(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 2021/071995 A1

(51) International Patent Classification:
G01B 11/24 (2006.01)

(21) International Application Number:
PCT/US2020/054645

(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). **VENKATARMAN, 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.**; 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, 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,

(54) Title: SYSTEMS AND METHODS FOR SURFACE NORMALS SENSING WITH POLARIZATION

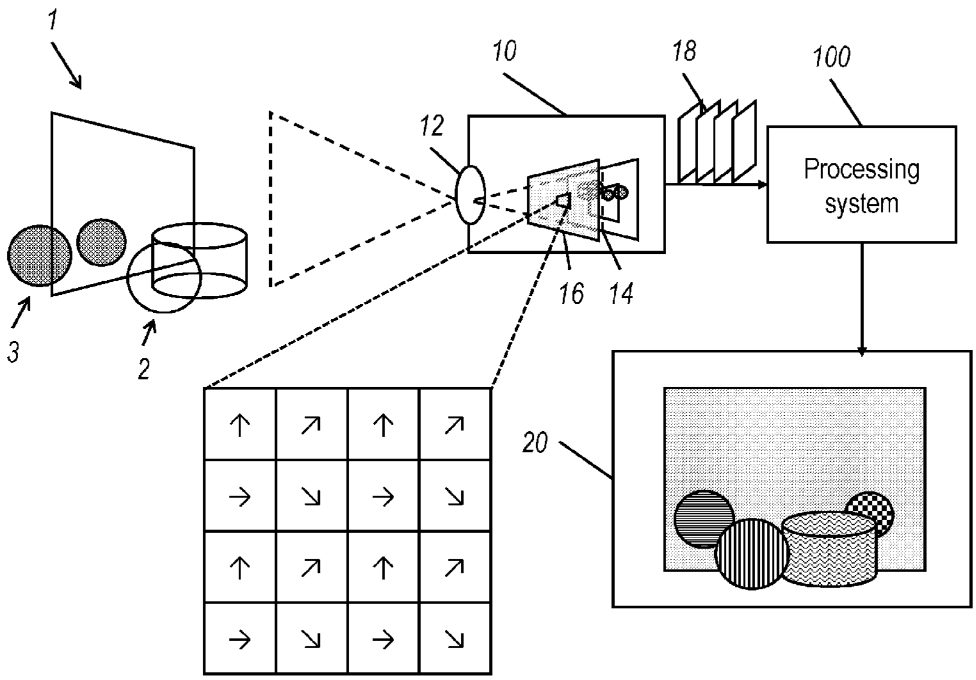

FIG. 1A

(57) **Abstract:** A method of performing surface profilometry includes receiving one or more polarization raw frames of a printed layer of a physical object undergoing additive manufacturing, the one or more polarization raw frames being captured at different polarizations by one or more polarization cameras, extracting one or more polarization feature maps in one or more polarization representation spaces from the one or more polarization raw frames, obtaining a coarse layer depth map of the printed layer, generating one or more surface-normal images based on the coarse layer depth map and the one or more polarization feature maps, and generating a 3D reconstruction of the printed layer based on the one or more surface-normal images.

WO 2021/071995 A1

[Continued on next page]

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 SURFACE NORMALS SENSING 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, the entire disclosures of which are incorporated by reference herein.

FIELD

[0002] Aspects of embodiments of the present disclosure are generally related to image sensor and processing systems and methods of using the same.

BACKGROUND

[0003] Sensor systems and imaging systems such as radar, lidar, cameras (e.g., visible light and/or infrared cameras), 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.

[0004] The above information disclosed in this Background section is only for enhancement of understanding of the present disclosure, and therefore it may contain information that does not form the prior art that is already known to a person of ordinary skill in the art.

SUMMARY

[0005] 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 such as digital cameras. Some aspects of embodiments of the present disclosure relate to camera systems configured to detect the polarization of light.

