorithm can be applied (as described in Kadambi et al. 2015) to estimate the revised depth coordinates of the surface. These revised depth coordinates have a higher resolution than the initial depth obtained from standard 3D imaging techniques (stereo, time of flight, etc.).

[00206] Surface defects and irregularities may then be detected based on detecting normals that are noisy or erroneous or that otherwise dis-obey pose consistency across the different camera modules of the stereo polarization camera system. In some circumstances, these sparse irregularities are especially apparent in standing out in different proportions across the DOLP images calculated for each of the views. In other words, portions of the normals map that violate the assumption of flatness or otherwise smoothness of the surface may actually be non-smooth surfaces, thereby enabling the detection of sparse irregularities in a surface that is assumed to be generally smooth.

[00207] *Augmenting Active Scanning Systems including Lidar, Radar, Time-of-Flight, and Structured Illumination with Polarization Imaging*

[00208] Some aspects of embodiments of the present disclosure relate to augmenting active scanning system based on polarization cues such as degree of linear polarization (DOLP) and angle of linear polarization (AOLP) as computed from polarization raw frames captured by polarization camera systems. In some circumstances, the augmentation of an imaging modality or the combination of multiple imaging modalities may be referred to as sensor fusion.

[00209] As one motivating example of augmenting active scanning systems, some autonomous driver assistance systems (ADAS) include scanning systems in order to capture information about the driving environment to offer provide the driver with features such as Forward Collision Warning (FCW), Automatic Emergency Braking (AEB), Adaptive Cruise Control (ACC), and Lane Departure Warning (LDW). Generally, these ADAS features respond to objects and the environment at a macro level, wherein the sensor system of the vehicle is able to reliably sense larger objects or features on the road such as other vehicles and lane separation signs. However, the ability to detect smaller vehicles such as motorbikes and cyclists may not be as robust. Furthermore, small objects and obstacles such as rocks, nails, potholes, and traffic cones may escape detection by such the sensor systems of

1

such ADAS of vehicles. Similar issues may arise for more autonomous or self-driving vehicles, which would benefit from the ability to detect, and navigate around, a wider variety of potential road hazards and to avoid harming other people, animals, or objects. Absent such ability, the autonomous driving systems of today cannot claim to provide a safer alternative given that an attentive driver, in many cases, can be more proactive in sensing and taking corrective action as required.

[00210] Accordingly, some aspects of the present disclosure relate to augmenting sensing systems using polarization to improve the detection of small and/or optically challenging objects in a scene.

[00211] FIG. 10A is a schematic diagram of a system in which an active scanning system is augmented with a polarization camera system according to one embodiment of the present disclosure. In the embodiment shown in FIG. 10A, a polarization camera system 1010 configured to capturer polarization raw frames 1018 (e.g., an individual polarization camera configured to capture different images corresponding to different polarization states, a polarization camera array, or a stereo polarization camera system, as described above) augments an active sensing system 1020. The active sensing system includes an active emitter 1022 configured to emit waves (e.g., electromagnetic waves or sound waves) toward a scene 1001 containing various objects 1002 and 1003, where reflections 1004 of the waves off of surfaces of the objects 1002 and 1003 are detected by a detector 1024 of the active sensing system 1020.

[00212] For example, in the case of a radar system, the active emitter 1022 may emit radio frequency or microwave frequency electromagnetic waves, and the detector 1024 may include an antenna array configured to detect the reflected signal. In the case of a lidar and/or time-of-flight depth camera system, the active emitter 1022 may include a laser configured to emit light into the scene 1001 (e.g. by scanning the laser over the scene 1001 or by flashing light over the scene) and computing depth based on the time that elapses until reflected light is received. In the case of an active stereo depth camera system, the active emitter 1022 may emit structured light or patterned light onto the scene, and the detector 1024 is configured to detect light in the wavelength emitted by the active emitter 1022 (e.g., infrared light). The processing system 1030 may be configured to control both the polarization camera system 1010 and the active sensing system 1020, including controlling the polarization camera system 1010 and/or the active scanning system 1020 to capture data and receiving raw data captured by the polarization camera system 1010 and the active sensing system 1020 to detect objects or otherwise analyze a scene.

[00213] As discussed above, in many instances, optically challenging objects may be substantially invisible to comparative or standard camera systems that do not capture information regarding the polarization of light. On the other hand, capturing polarization raw frames of scenes and computing polarization cues from the polarization raw frames can increase the contrast of optically challenging objects in a scene, because these optically challenging objects may present distinctive polarization signatures (e.g., in the case of a transparent glass ball, shiny metal on the ground, and the like).

[00214] Accordingly, some aspects of embodiments of the present disclosure relate to the use of a predictor 750 configured to receive first tensors 50 in polarization feature spaces as input (e.g., a trained polarization CNN) to identify optically challenging objects in a scene and to generate a characterization output 20 such as a segmentation map that can be combined or fused with the outputs of other sensors operating on the same scene (e.g., radar and/or lidar data showing the distances and velocities of various objects in a scene, a depth map computed based on an RGB/gray-based camera system, or a segmentation map computed by an RGB/gray-based camera system).

[00215] Continuing the above example of ADAS and automated driving for vehicles, a polarization CNN is trained to detect road obstacles at a micro level, in cluttered environments, of optically challenging objects such as shiny metal and glass, as well as other objects that may prove to be hazardous, such as semantically inconsistent objects that are transient and even more hazardous, such as rocks, water, ice, or oil slicks on the road. These represent road conditions that are hazardous and unexpected, and that call for caution. Accordingly, fusing the large scale, coarse data from active scanning systems such as lidar and radar with segmentation maps computed using polarization cues enables ADAS and automated driving systems to detect a wider range of objects in a driving environment, from large scale vehicles down to micro-scale nails, flat metal panels, and the like.

[00216] As noted above, lidar and radar systems are generally limited to generating sparse point clouds of information regarding the scene. In particular, the active emitter 1022 of the active scanning device 1020 may have various physical constraints that force a tradeoff between breadth of coverage (e.g., solid angle swept over by the active emitter 1022), resolution, and latency. For example, physical constraints may limit the speed at which a laser emitter of a lidar can sweep over a scene, and sweeping more quickly causes fewer rays of light to be emitted toward any one part of the scene, thereby reducing the signal received by the detector 1024 for any given angle. Sweeping more slowly can increase the time between

1

successive scans of an area (or reduce a frame rate of the data), thereby increasing latency, which may cause problems in applications that require real-time control. Likewise, in the case of active stereo depth sensors, a structured light projector emits patterned light toward a scene, where the pattern may include dots and/or lines where local patches are globally unique across the pattern. However, the angular resolution of the system is constrained by the resolution of the dot pattern.

[00217] Therefore, some aspects of embodiments of the present disclosure relate to generating to generate higher quality 3-D models by fusing sparse point clouds captured by 3-D active scanning systems with surface normals computed from polarization data captured by a polarization camera system.

[00218] FIG. 10B is a flowchart depicting a method 1050 for fusing the 3-D model captured by an active 3-D scanning system with surface normals captured by a polarization camera according to one embodiment of the present disclosure. Referring to FIG. 10B, in operation 1051, a processing system 1030 controls an active scanner 1020 (e.g., lidar, radar, active stereo, time-of-flight, etc.) is used to capture a point cloud of a scene 1001. In operation 1052, the processing system 1030 controls the polarization camera system 1010 to capture polarization raw frames 1018. In some embodiments, operations 1051 and 1052 are performed concurrently, such that substantially the same scene is sensed by both systems (e.g., in dynamic situations such as driving through traffic, the scene may continuously change over time due to the movement of the vehicle on which the ADAS or automated driving system is mounted, as well as the movement of other vehicles and objects in the environment). In operation 1053, surface normals are computed from the polarization raw frames, such as by applying the techniques described above (e.g., extracting tensors in AOLP ϕ and DOLP ρ representation spaces and computing surface normals in accordance with equations (2), (3), (4), and (5)). In operation 1054, the surface normals are integrated over sample space to compute a 3-D surface, and in operation 1055, the 3-D surface computed based on the surface normals is corrected in accordance with constraints of the sparse point cloud captured by the active scanner 1020 to generate a 3-D model of the scene that is of higher quality than the point cloud captured by the active scanner 1020 operating alone.

[00219] Some aspects of embodiments of the present disclosure further relate to generating higher quality 3-D models of a scene through beam steering of the active scanner 1020 to perform higher-resolution scans of particular regions of interest within the scene. In more detail, in some embodiments of the present disclosure, a predictor 750 is configured to identify regions of interest within a given scene. For example, in some embodiments, a polarization convolutional neural network is trained to compute a segmentation mask that identifies regions of interest, based on labeled training data. Continuing the example of sensors for ADAS and automated driving systems, a polarization CNN may be trained to identify objects of interest in a scene that may be hard to detect or analyze using a comparative active scanning system operating alone with its active emitter 1022 configured to emit a wide beam (e.g., small vehicles such as motorcycles and bicycles, small and medium sized objects such as traffic cones, chains, and other lightweight barriers, and the like).

[00220] FIG. 10C is a flowchart illustrating a method for augmenting an active sensing system using polarization according to one embodiment of the present disclosure. As shown in FIG. 10C, according to one method 1060, the processing system 1030 controls the polarization camera system 1010 to capture polarization raw frames in operation 1061. In operation 1062, the processing system 1030 identifies regions of interest in the polarization raw frames, such as by extracting polarization cues (or tensors in polarization representation spaces) from the polarization raw frames and supplying the polarization cues to a polarized convolutional neural network trained to identify objects and regions of interest for a given use case. In operation 1063, the processing system 1030 steers the active scanning system (e.g., steers the direction and/or solid angle of a beam emitted by the active emitter 1022) to scan the identified regions of interest to generate high-resolution active scans of the regions of interest. These active scans may provide more accurate information than the output of the low-resolution, wide scans of the scene performed by the active sensing system 1020 without using beam steering, and may also provide different information than determined by the polarization camera system 1010. For example, the polarization camera system 1010 may provide high precision angular data regarding the direction of a particular surface or region of interest within a scene, but may not be able to provide accurate information about the distance or three-dimensional shape of the surface of interest. On the other hand, focusing or steering the beam emitted by an active emitter 1022 toward the surface or region of interest enables high-precision measurements to be made of those surfaces (e.g., base on time-of-flight to points within the identified regions). As specific examples of beam steering, in the case of lidar, a standard scanning cycle of the lidar system can be focused on scanning only the identified regions of interest, thereby increasing the energy emitted toward those regions of interest, and thereby increasing the signal that can be detected by the detector 1024. Similarly, in the case of radar, beamforming techniques (e.g., using a phased array) may be used to form or shape the beam to be emitted over a wider or narrower shape, as well as to steer the direction in which the signals are emitted. In the case of an active stereo scanning system, the structured light pattern may be focused to cover only the particular regions of interest, thereby increasing the resolution of the pattern in those regions of interest, and thereby enabling higher resolution block matching to be performed over the surfaces in the region of interest. In some embodiments, selective beam steering across multiple frames through tracking in the registered/fused polar data. Accordingly, a steered beam can provide, for example, higher resolution 3-D point clouds of particular regions of interest. In some embodiments, these higher resolution 3-D point clouds are further enhanced using surface normals computed from the polarization raw frames captured by the polarization camera system 1010.

[00221] Capturing higher quality 3-D models of objects using active scanning systems (e.g., through beam steering and/or through the improvement of surface shape detection using surface normals computed from polarization) provides improvements to object detection and classification due to higher resolution models supplied as input to the classifier. For example, a neural network trained to classify objects (e.g., distinguishing between a dog and a fire hydrant), will produce more accurate results when the input to the neural network is of higher quality. Accordingly, aspects of embodiments of the preset disclosure implementing sensor fusion between active sensing systems 1020 and polarization camera systems 1010 can improve the accuracy and responsiveness of object detection systems.

[00222] *File Formats with Surface Normals from Polarization*

[00223] Shape from Polarization approaches, as described above, recover surface normals from polarization cues captured by a polarization imaging system. Surface normals provide a valuable prior about the surface being imaged that opens a wide range of possibilities for visual inspection of the scanned surfaces that are needed across different manufacturing verticals such as: vehicles (e.g., automobiles, aircraft, and watercraft) and vehicle parts (ex: tires, engine blocks, transmissions, painted surfaces, etc..), 3D metal or polymer-based printing, printed circuit boards (PCBs), and mission-critical medical devices (e.g., ventilators, pumps, stents, and the like).

[00224] Storing surface normals together with the images enables interactive post-visualization of the scanned surface for a variety of use cases. Examples include: inspection for surface anomalies on the scanned surface; and the ability to re-light the scanned surface for arbitrary lighting conditions.

[00225] Surface profile examination are important in analyzing the intrinsic shape and curvature properties of surfaces. These frequently reveal anomalous behavior which is not immediately apparent to the naked eye in simple iso-parametric surface digitizations of objects. The ability of polarization enhanced imaging to provide order of magnitude improvements to surface normals accuracy can enable their application in high precision manufacturing of industrial parts for smoothness detection and shape fidelity.

[00226] The need for fair or smooth surface shapes can be motivated by different considerations in different market verticals. In automotive design aesthetics dominate, while in aircraft and ship-building industry, aerodynamic flow is critical for fuel efficiency. The smoothness of surfaces is related to many different features such as: continuity between adjacent patches in tangents and curvature; curvature distribution; flat points; and convexity, which are driven directly by surface differential geometry. Aesthetically smooth surfaces cannot have bumps or dents, which are essentially variations in local curvature which in turn are defined by their surface normal representations. Other applications such as robotic bin picking rely on accurate surface profile of the parts being picked and the recognition of surface imperfections like exceeding curvature bounds and high variation of curvature can prevent failure of tool-path generation algorithms for controlling the robotic picker.

[00227] Visualization of differential geometric features such as curvature behavior, parabolic lines, and iso- or geodesic lines and their variational behavior in the presence of environmental stimuli such as temperature or humidity are critical in the analysis of surface properties of manufactured parts in mission critical use conditions. In all of the above cases, the ability to scan the surface in real-time made available by polarization enhanced imaging can provide major improvements to real-time monitoring and control. In some embodiments, real-time surface normals tracking provides significant improvements in surgery by providing surgeons with real-time information about induced or built-in stress on the operated upon surface tissues (such as brain tissues). This information is often critical for surgeons as they can now respond, in real-time, to their surgical path planning in response to variations in surface stress of the affected regions.

[00228] Accordingly, some aspects of embodiments of the present disclosure relate to file formats for storing information regarding surface normals captured by polarization camera systems. In addition to storing the surface normals of the scanned surface along with the image information (e.g., red, green, blue, and, in some cases, alpha color channels) in commonly used file formats (such as JPEG, TIFF, PNG) additional information from polarization include the degree and angle of polarization (DOLP ρ and AOLP ϕ). These additional channels of information provide visual cues in surface inspection, often providing enhanced contrast (even when the original image is of poor quality due to lack of adequate exposure or glare). Storing

1

these images in compressed form enables significant latitude in producing visual overlays to allow for increased visual understanding of surface anomalies. These can also help to refine depth maps to a much higher degree of precision than is possible with conventional non-polarization technologies currently available.

[00229] One example embodiment for storing polarization data (e.g., DOLP and AOLP) alongside color information relates to using the “application markers” of the JPEG File Interchange Format (JFIF). Metadata can be stored in JFIF using sixteen “application markers,” which makes it possible for a decoder to parse the interchange format and decode only required segments of image data. Although the current JFIF standard limits application markers to 64K bytes each, it is possible to use the same marker ID multiple times and refer to different memory segments. This enables one to store surface normal information in compressed or uncompressed formats. In addition, in some embodiments of the present disclosure, one of the application markers is used to store specific additional information such as degree and angle of polarization images. While one embodiments of the present disclosure is described herein with regard to storing polarization data along color image data using the JFIF standard, embodiments of the present disclosure are not limited thereto.

[00230] In some embodiments, surface normals are represented as vectors in three dimensions and can therefore take up a lot more space than the size of the original image (e.g., larger than the size of a corresponding RGB color or grayscale image). To save space, a compressed form of surface normal can be stored by sharing the exponent across the three dimensions and using 8 bits for each of the exponents in fixed-point format. Further reduction may be achieved, at the expense of re-computing the normals along the 3 dimensions at render time, by storing just the azimuth and zenith angles, or by just storing the DOLP and AOLP.

[00231] In some embodiments of the present disclosure, an existing image format that is capable of storing three channel information (e.g., red, green, and blue color information) is repurposed to store three-dimensional surface normal information. Although such embodiments may no longer be able to store color information (e.g., color information may be stored in a separate file), these approaches enable the leveraging of existing features such as image compression and progressive display, and also allowing for the use of existing software tools for writing data to such formats and parsing data from such formats. In some embodiments of the present disclosure, the mapping of the channels of the image format and the three-dimensional directions (e.g., x-, y-, and z-directions) is fixed or otherwise previously agreed-upon, in order to avoid the misinterpretation of the orientations of the surface normals.

[00232] *Augmenting DSLR and Video Cameras with Polarized Imaging*

[00233] Some aspects of embodiments of the present disclosure relate to augmenting digital camera systems such as digital single-lens reflex (DSLR) cameras and video cameras with polarized imaging. In some embodiments of the present disclosure, this relates to mounting a polarization camera system onto an underlying standard color camera system (e.g., the polarization camera system may include a single polarization camera capable of capturing data at multiple different polarization angles, a polarization camera array in which each camera array is configured to capture light of a different polarization state, and/or a stereo polarization camera system as described above). According to some embodiments of the present disclosure, the polarization camera system and the underlying camera system may be registered by imaging calibration targets (e.g., a checkerboard pattern) to compute the extrinsic camera parameters for mapping between the multiple views captured by the polarization camera system and the underlying camera system. This enables an augmented DSLR camera and/or an augmented video camera system to perform surface normal extraction and, in the case of stereo camera systems, perform depth estimation.

[00234] Various applications of combining color images with polarization images are described above. In embodiments of the present disclosure in which a color camera system is augmented with a polarization camera system, the polarization camera system is configured to synchronize its capture of polarization raw frames with the capture of images by the underlying system. For example, a same shutter release button on the DSLR may be used to trigger both the capture of images by the DSLR and by the polarization camera system. Likewise, a same "record" button on a video camera system may be used to also control the polarization camera system to capture polarization raw frames that are time indexed to match the frames captured by the underlying video camera system. In some embodiments of the present disclosure, a polarization camera system is configured to fit into the flash hot shoe or otherwise configured to be controlled by the flash system of the underlying camera system to provide the synchronization between the triggering of the underlying camera and the polarization camera systems.

[00235] As discussed above, various file formats may be used to store the polarization data and/or surface normals as captured by the polarization camera system alongside the color or grayscale image data captured by the underlying camera, and/or by re-purposing standard color image file formats to store surface normal data.

[00236] Capturing polarization raw frames (and, accordingly, surface normals of objects in a scene) concurrently with color video data may be used to perform further analysis of a scene, such as by providing more accurate depth estimations, which may be used for providing additional focus control (e.g., predicting which portions of a scene will be in focus based on depth from the camera). Capturing stereo polarization raw frames may also enable the concurrent capture of depth maps along with color texture information of a scene.

[00237] In some embodiments of the present disclosure, the surface normals computed from the polarization raw frames captured by the camera system are used to assist in motion capture of the movement of people or objects in a scene. In more detail, the surface normals may provide information about the location and orientation of motion capture markers on objects in a scene, whereas comparative techniques may track only the location of the markers. Similarly, in the case of performing motion capture without markers, surface normals may also provide information about the orientation of various surfaces, thereby also improving the fidelity of the capture. Polarization based imaging may also avoid the temporary invisibility of markers or surfaces of motion captured objects in a scene due to specular reflections or glare, as discussed above. Accordingly, applying polarization camera systems to motion capture techniques can improve the accuracy and richness of motion captured data.

[00238] *Computational Photography with Polarization*

[00239] When polarized light enters a polarization filter, the intensity of outgoing light changes depending on the relative angle between the polarized incoming light and the polarizing filter. By capturing the scene at multiple angles of polarization, it is possible to eliminate glare, specular reflections, multiple reflections in at least one of the multiple angles of polarization images captured. These approaches may be used in industrial imaging applications to visualize hard-to-image scenes that are constrained by glare or specular reflections. However, polarization can also be used in computational photography to improve the capture of photographs for human visual consumption. Some aspects of embodiments of the present disclosure relate to the synthesis of high dynamic range images using multiple polarized input images, which may also improve the sharpness of the captured images.

[00240] High dynamic range (HDR) imaging generally involves capturing multiple images of a scene at different exposures to capture the full extent of the intrinsic dynamic range of the scene. However, the different exposures provide edges of different degrees of sharpness across the images, making them hard to align or at best aligning them imperfectly, thereby resulting in softer-looking (e.g., blurrier) images. Some comparative approaches capture frames of constant exposure, which makes alignment more robust, where the exposures are set low enough to avoid blowing out the highlights. The resulting merged HDR image has clean shadows and high bit depth. However, this works as long as there is no motion present in the scene (e.g., motion in the camera and/or motion of objects in the scene). Any motion while capturing frames of constant exposure can create blur and introduce unwanted artifacts (e.g., ghosting) in the edges.

[00241] Accordingly, some aspects of embodiments of the present disclosure relate to the use of capturing polarization raw frames at the same exposure settings and synthesizing high dynamic range images in the presence of scene motion using the polarization raw frames.

[00242] Having multi-camera systems with different polarization filters along with a reference camera with nominal Bayer filters allow for capturing multiple images concurrently (e.g., simultaneously) with the same exposure settings across all cameras. This allows for fusion of these multiple images without suffering the impact of anomalies created by motion across temporal frames (e.g., frames captured at different times). In addition, the different polarization filters allow for capturing the same scene at different intensities as if they were captured with different “neutral density” filters. Because the polarization filters have a natural effect of reducing the intensity of light incident on the camera system (e.g., incident on the image sensor), the exposure at which these sets of images are captured may not need to be as low as what would be needed in the comparative approach outlined above, as in that case there is nothing that attenuates the incident light on the camera system thereby necessitating carefully calibrated lower exposures.

[00243] Furthermore, merging polarization raw frames enables the recovery of detail in glare or otherwise saturated portions of the scene due to specular reflection. For example, under the comparative approach in which multiple lower exposure images are captured of a scene, it is still possible that portions of the scene will be overexposed to the point of saturation due to specular highlights. On the other hand, by capturing polarization raw frames of the scene using different polarization states and, in the case of a camera array and/or a stereo polarization camera system, from different viewpoints, it is unlikely that a given surface patch of the scene will exhibit specular glare from all perspectives. This, in turn, enables the recovery of detail from regions that would otherwise be lost in the case of a standard camera system that did not use polarization raw frames.

[00244] FIG. 11 is a flowchart illustrating a method for synthesizing a high dynamic range (HDR) image from polarization raw frames. In operation 1101, a processing system (e.g., a controller of a camera system) controls a polarization camera system to capture polarization raw frames and, in operation 1102, synthesizes a high dynamic range (HDR) image based on the polarization raw frames. As noted above, in some embodiments of the present disclosure, the polarization raw frames are combined based on techniques for combining multiple low-exposure frames (e.g., captured with a small aperture and/or short exposure time). In some embodiments of the present disclosure, the impact of glare and/or specularity can be removed and detail can be recovered in regions of the image that would be overexposed by dropping or portions of the images that are saturated in some polarization raw frames and more likely to be properly exposed in other polarization raw frames.

[00245] While the present invention has been described in connection with certain exemplary embodiments, it is to be understood that the invention is not limited to the disclosed embodiments, but, on the contrary, is intended to cover various modifications and equivalent arrangements included within the spirit and scope of the appended claims, and equivalents thereof.

WHAT IS CLAIMED IS:

1. A multi-modal sensor system comprising:

an underlying sensor system;

a polarization camera system configured to capture polarization raw frames corresponding to a plurality of different polarization states; and

a processing system comprising a processor and memory, the processing system being configured to control the underlying sensor system and the polarization camera system, the memory storing instructions that, when executed by the processor, cause the processor to:

control the underlying sensor system to perform sensing on a scene and the polarization camera system to capture a plurality of polarization raw frames of the scene;

extract first tensors in polarization representation spaces based on the plurality of polarization raw frames; and

compute a characterization output based on an output of the underlying sensor system and the first tensors in polarization representation spaces.

2. The multi-modal sensor system of claim 1, wherein the polarization camera system comprises a polarization camera module comprising:

a first polarization camera comprising a first polarizing filter at a first polarization orientation, the first polarization camera having a first optical axis;

a second polarization camera comprising a second polarizing filter at a second polarization orientation, the second polarization camera having a second optical axis substantially parallel to the first optical axis; and

a third polarization camera comprising a third polarizing filter at a third polarization orientation, the third polarization camera having a third optical axis substantially parallel to the first optical axis.

3. The multi-modal sensor system of claim 2, wherein the polarization camera module further comprises a fourth polarization camera comprising a fourth polarizing filter at a fourth polarization orientation, the fourth polarization camera having a fourth optical axis substantially parallel to the first optical axis.

4. The multi-modal sensor system of claim 2, wherein the first tensors comprise a degree of linear polarization (DOLP) and an angle of linear polarization (AOLP), and

Date reçue/Date received 2023-02-17

wherein the memory further stores instructions that, when executed by the processor, cause the processor to compute the DOLP and the AOLP based on polarization raw frames captured by the first polarization camera, the second polarization camera, and the third polarization camera, the instructions comprising instructions to:

initialize an estimated DOLP and an estimated AOLP based on stakes vectors;

estimate a scene geometry based on parallax shifts in the polarization raw frames to generate a coarse model; and

iteratively:

refine the coarse model based on the estimated DOLP and the estimated AOLP to generate an estimated geometry; and

update the estimated DOLP and the estimated AOLP based on the estimated geometry,

until a change in the estimated DOLP and a change in the estimated AOLP are both less than corresponding threshold values.

5. The multi-modal sensor system of claim 1, wherein the polarization camera system comprises a stereo polarization camera system comprising:

a first polarization camera module having a first optical axis, the first polarization camera module being configured to capture a first plurality of polarization raw frames corresponding to a first plurality of different polarization states; and

a second polarization camera module having a second optical axis and spaced apart from the first polarization camera module along a baseline, the second polarization camera module being configured to capture a second plurality of polarization raw frames corresponding to a second plurality of different polarization states,

the first optical axis being substantially parallel to the second optical axis.

6. The multi-modal sensor system of claim 5, wherein the first polarization camera module comprises a first plurality of color filters configured to transmit light in three or more different first color spectra, and

wherein the second polarization camera module comprises a second plurality of color filters configured to transmit light in three or more different second color spectra,

Date reçue/Date received 2023-02-17

wherein the three or more second color spectra are different from the three or more first color spectra.

7. The multi-modal sensor system of claim 6, wherein the memory further stores instructions that, when executed by the processor, cause the processor to:

control the stereo polarization camera system to capture multi-spectral stereo polarization imaging data in the first color spectra and in the second color spectra; and

extract first tensors in polarization representation spaces from the multi-spectral stereo polarization imaging data.

8. The multi-modal sensor system of claim 1, wherein the underlying sensor system comprises an active scanning system comprising an active emitter and a detector.

9. The multi-modal sensor system of claim 8, wherein the active scanning system comprises a radar system.

10. The multi-modal sensor system of claim 8, wherein the active scanning system comprises a lidar system.

11. The multi-modal sensor system of claim 8, wherein the active scanning system comprises an active stereo depth camera system.

12. The multi-modal sensor system of claim 8, wherein the multi-modal sensor system is mounted on a vehicle.

13. The multi-modal sensor system of claim 8, wherein the memory further stores instructions that, when executed by the processor of the processing system, cause the processor to:

compute a sparse point cloud based on the output of the underlying sensor system;

compute surface normals from the polarization raw frames;

compute a 3-D surface based on the surface normals; and

correct the 3-D surface based on the sparse point cloud to compute a 3-D model of the scene.

Date reçue/Date received 2023-02-17

14. The multi-modal sensor system of claim 8, wherein the memory further stores instructions that, when executed by the processor of the processing system, cause the processor to:

compute a segmentation map based on the first tensors in the polarization representation spaces, the segmentation map identifying one or more regions of interest in the scene;

steer the active emitter to emit beams toward the one or more regions of interest; and

detect a reflection of the beams emitted by the active emitter using the detector of the active scanning system.