[0006] According to some embodiments of the present invention, there is provided a method of performing surface profilometry, the method including: receiving one or more polarization raw frames of a printed layer of a physical object undergoing additive manufacturing, the one or more polarization raw frames being captured at different polarizations by one or more polarization cameras; extracting one or more polarization feature maps in one or more polarization representation spaces from the one or more polarization raw frames; obtaining a coarse layer depth map of the printed layer; generating one or more surface-normal images based on the coarse layer depth map and the one or more polarization feature maps; and generating a 3D reconstruction of the printed layer based on the one or more surface-normal images.

[0007] In some embodiments, the one or more polarization feature maps in the one or more polarization representation spaces include: a degree of linear polarization (DOLP) image in a DOLP representation space; and an angle of linear polarization (AOLP) image in an AOLP representation space.

[0008] In some embodiments, the one or more polarization cameras include: a first polarization camera configured to capture a first partition of a print bed on which the printed layer of the physical object resides; and a second polarization camera at a distance from the first polarization camera and configured to capture a second partition of the print bed.

[0009] In some embodiments, the obtaining the coarse layer depth map of the printed layer includes: constructing the coarse layer depth map based on a parallax shift between polarization raw frames captured by the first and second polarization cameras.

[0010] In some embodiments, the obtaining the coarse layer depth map of the printed layer includes: receiving a computer-aided-design (CAD) layer model corresponding to the printed layer of the physical object.

[0011] In some embodiments, the generating the one or more surface-normal images includes: determining a pose of the printed layer with respect to each of the one or more polarization cameras; transforming the coarse layer depth map into one or more camera spaces corresponding to the one or more polarization cameras to generate one or more transformed coarse layer depth maps; and correcting the one or more surface-normal images based on the one or more transformed coarse layer depth maps.

[0012] In some embodiments, the generating the 3D reconstruction of the printed layer includes: integrating surface normals of the one or more surface-normal images over a sample space to determine a shape of a surface of the printed layer.

[0013] According to some embodiments of the present invention, there is provided a surface profilometry system including: one or more polarization cameras including a polarizing filter, the one or more polarization cameras being configured to capture polarization raw frames at different polarizations; and a processing system including a processor and memory storing instructions that, when executed by the processor, cause the processor to perform: receiving one or more polarization raw frames of a printed layer of a physical object undergoing additive manufacturing, the one or more polarization raw frames being captured at different polarizations by the one or more polarization cameras; extracting one or more polarization feature maps in one or more polarization representation spaces from the one or more polarization raw frames; obtaining a coarse layer depth map of the printed layer; generating one or more surface-normal images based on the coarse layer depth map and the one or more polarization feature maps; and generating a 3D reconstruction of the printed layer based on the one or more surface-normal images.

[0014] In some embodiments, the one or more polarization feature maps in the one or more polarization representation spaces include: a degree of linear polarization (DOLP) image in a DOLP representation space; and an angle of linear polarization (AOLP) image in an AOLP representation space.

[0015] In some embodiments, the one or more polarization cameras include: a first polarization camera configured to capture a first partition of a print bed on which the printed layer of the physical object resides; and a second polarization camera at a distance from the first polarization camera and configured to capture a second partition of the print bed.

[0016] In some embodiments, the obtaining the coarse layer depth map of the printed layer includes: constructing the coarse layer depth map based on a parallax shift between polarization raw frames captured by the first and second polarization cameras.

[0017] In some embodiments, the obtaining the coarse layer depth map of the printed layer includes: receiving a computer-aided-design (CAD) layer model corresponding to the printed layer of the physical object.

[0018] In some embodiments, the generating the one or more surface-normal images includes: determining a pose of the printed layer with respect to each of the one or more polarization cameras; transforming the coarse layer depth map into one or more camera spaces corresponding to the one or more polarization cameras to generate one or more transformed coarse layer depth maps; and correcting the one or more surface-normal images based on the one or more transformed coarse layer depth maps.

[0019] In some embodiments, the generating the 3D reconstruction of the printed layer includes: integrating surface normals of the one or more surface-normal images over a sample space to determine a shape of a surface of the printed layer.

[0020] In some embodiments, the memory further stores instructions that, when executed by the processor, cause the processor to further perform: providing the 3D reconstruction of the printed layer to a 3D printing system as a control feedback, wherein the 3D printing system being configured to additively manufacture the physical object layer by layer.