15. The multi-modal sensor system of claim 1, wherein the memory stores instructions that cause the processor to implement a convolutional neural network trained to compute a segmentation map based on the first tensors.

16. The multi-modal sensor system of claim 1, wherein the underlying sensor system comprises a color camera system.

17. The multi-modal sensor system of claim 16, wherein the color camera system is a digital single lens reflex camera or a video camera.

18. The multi-modal sensor system of claim 16, wherein the output of the color camera of the underlying sensor system comprises a color image, and

wherein the memory further stores instructions that, when executed by the processor, cause the processor to:

compute a plurality of surface normals of the scene based on the first tensors; and

store the computed surface normals of the scene in a same file as the color image captured by the color camera.

Date reçue/Date received 2023-02-17

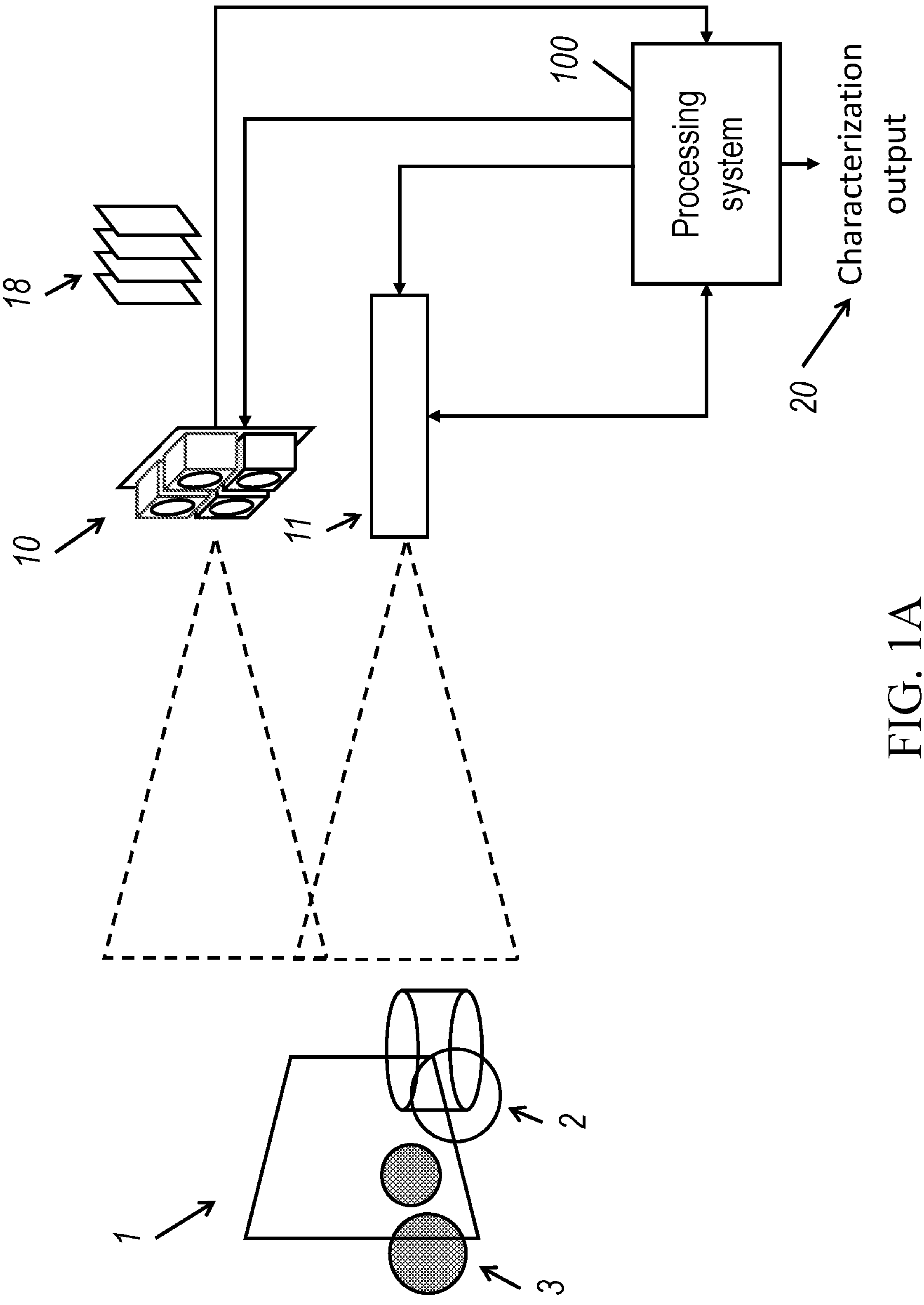

100
Processing system
Characterization output