[0021] In some embodiments, operations of the one or more polarization cameras, the processing system, and the 3D printing system are synchronized via a synchronization signal.

[0022] According to some embodiments of the present invention, there is provided a method of capturing a 3D image of a face, the method including: receiving one or more polarization raw frames of the face, the one or more polarization raw frames being captured at different polarizations by a polarization camera at a distance from the face; extracting one or more polarization feature maps in one or more polarization representation spaces from the one or more polarization raw frames; and generating a 3D reconstruction of the face based on the one or more polarization feature maps via a facial reconstruction neural network.

[0023] In some embodiments, the one or more polarization feature maps in the one or more polarization representation spaces include: a degree of linear polarization (DOLP) image in a DOLP representation space; and an angle of linear polarization (AOLP) image in an AOLP representation space.

[0024] In some embodiments, the one or more polarization feature maps include estimated surface normals, and the extracting the one or more polarization feature maps includes: generating the estimated surface normals based on the one or more polarization raw frames.

[0025] In some embodiments, the facial reconstruction neural network is trained to compute corrected surface normals based on the estimated surface normals, and to generate the 3D reconstruction of the face based on the corrected surface normals.

[0026] In some embodiments, the facial reconstruction neural network includes a trained polarized convolutional neural network (CNN).

[0027] According to some embodiments of the present invention, there is provided a 3D imaging system for capturing a 3D image of a face, the 3D imaging system including: a polarization camera including a polarizing filter and configured to capture one or more polarization raw frames at different polarizations; and a processing system including a processor and memory storing instructions that, when executed by the processor, cause the processor to perform: receiving one or more polarization raw frames of the face, the one or more polarization raw frames being captured at different polarizations by the polarization camera; extracting one or more polarization feature maps in one or more polarization representation spaces from the one or more polarization raw frames; and generating a 3D reconstruction of the face based on the one or more polarization feature maps via a facial reconstruction neural network.

[0028] In some embodiments, the one or more polarization feature maps in the one or more polarization representation spaces include: a degree of linear polarization (DOLP) image in a DOLP representation space; and an angle of linear polarization (AOLP) image in an AOLP representation space.

[0029] In some embodiments, the one or more polarization feature maps include estimated surface normals, and the extracting the one or more polarization feature maps includes: generating the estimated surface normals based on the one or more polarization raw frames.

[0030] In some embodiments, the facial reconstruction neural network is trained to compute corrected surface normals based on the estimated surface normals, and to generate the 3D reconstruction of the face based on the corrected surface normals.

[0031] In some embodiments, the facial reconstruction neural network includes a trained polarized convolutional neural network (CNN).

[0032] According to some embodiments of the present invention, there is provided a method of capturing a 3D image of a face, the method including: receiving one or more polarization raw frames of the face, the one or more polarization raw frames being captured at different polarizations by a polarization camera; extracting estimated polarization cues from the one or more polarization raw frames; generating estimated surface normals based on the estimated polarization cues; generating an initial coarse depth map of the face; refining the estimated polarization cues and the initial coarse depth map to generate refined polarization cues and a refined depth map; generating corrected surface normals based on the refined polarization cues and the refined depth map; and generating a 3D reconstruction of the face based on the corrected surface normals.

[0033] In some embodiments, the generating the initial coarse depth map of the face includes: receiving a 2D color image of the face; and computing the initial coarse depth map based on the 2D color image of the face.

[0034] In some embodiments, the generating the initial coarse depth map of the face includes: receiving a 3D model of a generic human face; and generating the initial coarse depth map based on the 3D model of the generic human face.

[0035] In some embodiments, the method further includes: providing the estimated surface normals and the corrected surface normals to a facial reconstruction neural network as a set of training data.

[0036] According to some embodiments of the present invention, there is provided a method of computing a prediction, the method including: receiving a surface-normal image corresponding to an object, the surface-normal image including surface-normal information for each pixel of the image; and computing the prediction based on the surface-normal image.

[0037] In some embodiments, a red color value, a green color value, and blue color values of a pixel of the surface normal image encode an x-axis component, a y-axis component, and a z-axis component of a surface normal of the object at the pixel.