18
10
11
20
1
2
3

FIG. 1A

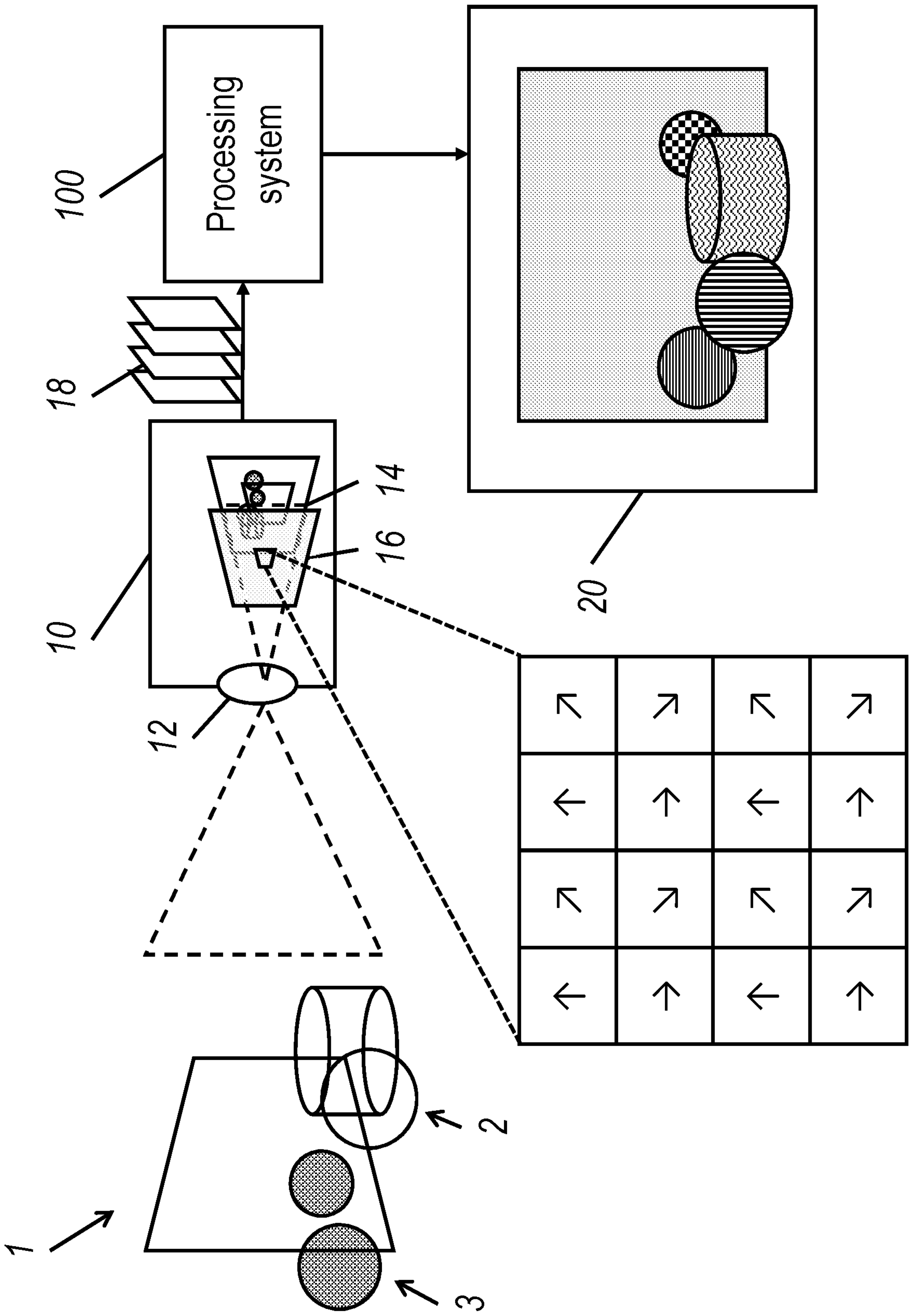


FIG. 1B

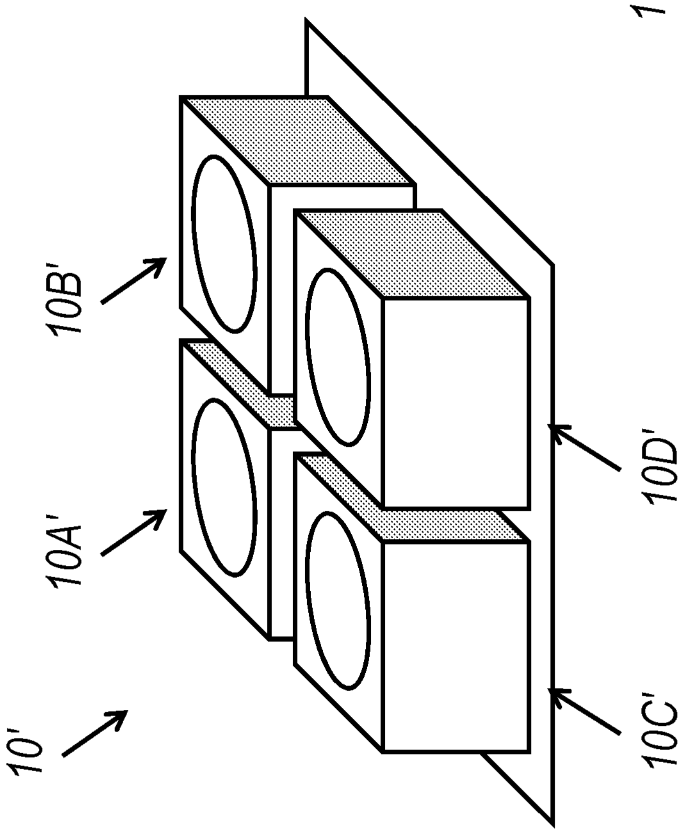


FIG. 1C

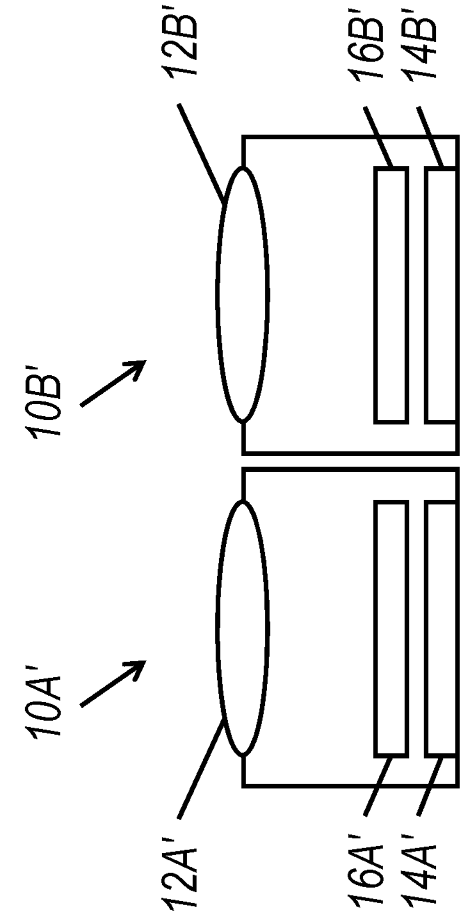


FIG. 1D

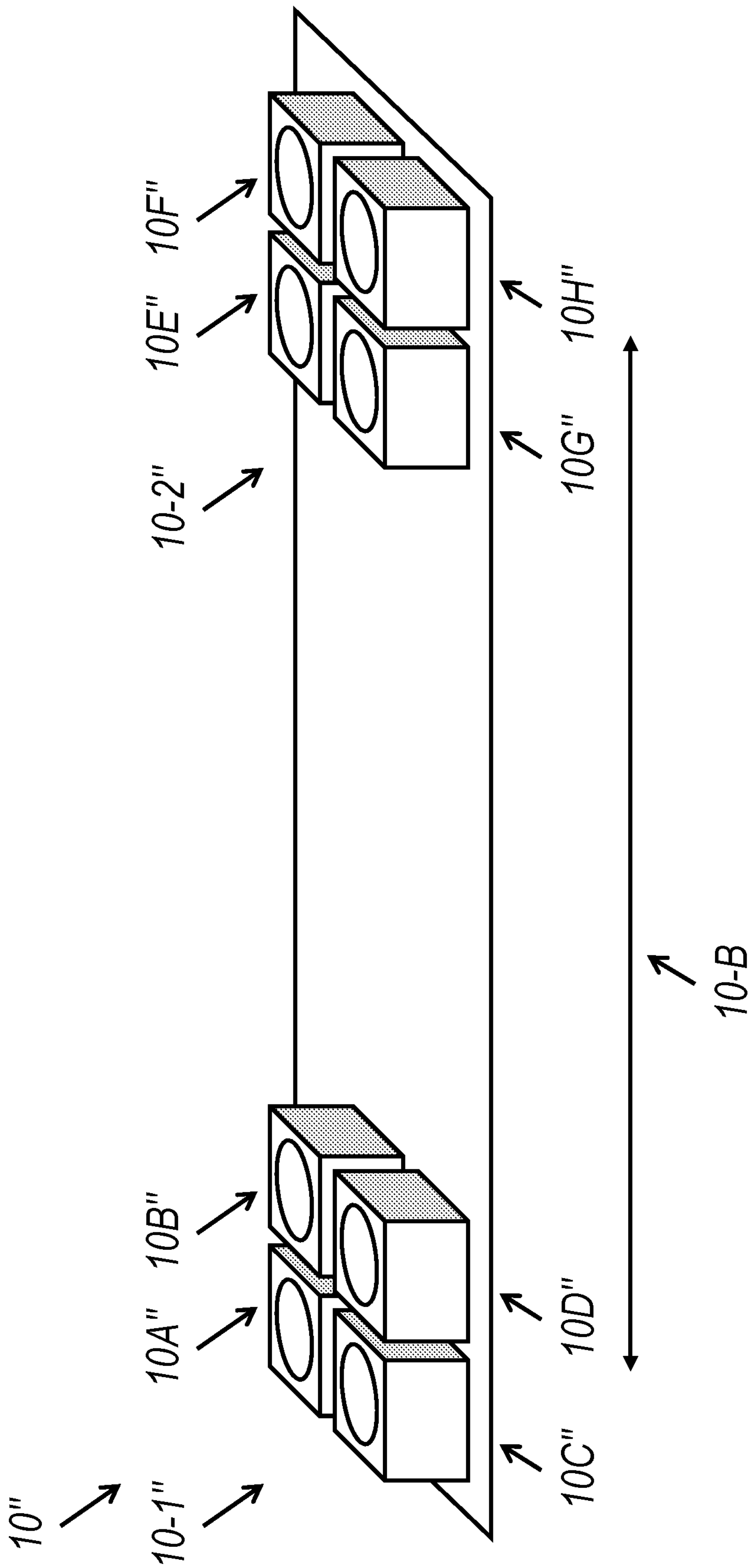


FIG. 1E

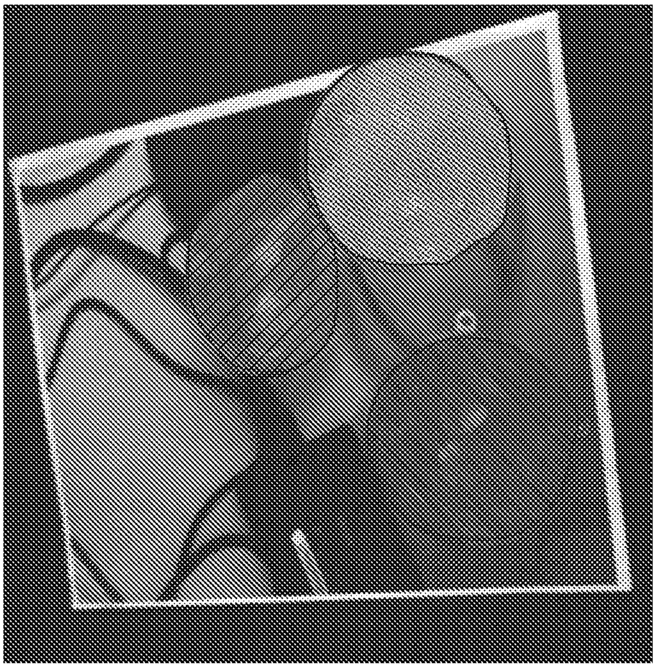


FIG. 2B

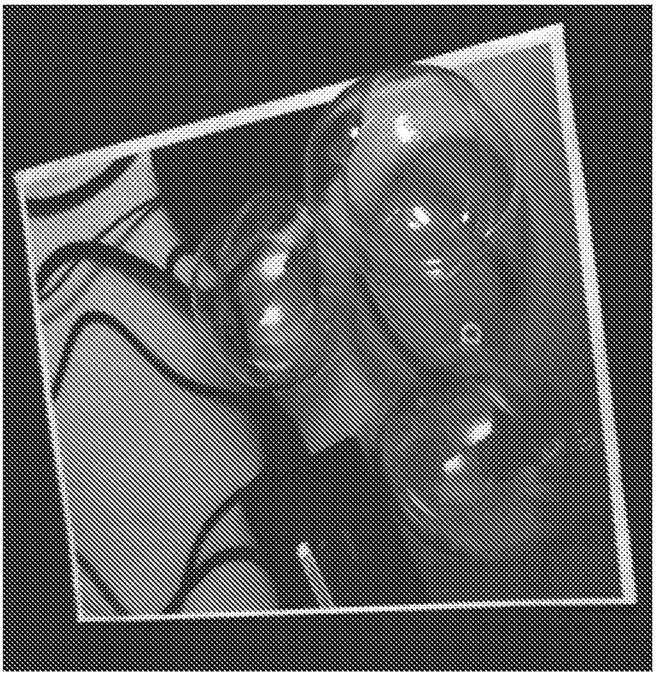


FIG. 2A

FIG. 2D

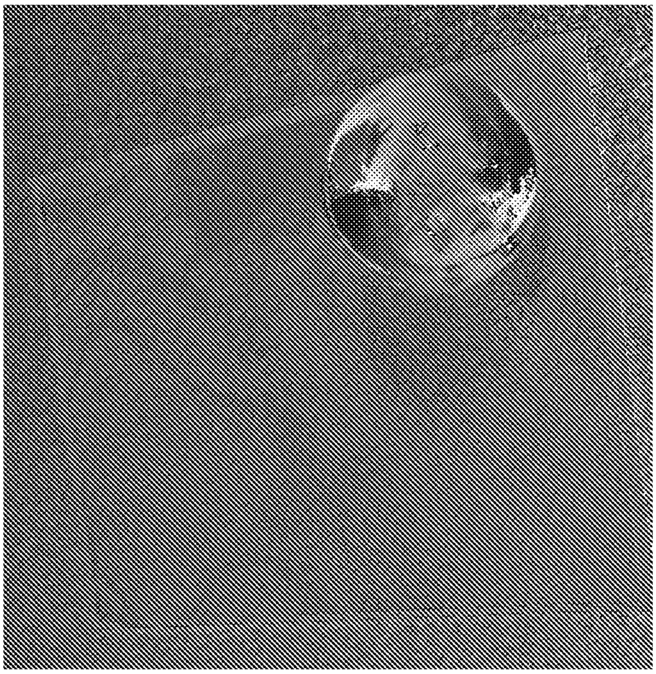


FIG. 2C

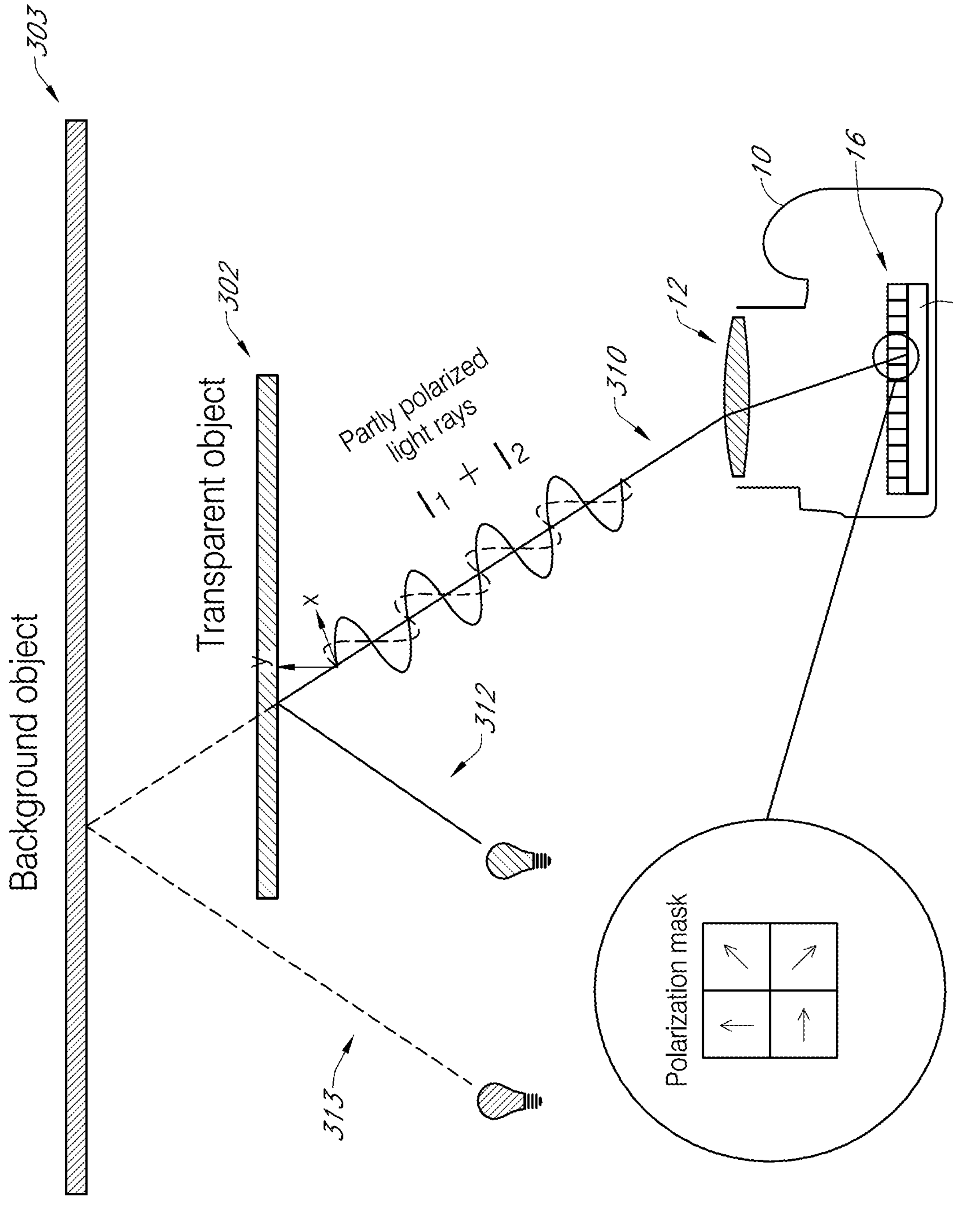

303
Background object
302
Transparent object
Partly polarized light rays
I1 + I2
x
y
310
312
313
12
10
16
14
Polarization mask

FIG. 3

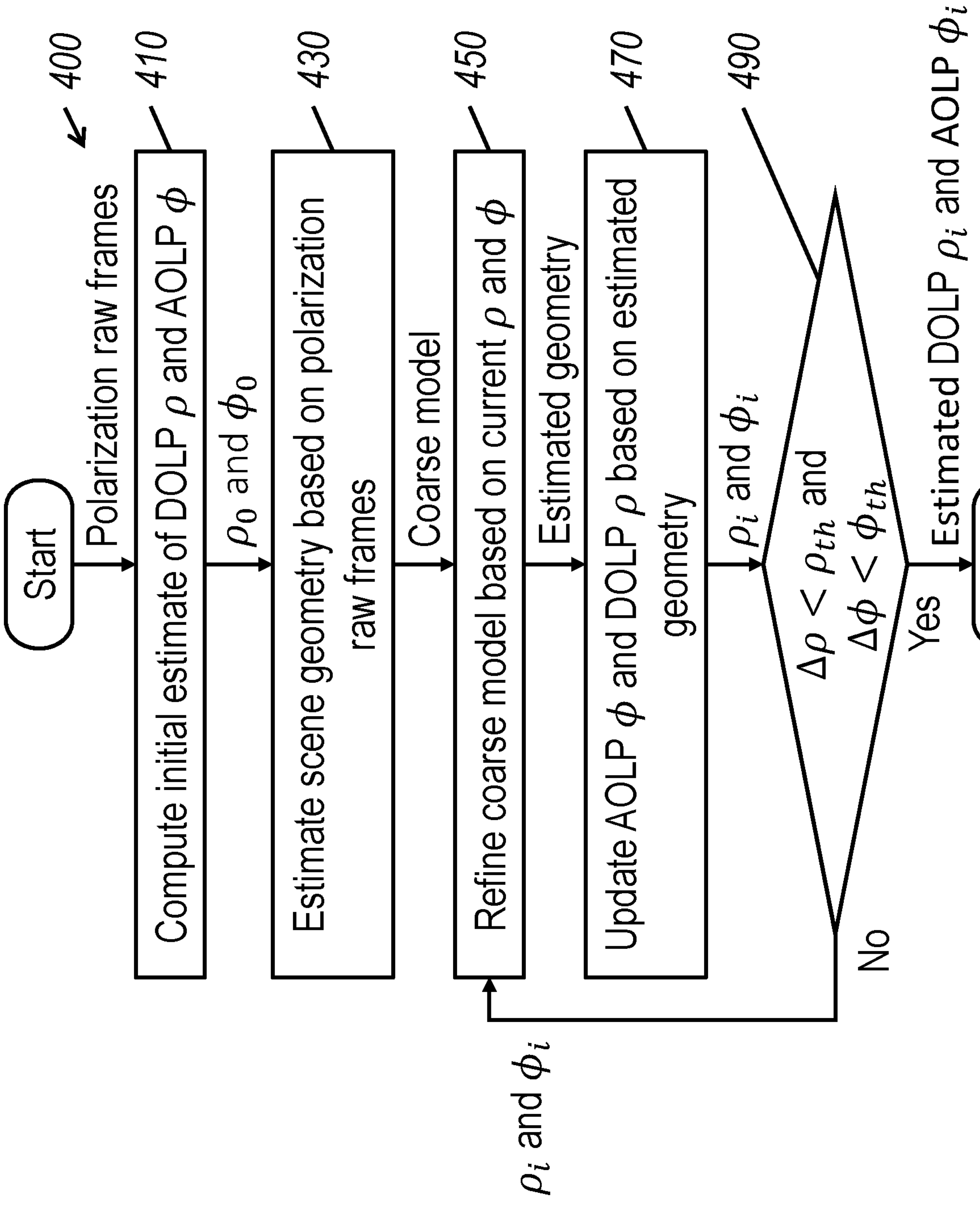

400
Start
Polarization raw frames
410
Compute initial estimate of DOLP ρ and AOLP ϕ
ρ_0 and ϕ_0
430
Estimate scene geometry based on polarization raw frames
Coarse model
450
Refine coarse model based on current ρ and ϕ
Estimated geometry
470
Update AOLP ϕ and DOLP ρ based on estimated geometry
ρ_i and ϕ_i
490
$\Delta\rho < \rho_{th}$ and $\Delta\phi < \phi_{th}$
Yes
No
ρ_i and ϕ_i
Estimated DOLP ρ_i and AOLP ϕ_i
End

FIG. 4

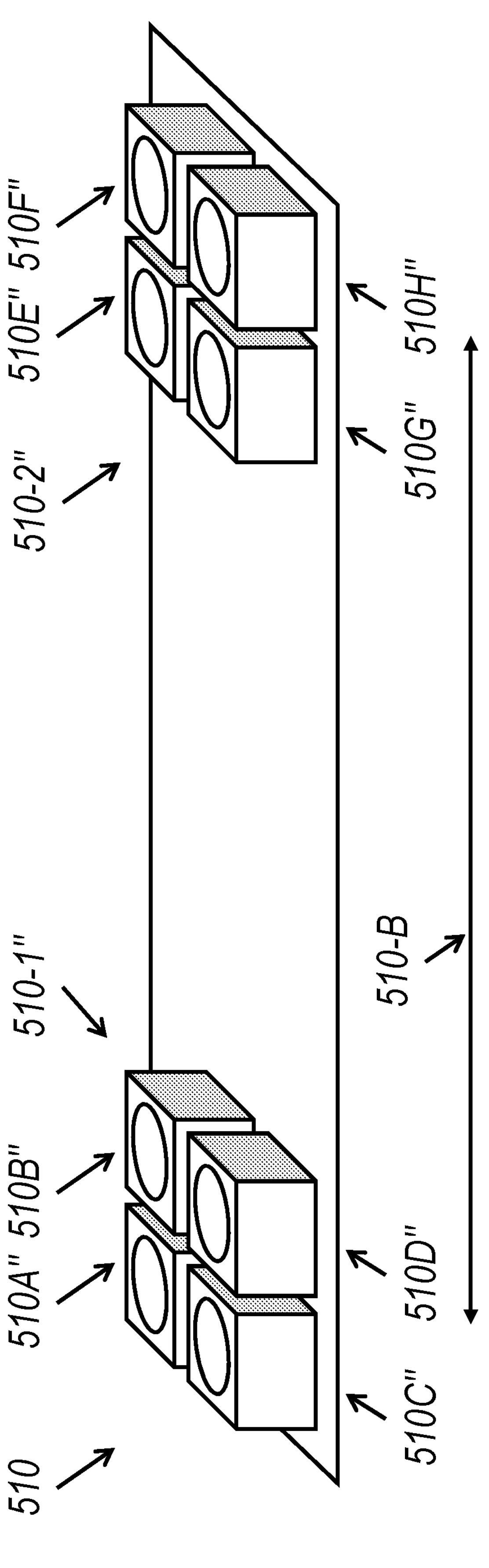

510
510A"
510B"
510C"
510D"
510-1"
510-B
510E"
510F"
510G"
510H"
510-2"

FIG. 5A

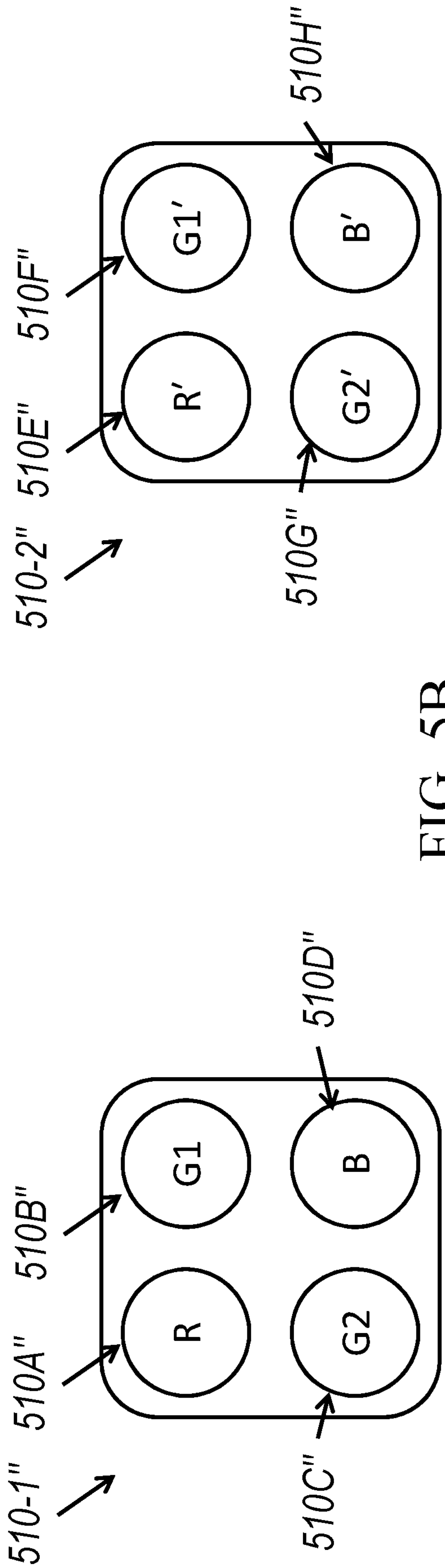

510-1"
510A"
510B"
510C"
510D"
R
G1
G2
B
510-2"
510E"
510F"
510G"
510H"
R′
G1′
G2′
B′

FIG. 5B

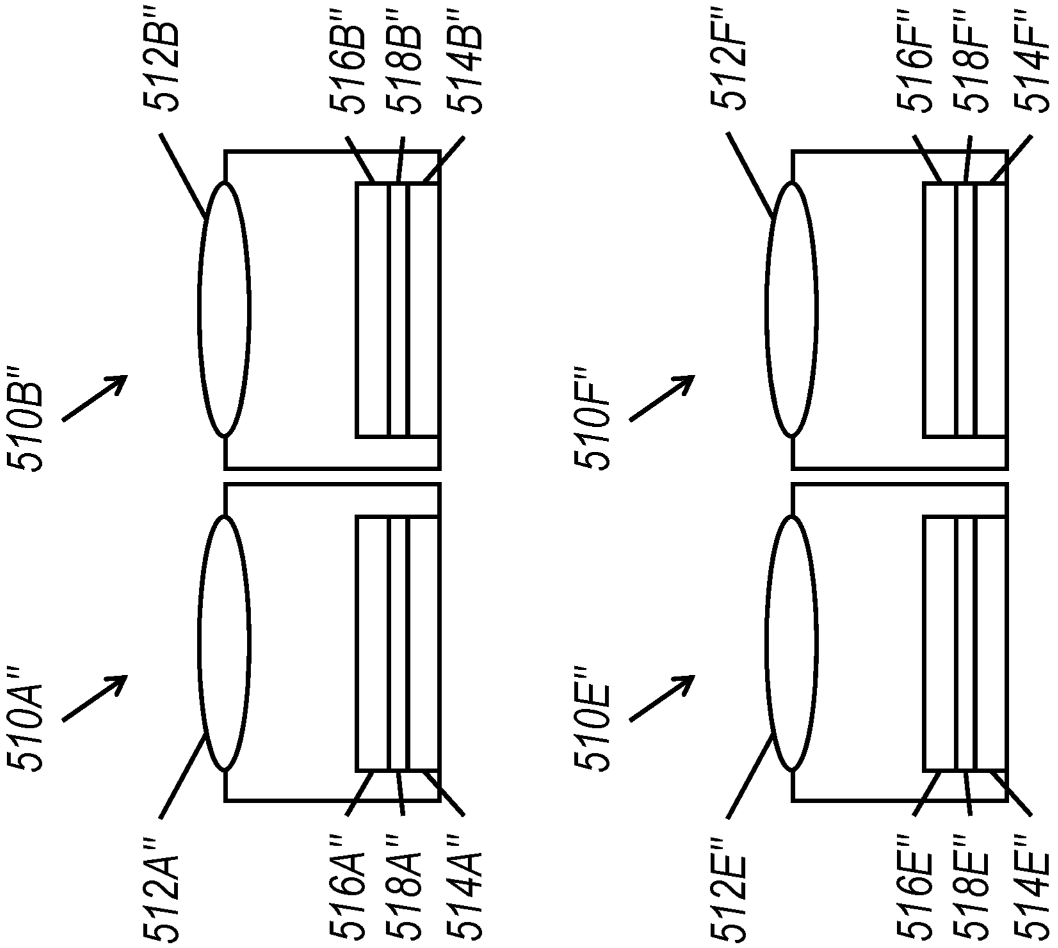

510A"
510B"
512A"
512B"
516A"
518A"
514A"
516B"
518B"
514B"
510E"
510F"
512E"
512F"
516E"
518E"
514E"
516F"
518F"
514F"

FIG. 5C

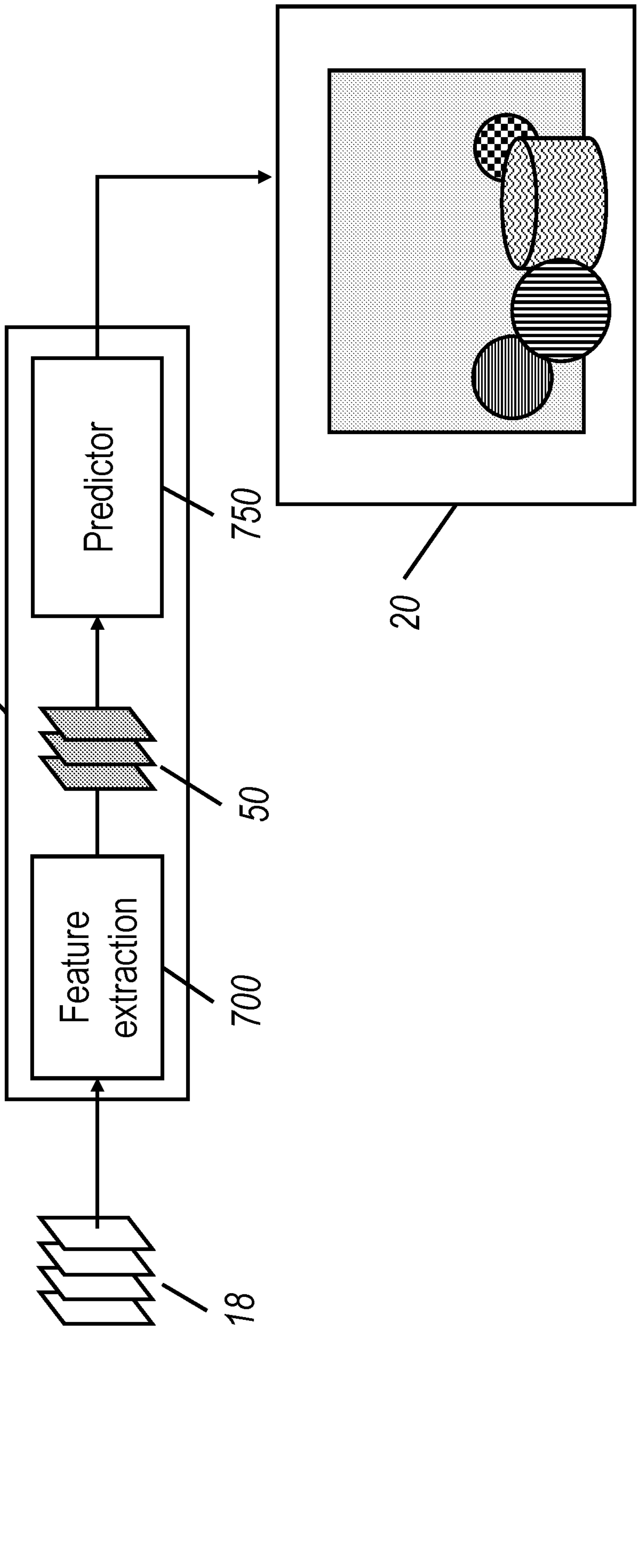


FIG. 6A

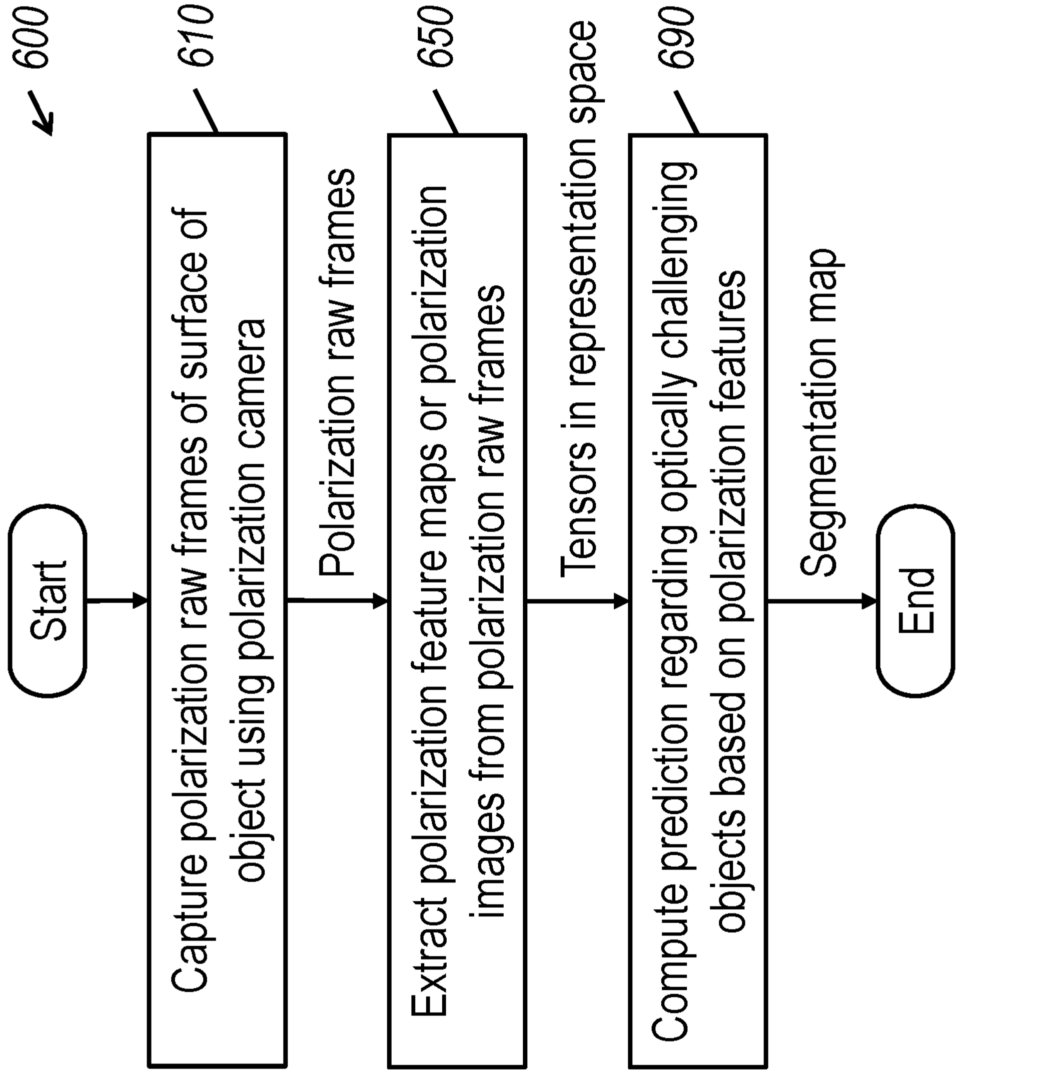

600
Start
610
Capture polarization raw frames of surface of object using polarization camera
Polarization raw frames
650
Extract polarization feature maps or polarization images from polarization raw frames
Tensors in representation space
690
Compute prediction regarding optically challenging objects based on polarization features
Segmentation map
End