[0038] In some embodiments, the red, green, and blue color values of the pixel are respectively the x-axis component, the y-axis component, and the z-axis component of the surface normal of the object at the pixel.

[0039] In some embodiments, the prediction is a probability vector, the computing the prediction includes supplying the surface-normal image to a trained classifier, and the trained classifier is configured to identify image characteristics of the surface-normal image and to output the probability vector, each element of the probability vector being a probability value corresponding to one of possible image characteristics.

[0040] In some embodiments, the trained classifier includes a plurality of statistical models corresponding to the possible image characteristics.

[0041] In some embodiments, the image characteristics include facial expressions or object types.

[0042] In some embodiments, the object includes a vehicle, a face, or a body.

BRIEF DESCRIPTION OF THE DRAWINGS

[0043] The accompanying drawings, together with the specification, illustrate example embodiments of the present disclosure, and, together with the description, serve to explain the principles of the present disclosure.

[0044] FIG. 1A is a schematic block diagram of a system using a polarization camera according to some embodiments of the present invention.

[0045] FIG. 1B is a perspective view of a polarization camera module according to some embodiments of the present disclosure.

[0046] FIG. 1C is a cross sectional view of a portion of a polarization camera module according to one embodiment of the present disclosure.

[0047] FIG. 1D is a perspective view of a stereo polarization camera system according to one embodiment of the present disclosure.

[0048] 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.

[0049] FIG. 3 is a high-level depiction of the interaction of light with transparent objects and non-transparent objects.

[0050] FIG. 4 is a flowchart of a method for estimating polarization cues under parallax ambiguities according to one embodiment of the present disclosure.

[0051] FIG. 5A is a perspective view of a multi-spectral stereo polarization camera system according to one embodiment of the present disclosure.

[0052] 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.

[0053] 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.

[0054] FIG. 6A is a block diagram of processing circuit for computing surface characterization outputs based on polarization data according to one embodiment of the present invention.

[0055] FIG. 6B is a flowchart of a method for performing surface characterization based on input images to compute a surface characterization output according to one embodiment of the present invention.

[0056] FIG. 7A is a block diagram of a feature extractor according to one embodiment of the present invention.

[0057] FIG. 7B is a flowchart depicting a method according to one embodiment of the present invention for extracting features from polarization raw frames.

[0058] FIG. 8A is an illustration of a Greek bust statue being scanned by an exemplary implementation of the imaging setup proposed in this invention.

[0059] FIG. 8B is a flowchart of a method for 3D surface reconstruction using polarization according to one embodiment of the present disclosure.

[0060] 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.

[0061] FIG. 9B is a flowchart of a method for 3D surface reconstruction of flat or geometrically simple surfaces using polarization according to one embodiment of the present disclosure.

[0062] FIG. 10A is a block diagram of various components of the surface profilometry system, according to some embodiments of the present disclosure.

[0063] FIG. 10B is a schematic diagram illustrating the spatial relation of the one or more polarization camera modules of the surface profilometry system and the print bed, according to some embodiments of the present disclosure.

[0064] FIG. 10C is a flow diagram of a method for performing surface profilometry based on polarization raw images, according to some embodiments of the present disclosure.

[0065] FIG. 11A illustrates the 3D imaging system utilizing a neural network for 3D reconstruction, according to some embodiments of the present disclosure.

[0066] FIG. 11B illustrates the 3D imaging system using a coarse depth map for 3D reconstruction, according to some embodiments of the present disclosure.

[0067] FIG. 12 is a block diagram of a predictor, according to one embodiment of the present invention.

DETAILED DESCRIPTION

[0068] The detailed description set forth below is intended as a description of example embodiments of a system and method for 3D imaging and processing using light polarization, provided in accordance with the present disclosure, and is not intended to represent the only forms in which the present disclosure may be constructed or utilized. The description sets forth the features of the present disclosure in connection with the illustrated embodiments. It is to be understood, however, that the same or equivalent functions and structures may be accomplished by different embodiments that are also intended to be encompassed within the scope of the disclosure. As denoted elsewhere herein, like element numbers are intended to indicate like elements or features.

[0069] FIG. 1A is a schematic block diagram of a system using a polarization camera according to some embodiments of the present invention. In the arrangement shown in FIG. 1A, 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).

[0070] 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 some 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).

[0071] As one example, FIG. 1A 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. 1A is capable of concurrently or simultaneously 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.

[0072] 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, 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 some 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).

[0073] 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 (also referred to as the observed object). 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 100 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.

[0074] 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.

[0075] The polarization raw frames 18 are supplied to a processing circuit 100, described in more detail below, which computes a characterization output 20 based on the polarization raw frames 18. In the embodiment shown in FIG. 1A, the characterization output 20 is an instance segmentation map identifying instances of different objects 2 and 3 that are present in the scene 1.

[0076] FIG. 1B is a perspective view of a polarization camera module according to some embodiments of the present disclosure. FIG. 1C is a cross sectional view of a portion of a polarization camera module according to one embodiment of the present disclosure.

[0077] 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 (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), 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.

[0078] For example, in the embodiment of the polarization camera module 10′ shown in FIG. 1B, four cameras 10A′, 10B′, 10C′, and 10D′ are arranged in a 2×2 grid to form a camera array or camera system, 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.

[0079] FIG. 1C shows a cross sectional view of two of the polarization cameras 10A′ and 10B′ shown in FIG. 1B. As seen in FIG. 1C, 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′.

[0080] 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. 1B and 1C, 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.

[0081] While not shown in FIG. 1C, in some embodiments of the present disclosure, each polarization camera may also include a color filter having 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.

[0082] FIG. 1D 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.

[0083] Accordingly, FIG. 1D 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. 1D, the first polarization camera module 10-1″ includes polarization cameras 10A″, 10B″, 10C″, and 10D″ arranged in a 2×2 array similar to that shown in FIGs. 1B and 1C. 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 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°.

[0084] Embodiments of the present disclosure are not limited to the particular embodiment shown in FIG. 1D. 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 an array and is 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.

[0085] 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. 1A. 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. 1A) such that each individual polarization camera captures images where the pixels detect light in accordance with a mosaic pattern of different polarizing filters (e.g., polarizing filters at different angles). For example, in the embodiment shown in FIG. 1A, 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. 1A, to enable the polarization cameras to capture polarization raw frames of different polarizations.

[0086] 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).

[0087] 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.

[0088] 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.

[0089] 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 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.

[0090] 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). 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.

[0091] 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.

[0092] *Polarization feature representation spaces*

[0093] 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).

[0094] 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.

[0095] 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 to the light 313 reflected off the background object 303 and passing through the transparent object 302.

[0096] 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).

[0097] 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 710 to detect transparent objects, as described in more detail below. In some embodiments, the predictor 710 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).

[0098] 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.

[0099] 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}.

[00100] The relationship between $I_{\phi_{pol}}$ and intensity I, DOLP ρ, and AOLP ϕ at each pixel can be expressed as:

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

[00101] 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.

[00102] 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:

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

$$\phi = \theta_a \tag{3}$$

and when the specular reflection is dominant:

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

$$\phi = \theta_a - \frac{\pi}{2} \tag{5}$$

[00103] 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.

[00104] 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.

[00105] 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.

[00106] 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:

$$I = I_r + I_t \tag{6}$$

[00107] 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:

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

[00108] 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:

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

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

[00109] 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.

[00110] 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.

[00111] 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).

[00112] 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.

[00113] 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.

[00114] 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:

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

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

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

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

$$\theta_z = a\sin dd$$

[00115] 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, the entire disclosures of which are incorporated by reference herein.

[00116] <u>Computing Polarization Cues from Multi-camera Arrays</u>

[00117] 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 polarization (DOLP) ρ generally assume that the polarization raw frames are captured from the same viewpoint.

[00118] 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.

[00119] 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.

[00120] FIG. 4 is a flowchart 400 of a method for estimating polarization cues under parallax ambiguities according to one embodiment of the present disclosure.

[00121] 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. 1B, 1C, and 1D 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.

[00122] 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 10′.

[00123] 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)).

[00124] 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 to compute new estimates DOLP ρ_i and AOLP ϕ_i.