FIG. 6B

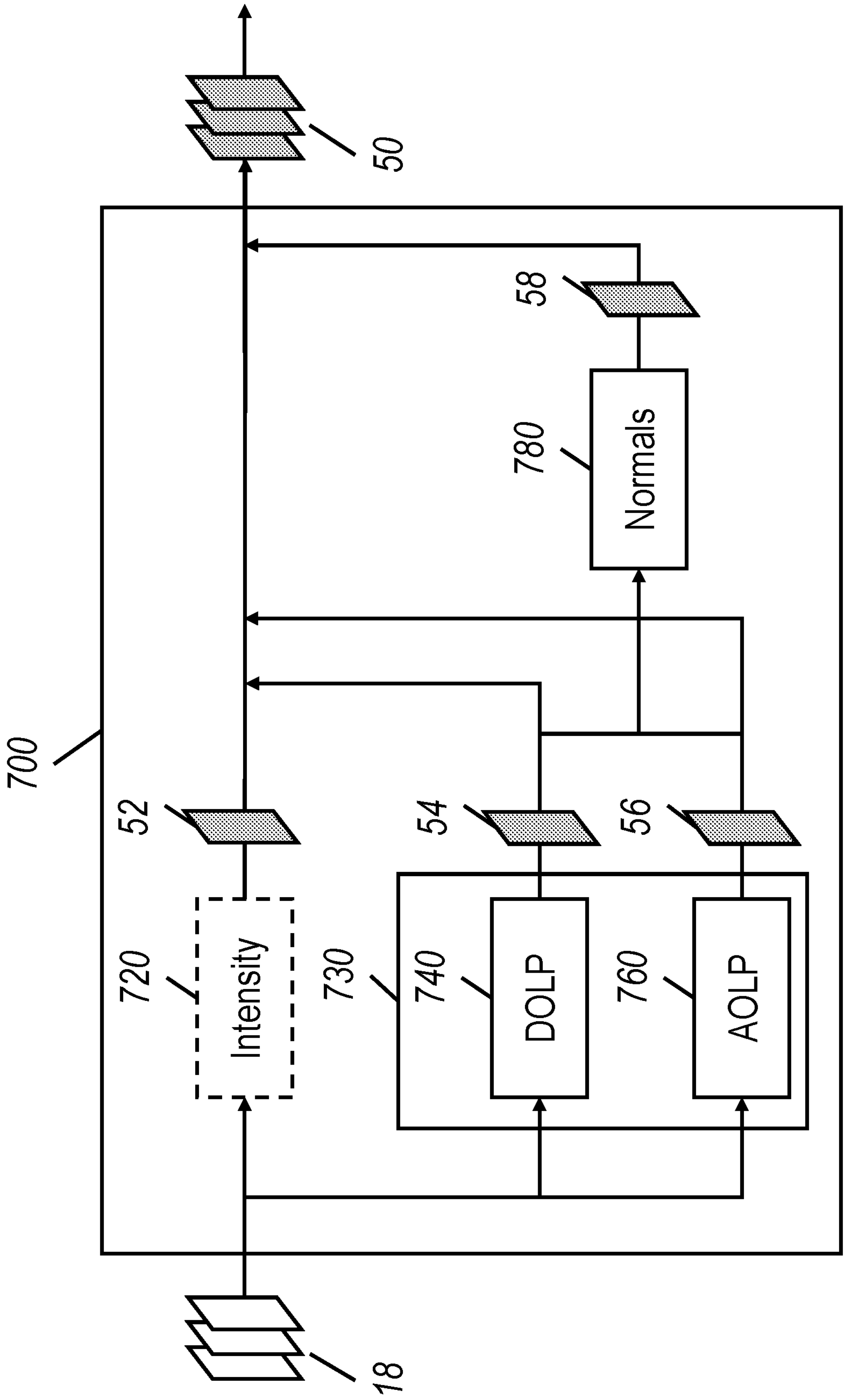

50
700
58
Normals
780
52
54
56
720
Intensity
730
740
DOLP
760
AOLP
18

FIG. 7A

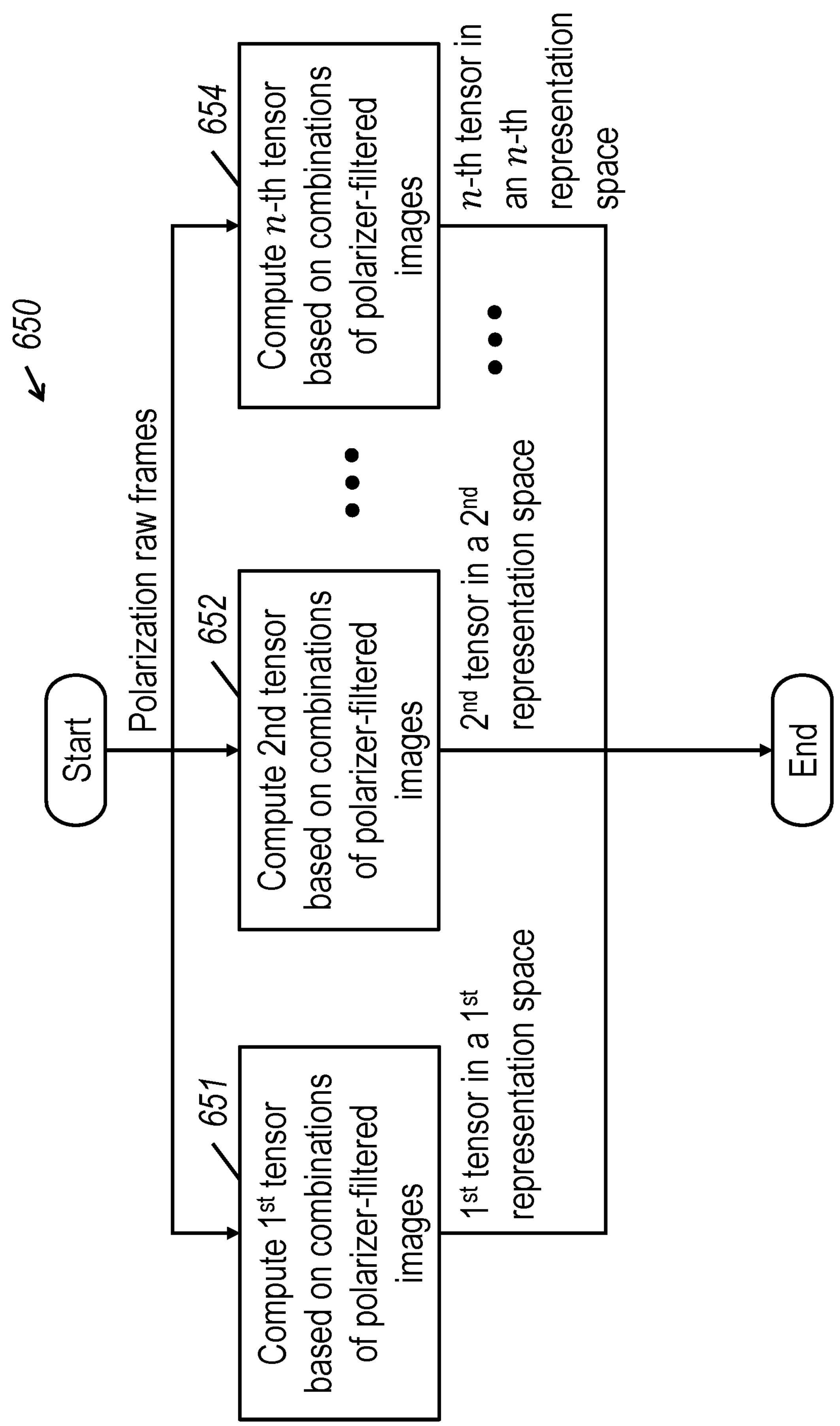

650
Start
Polarization raw frames
651
Compute 1st tensor based on combinations of polarizer-filtered images
1st tensor in a 1st representation space
652
Compute 2nd tensor based on combinations of polarizer-filtered images
2nd tensor in a 2nd representation space
654
Compute n-th tensor based on combinations of polarizer-filtered images
n-th tensor in an n-th representation space
End

FIG. 7B

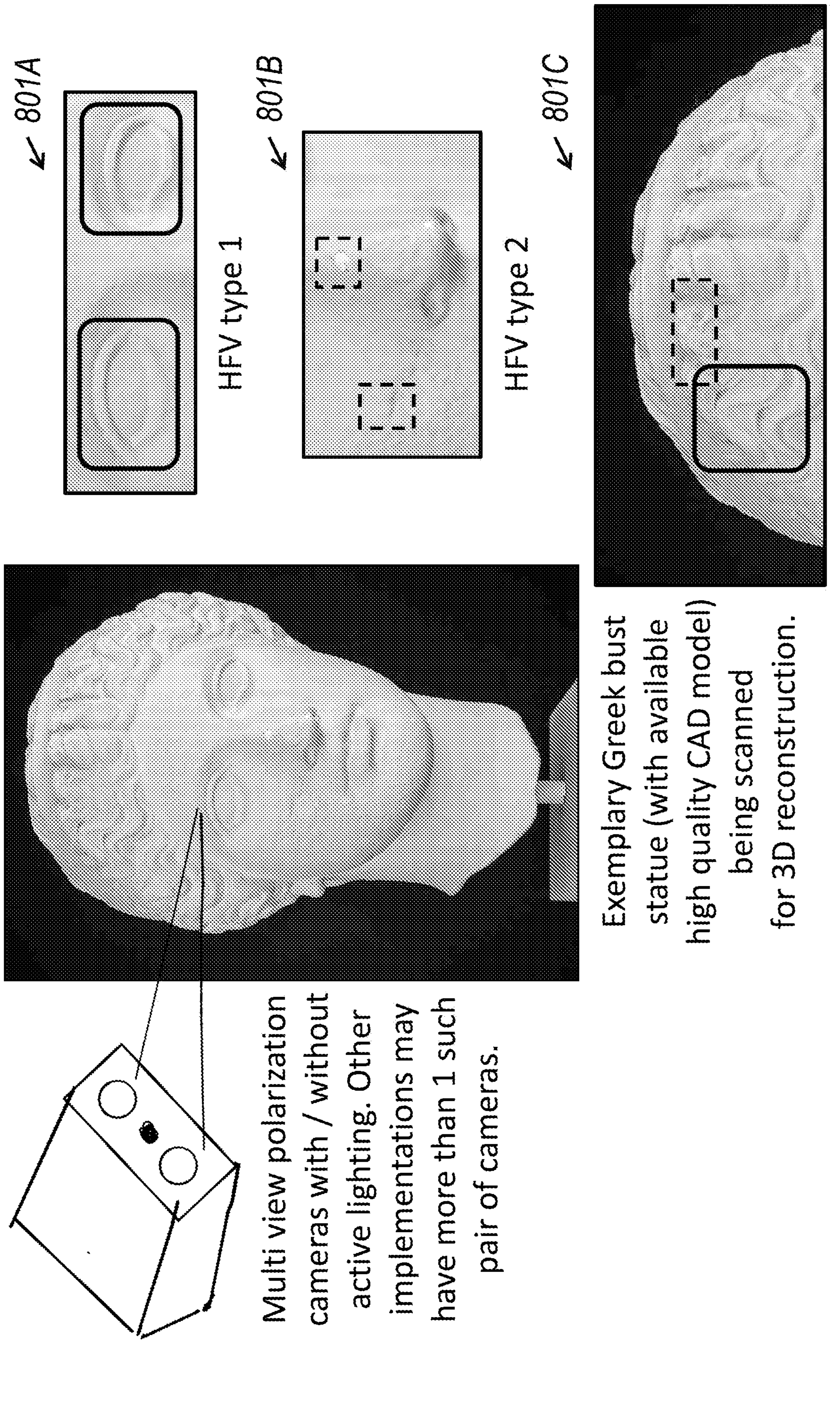

801A
HFV type 1
801B
HFV type 2
801C
HFV type 3
Multi view polarization cameras with / without active lighting. Other implementations may have more than 1 such pair of cameras.
Exemplary Greek bust statue (with available high quality CAD model) being scanned for 3D reconstruction.