[00125] 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., $|\rho_i - \rho_{i-1}|$). 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., $|\phi_i - \phi_{i-1}|$). 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 surface normal maps, instance segmentation maps, etc.

[00126] *Multi-spectral stereo with Polarization Imaging*

[00127] 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.

[00128] 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).

[00129] 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:

$$s \approx \sum_{i=0}^{n} w_i B_i \tag{1}$$

[00130] 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.

[00131] 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.

[00132] 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°.

[00133] 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.

[00134] 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.

[00135] 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″.

[00136] 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″.

[00137] 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.

[00138] 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).

[00139] 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.

[00140] 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.

[00141] 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

$$p = TS = TBw \tag{2}$$

[00142] 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), B represents the spectral basis functions, and w represents the coefficients for the basis functions.

[00143] Accordingly, applying equation (13) above enables computation of per-pixel polarization information as well as spectral information.

[00144] 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.

[00145] 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.

[00146] 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).

[00147] 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.

[00148] Image Segmentation using Polarimetric Cues

[00149] 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, the entire disclosures of which are incorporated by reference herein.

[00150] 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.

[00151] 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 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.

[00152] As shown in FIG. 6A, in some embodiments, a processing circuit 100 includes a feature extractor or feature extraction system 700 and a predictor 710 (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 710 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.

[00153] 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.

[00154] 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.

[00155] 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.

[00156] 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.

[00157] 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 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.

[00158] 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 710 for detecting surface characteristics based on the feature maps 50.

[00159] 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).

[00160] 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.

[00161] 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 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).

[00162] *<u>Augmenting 3D Surface Reconstruction with Polarization Imaging</u>*

[00163] 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 (3D) 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.

[00164] Capturing a high quality 3D reconstruction of a physical object for which a high-quality 3D 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 3D printing, a designer may create a 3D model of an object and supply the 3D model to a 3D printer, which fabricates a physical object based on the 3D model. During or after the 3D printing process, the physical object fabricated by the 3D 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 3D reconstruction of the surfaces of the physical object. This 3D reconstruction can then be compared, in software, to the designed 3D model to detect defects in the 3D printing process. Similar techniques may be applied to other manufacturing processes, such as for creating 3D reconstructions of the shapes of objects created through other manufacturing processes such as injection molding, die-casting, bending, and the like.

[00165] As one example, a stereo polarization camera system, such as that described above with respect to FIG. 1D, is used to image an object that is intended to be reconstructed in 3D, e.g., to create a 3D 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:

[00166] 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 3D model may also be present in the high-quality 3D model that was used to guide the fabrication of the physical object.

[00167] 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 3D 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 3D model.

[00168] 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).

[00169] 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).

[00170] 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 3D surface reconstruction using polarization according to one embodiment of the present disclosure.

[00171] 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. 1D. 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).

[00172] 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}. 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}). These surface normal from the two viewpoints may be denoted as $N_{Pol_{C1}}$ and $N_{Pol_{C2}}$.

[00173] 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.

[00174] 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.

[00175] 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.

[00176] In addition to only relying on the CAD model for resolving ambiguities and errors in 3D 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 3D reconstruction by enforcing view-point consistency between the cameras of the stereo polarization camera system.

[00177] 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 (P*n*P) (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 P*n*P 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 $N_{CAD_{C2}}$.

[00178] 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 P*n*P, 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.

[00179] 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. 1D, where different cameras have different viewpoints of surfaces of an object.

[00180] 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.

[00181] 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.

[00182] 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.

[00183] 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.

[00184] 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:

[00185] **N_P_C1** - shape-from-polarization (SFP) normals in Camera1

[00186] **N_P_C2** - SFP normals in Camera2

[00187] N_CAD_C1 - Normals in CAD reference model aligned with the object pose with respect to Camera1

[00188] N_CAD_C2 - Normals in CAD reference model aligned with the object pose with respect to Camera2

[00189] 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)

[00190] Trans_CAD_C1 - Transform used to align CAD reference model with the object in Camera1 image space obtained through multiview P*n*P

[00191] Trans_CAD_C2 - Transform used to align CAD reference model with the object in Camera2 image space obtained through multiview P*n*P

[00192] (~) - Consistency operator

[00193] 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 (~).