FIG. 8A

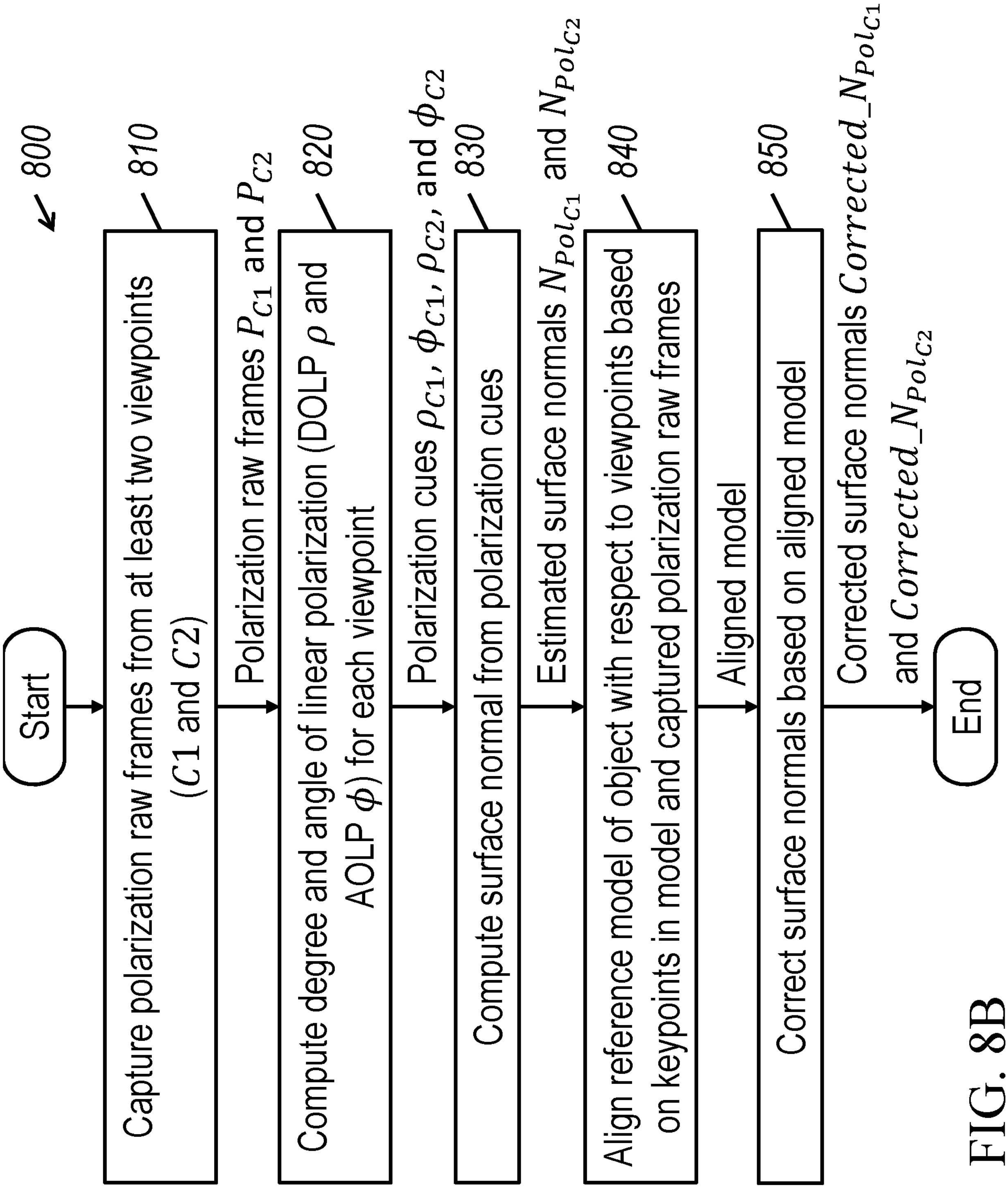

800
Start
810
Capture polarization raw frames from at least two viewpoints ($C1$ and $C2$)
Polarization raw frames P_{C1} and P_{C2}
820
Compute degree and angle of linear polarization (DOLP ρ and AOLP ϕ) for each viewpoint
Polarization cues ρ_{C1}, ϕ_{C1}, ρ_{C2}, and ϕ_{C2}
830
Compute surface normal from polarization cues
Estimated surface normals $N_{Pol_{C1}}$ and $N_{Pol_{C2}}$
840
Align reference model of object with respect to viewpoints based on keypoints in model and captured polarization raw frames
Aligned model
850
Correct surface normals based on aligned model
Corrected surface normals $Corrected\_N_{Pol_{C1}}$ and $Corrected\_N_{Pol_{C2}}$
End

FIG. 8B

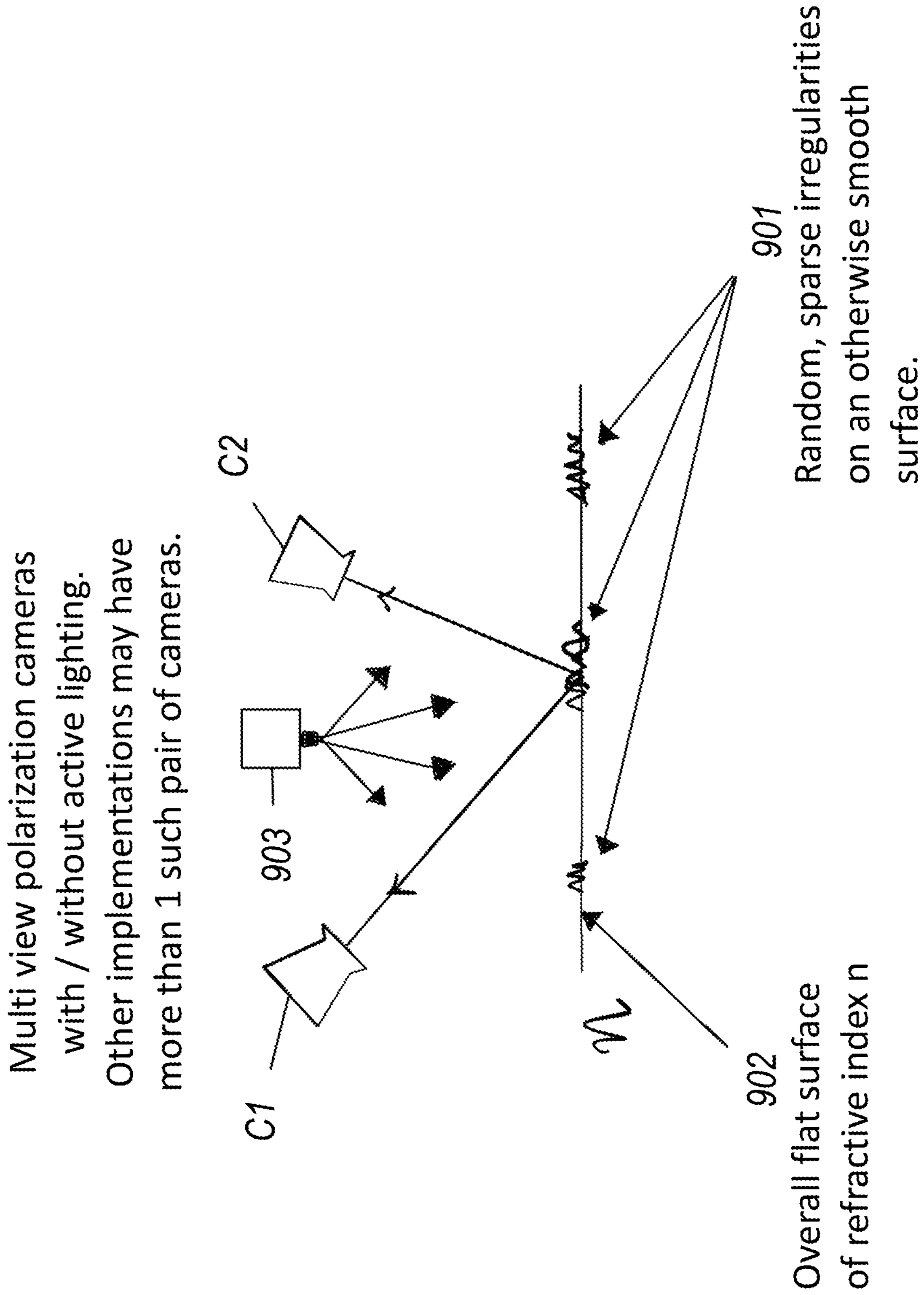

Multi view polarization cameras with / without active lighting. Other implementations may have more than 1 such pair of cameras.
C1
C2
903
901
Random, sparse irregularities on an otherwise smooth surface.
902
Overall flat surface of refractive index n

FIG. 9A

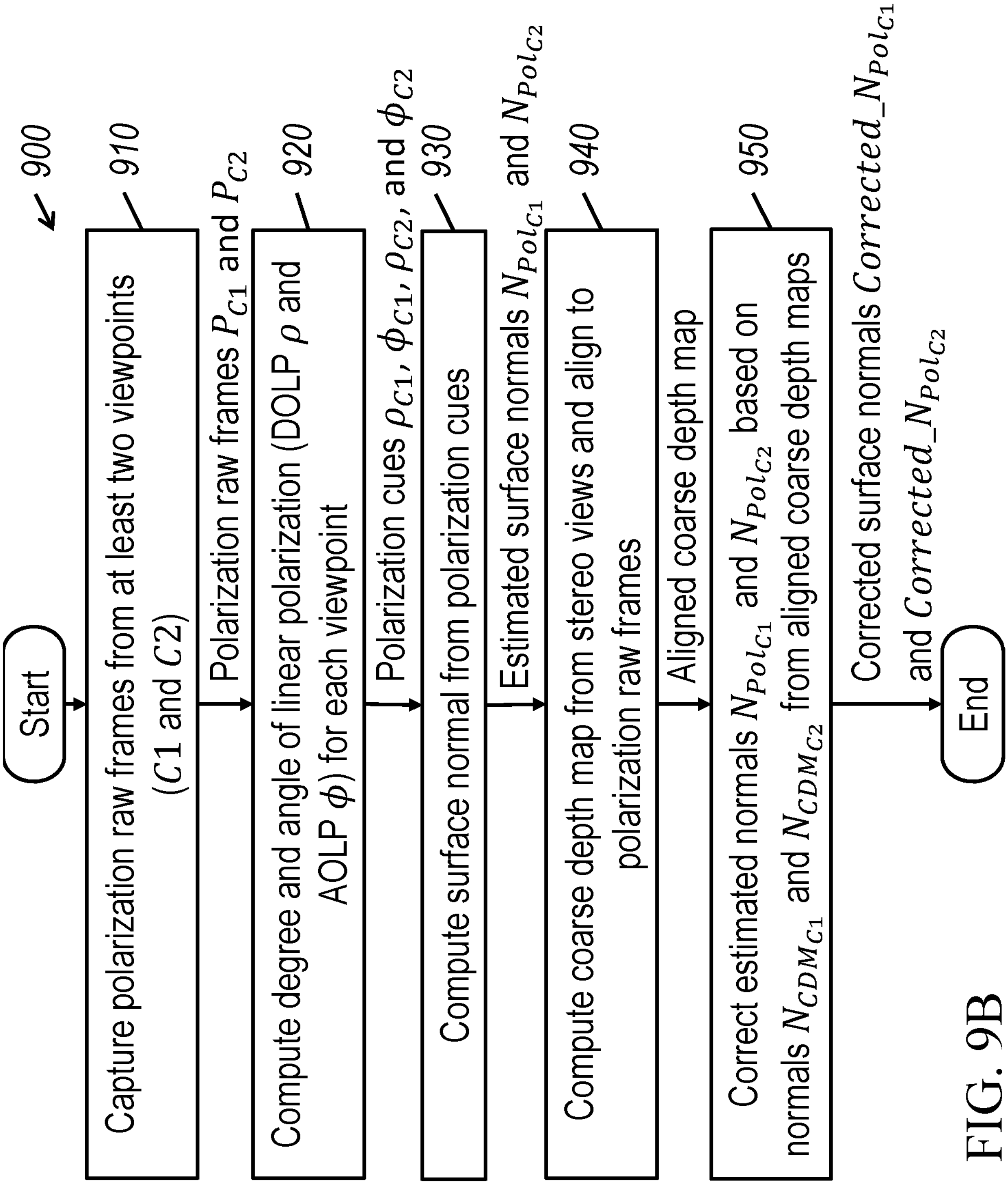

900
Start
910
Capture polarization raw frames from at least two viewpoints ($C1$ and $C2$)
Polarization raw frames P_{C1} and P_{C2}
920
Compute degree and angle of linear polarization (DOLP ρ and AOLP ϕ) for each viewpoint
Polarization cues ρ_{C1}, ϕ_{C1}, ρ_{C2}, and ϕ_{C2}
930
Compute surface normal from polarization cues
Estimated surface normals $N_{Pol_{C1}}$ and $N_{Pol_{C2}}$
940
Compute coarse depth map from stereo views and align to polarization raw frames
Aligned coarse depth map
950
Correct estimated normals $N_{Pol_{C1}}$ and $N_{Pol_{C2}}$ based on normals $N_{CDM_{C1}}$ and $N_{CDM_{C2}}$ from aligned coarse depth maps
Corrected surface normals $Corrected\_N_{Pol_{C1}}$ and $Corrected\_N_{Pol_{C2}}$
End

FIG. 9B

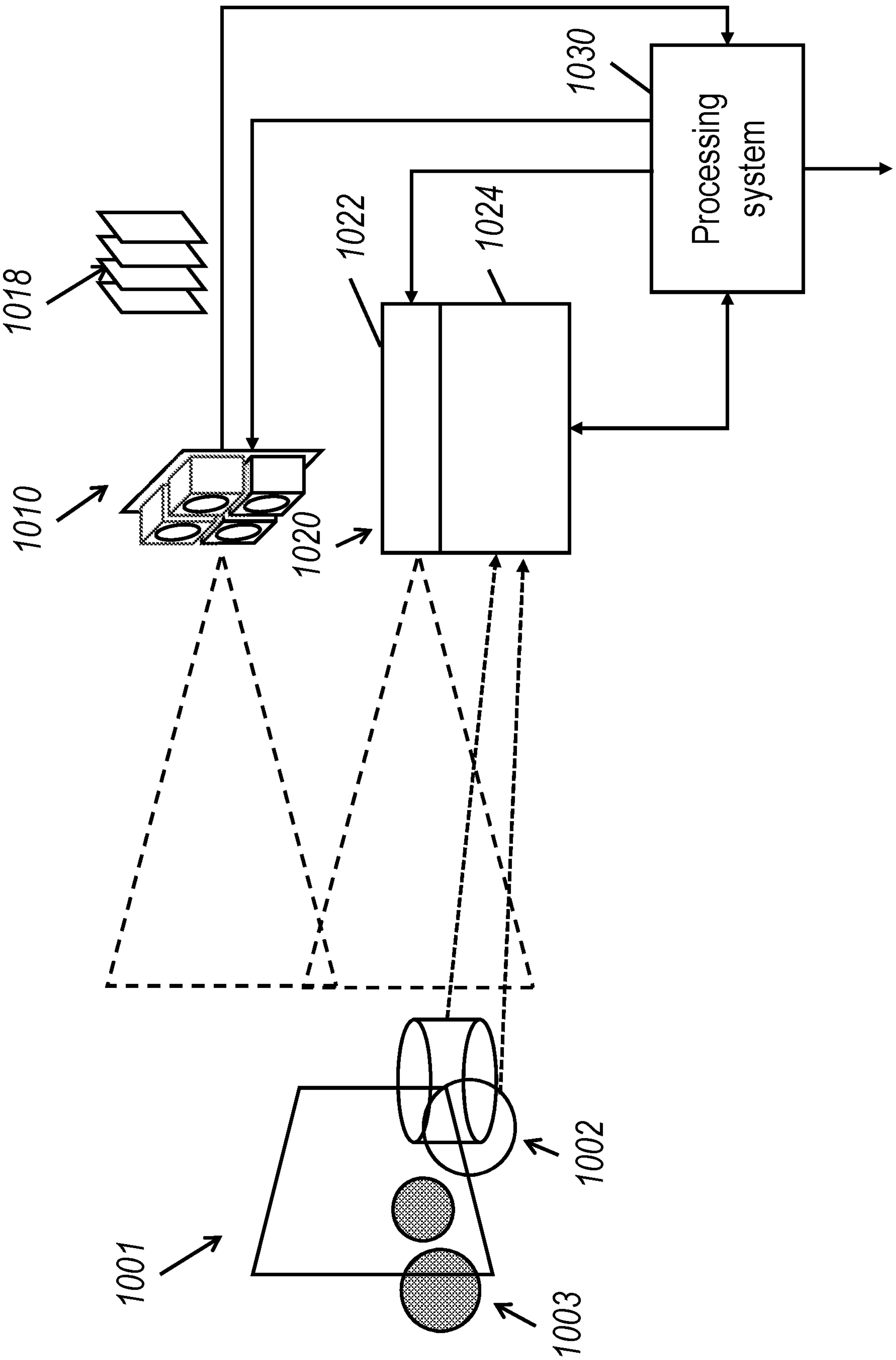

1030
Processing system
1018
1022
1024
1010
1020
1001
1002
1003

FIG. 10A

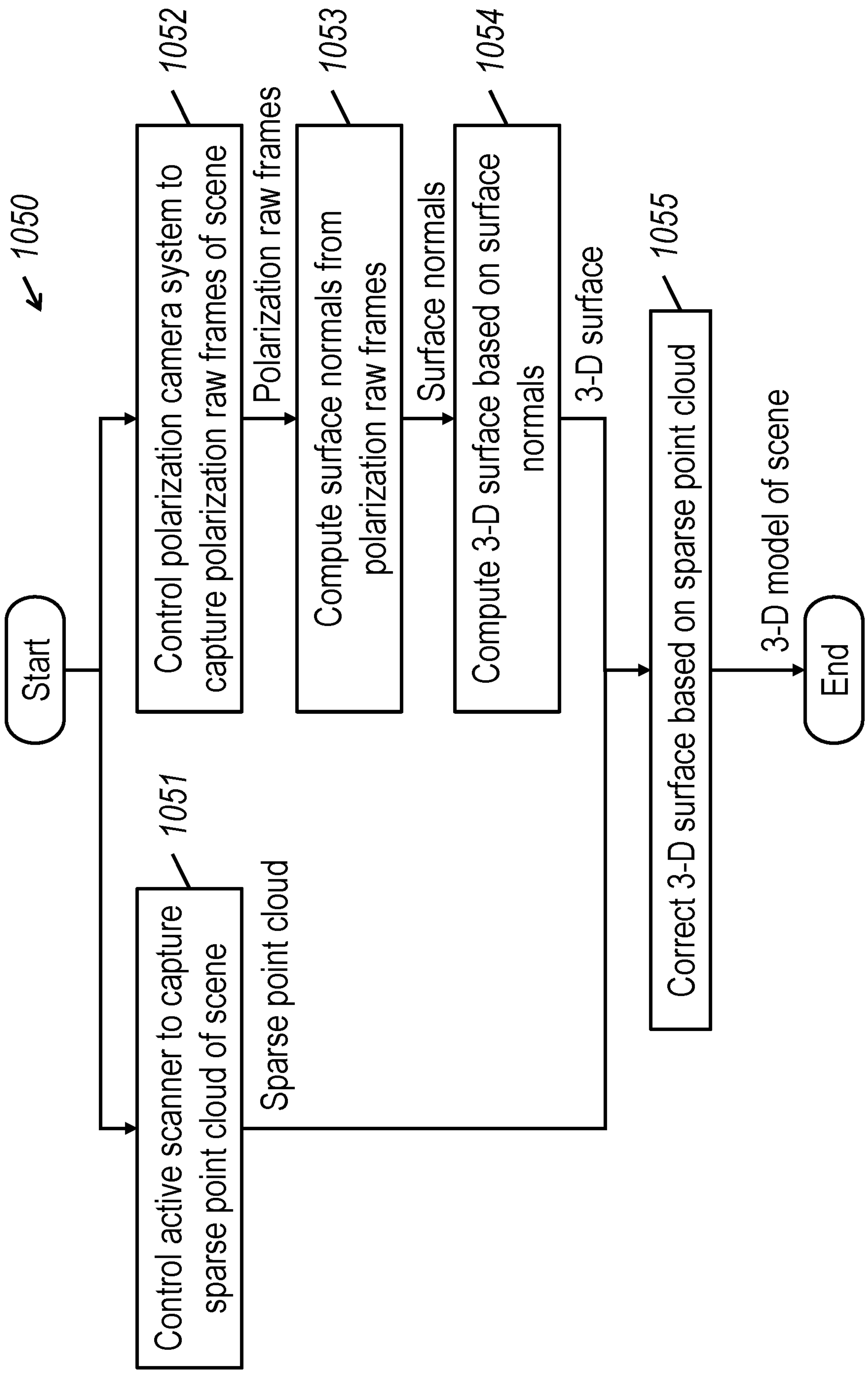

1050
Start
Control active scanner to capture sparse point cloud of scene
1051
Sparse point cloud
Control polarization camera system to capture polarization raw frames of scene
1052
Polarization raw frames
Compute surface normals from polarization raw frames
1053
Surface normals
Compute 3-D surface based on surface normals
1054
3-D surface
Correct 3-D surface based on sparse point cloud
1055
3-D model of scene
End

FIG. 10B

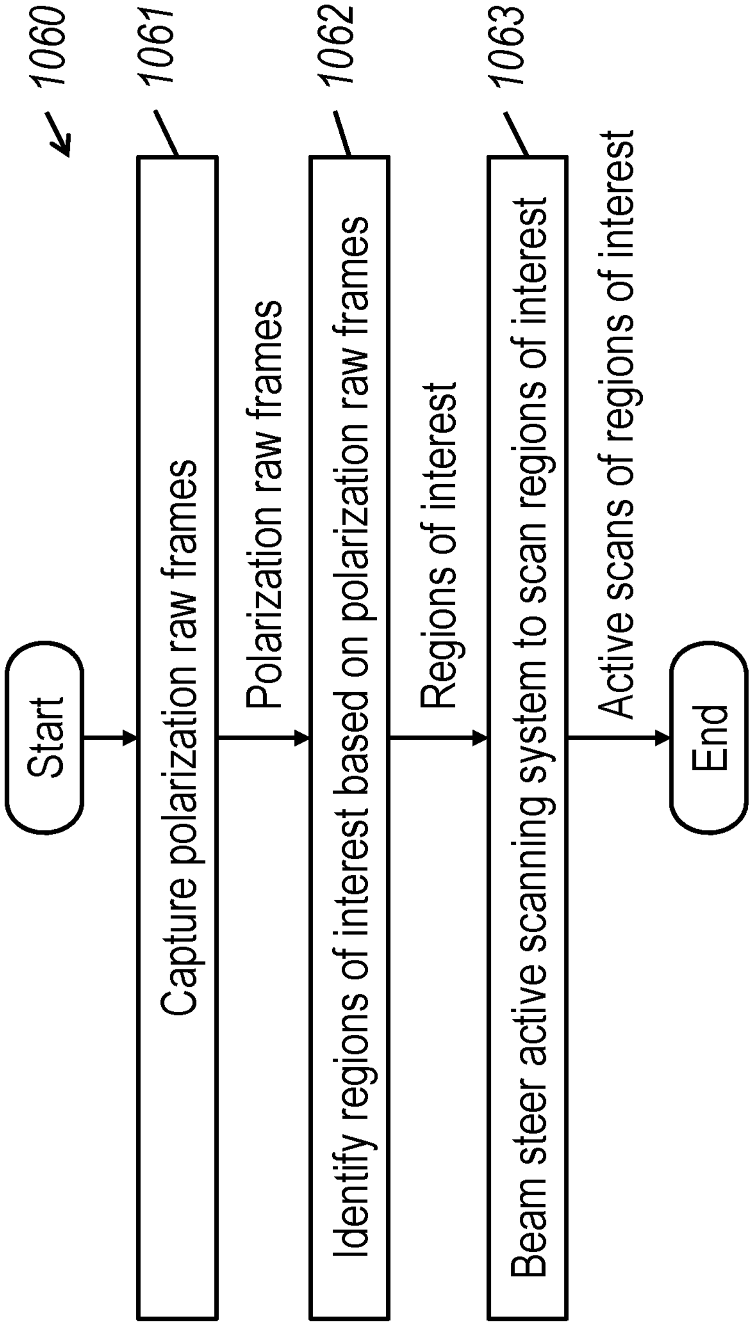

1060
Start
1061
Capture polarization raw frames
Polarization raw frames
1062
Identify regions of interest based on polarization raw frames
Regions of interest
1063
Beam steer active scanning system to scan regions of interest
Active scans of regions of interest
End

FIG. 10C

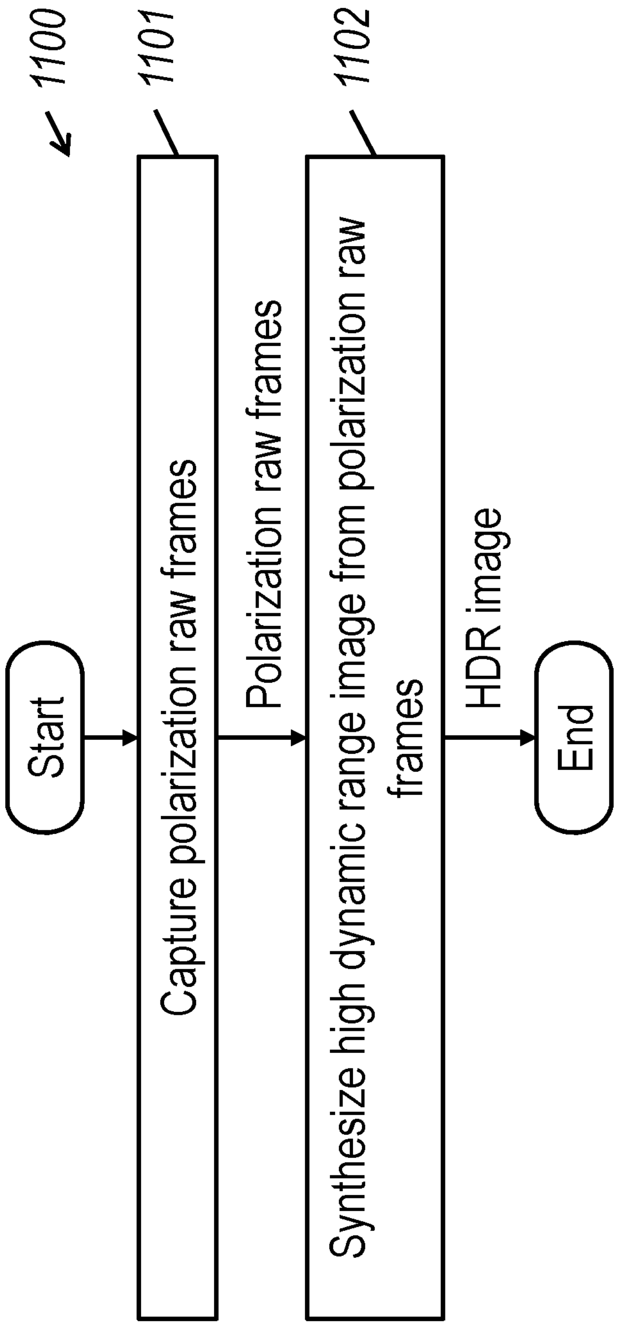
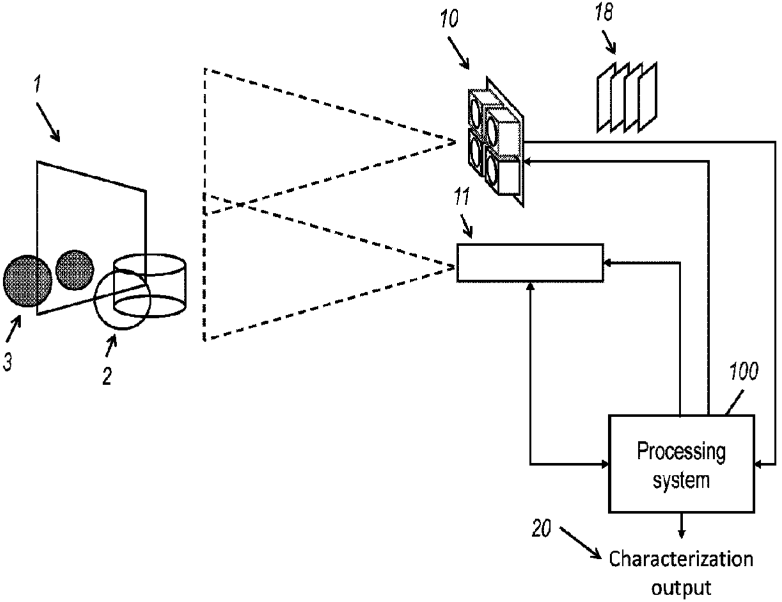

1100
1101
1102
Start
Capture polarization raw frames
Polarization raw frames
Synthesize high dynamic range image from polarization raw frames
HDR image
End

FIG. 11


1
10
18
11
3
2
100
Processing system
20
Characterization output