[00194] 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

[00195] 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 3D reconstruction of the object imaged by the stereo polarization camera system.

[00196] While the embodiments discussed above relate to the 3D reconstruction of 3D objects based on a high-quality 3D model such as a CAD design model, some aspects of embodiments of the present disclosure further relate to 3D reconstruction of generally flat surfaces or surfaces having known, simple geometry, using multi-view polarized camera system such as that shown in FIG. 1D. 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.

[00197] 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.

[00198] 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.

[00199] 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.

[00200] FIG. 9B is a flowchart of a method 900 for 3D surface reconstruction of flat or geometrically simple surfaces using polarization according to one embodiment of the present disclosure.

[00201] 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 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. 1D. 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).

[00202] In operation 920, 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}.

[00203] 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.

[00204] 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.

[00205] 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}.

[00206] 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.

[00207] 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:

[00208] **N_P_C1** - shape-from-polarization (SFP) normals in Camera1

[00209] **N_P_C2** - SFP normals in Camera2

[00210] **CDM** - Coarse Depth Map

[00211] **N_CDM_C1** - Normals in CDM in Camera1 image space

[00212] **N_CDM_C2** - Normals in CDM in Camera2 image space

[00213] **Trans_C2_C1** - Relative pose of Camera2 with respect to Camera1 obtained through extrinsic and intrinsic camera calibration

[00214] **Trans_CDM_C1** - Transform used to align CDM with the object in Camera1 image space

[00215] **Trans_CDM_C2** - Transform used to align CDM with the object in Camera2 image space

[00216] **(~)** - Consistency operator

[00217] **obeys_flatness()** - operator that checks if the normals being selected obey a flatness constraint

[00218] 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 (~).

[00219] **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

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

[00220] 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.).

[00221] 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.

[00222] *A Multi-Camera Polarization Enhanced System for Real-time Visualization based Feedback Control*

[00223] A number of industrial applications of imaging involve viewing volumes that are fixed and in controlled lighting conditions. For example, 3D printers have a fixed size print bed and the viewing volume that corresponds to this print bed is a controlled environment in terms of illumination, temperature, and humidity.

[00224] Aspects of the present disclosure are directed to a surface profilometry system configured to capture the surface profiles, layer-by-layer, of the objects being printed. This allows for the synthesis of a 3D model at the end of the print job. This may be of great use to an end customer as it allows one to understand the deviation of the printed object from its design target and thereby capture manufacturing tolerances on an object by object basis. For manufacturing of mission critical parts this information assumes even more importance. In addition, the layer-by-layer surface profile captured by the surface profilometry system allows the 3D printer to use that information to take corrective action if necessary to improve the print process and correct for any printed anomalies.

[00225] FIG. 10A is a block diagram of various components of the surface profilometry system 1000, according to some embodiments of the present disclosure. FIG. 10B is a schematic diagram illustrating the spatial relation of the one or more polarization camera modules 10′ of the surface profilometry system 1000 and the print bed 1004, according to some embodiments of the present disclosure.

[00226] Referring to FIGS. 10A-10B, in some embodiments, the surface profilometry system 1000 includes a 3D printing system 1002 for printing (e.g., additively manufacturing) a physical object 1006 on a print bed 1004, one or more polarization camera modules 10′ for imaging the surface profiles, layer-by-layer, of the objects being printed on the print bed 1004, and a processing system 100 for processing the polarization raw images 18 captured by the one or more polarization camera modules 10′. As described above with reference to FIGS. 1B-1C, each of the one or more polarization camera modules 10′ is an array (e.g., a 2x2 array) of polarization cameras, where each polarization camera has a linear polarization filter of a specific angle (for example, 0°, 45°, 90°, and 135°).

[00227] For each layer of the object 1006 being printed by the 3D printing system 1002, the processing circuit 100 receive polarization raw frames 18 that correspond to said printed layer of the object 1006. The processing circuit 100 extracts polarization cues from the polarization raw frames 18 and, with the aid of a coarse layer depth map, generates surface-normal images (also referred to as “normal images”). The processing circuit 100 then generates a corresponding 3D surface profile of the layer being printed. In some examples, this 3D surface profile may be used in post-processing analysis to determine deviations of the printed object 1006 from its design target on a layer-by-layer basis, which can be used to improve the printing process or enhance understanding of printing tolerances. Further, the processing circuit 100 may provide the 3D surface profile to the printing system 1002 as feedback (e.g., as a control feedback), which may allow the printing system 1002 to take corrective action in real time (e.g., by correcting printed anomalies). The operations of the printing system 1002, the one or more polarization camera modules 10′, and the processing circuit 100 may be synchronized by virtue of a synchronization signal supplied by the printing system 1002 or an external controller. In some examples, once a layer of the object 1006 is printed, the one or more polarization camera modules 10′ capture the polarization raw images, and the processing circuit 100 completes the processing of these images to generate a 3D profile of the layer, before the printing system 1002 prints the subsequent layer (or before it begins printing the subsequent layer).

[00228] In some examples, in order to capture details of the object 1006 with sufficient resolution, the surface profilometry system 1000 may be positioned close enough to the print bed 1004 that the diagonal field of view subtended from the camera may exceed an angular threshold (e.g., 100°). In such scenarios, using a single camera with a large field of view lens may present optical distortions or modulation transfer function (MTF) problems, wherein at large field heights the lens distortion effects may become so severe that the system MTF degrades significantly.

[00229] Accordingly, in some examples, the active area of the print bed 1004 is partitioned into two one more regions, each region of which is covered by a different polarization camera module 10′ in a manner that the diagonal field of view of the region to be covered does not exceed the angular threshold (e.g., 100°). This may lead to good control over field distortion and system MTF across all field heights. In some examples, each adjacent pair of partitions has an overlapping region which falls within the field of view of the corresponding adjacent polarization camera modules 10′. The overlapping region may aid in the alignment of images captured by the adjacent polarization camera modules 10′. However, embodiments of the present disclosure are not limited thereto, and the overlapping region may be negligibly small or non-existent, and alignment may be performed without the aid of an overlap region. According to some examples, image alignment may be performed based on prior calibrations using knowledge of the position and orientation of each polarization camera modules 10′ relative to the print bed 1004. The prior calibrations may be encoded as an extrinsic matrix of rotation and translation, which may be stored at the processing system 100.

[00230] In embodiments in which a single polarization camera modules 10′ is used per partition, the processing circuit 100 may derive the coarse depth map from the 3D CAD layer model used by the 3D printing system to print the layer of the object 1006 being observed. However, embodiments of the present invention are not limited thereto. According to some embodiments, a pair of polarization camera modules 10′ (such as the stereo polarization camera system 10″ described with reference to FIG. 1D) may be utilized by the surface profilometry system 1000 to capture the polarization cues of each partition of the print bed 1004. In such embodiments, the processing circuit 100 may construct the coarse depth map based on the parallax shift between the polarization raw frames captured by the pair (e.g., by the stereo polarization camera system 10″), as described above with respect to FIG. 1D, or may derive the coarse depth maps from the 3D CAD layer model.

[00231] FIG. 10C is a flow diagram of a method 1100 for performing surface profilometry of an object undergoing additive manufacturing based on polarization raw images, according to some embodiments of the present disclosure.

[00232] In some embodiments, the processing circuit 100 performs the method 1100 for each layer, or for each one of a subset of layers, of the object 1006 being printed.

[00233] In operation 1102, the processing circuit 100 receives polarization raw frames captured of a printed layer of the 3D object using one or more polarization camera modules 10′. The processing circuit 100 then extracts polarization feature maps or polarization images from polarization raw frames, in operation 1104. In so doing, the processing circuit 100 computes, based on the polarization images, degree and angle of linear polarization (DOLP ρ and AOLP ϕ) from Stokes vector formulation for each of the one or more cameras polarization camera modules 10′ (similarly to operation 920 of FIG. 9B).

[00234] In operation 1106, the processing circuit 100 obtains the coarse layer depth map corresponding to the layer of the object 1006 being printed. In some examples, the coarse layer depth map may be the CAD layer model provided by the printing system 1002. In examples in which the surface profilometry system 1000 utilizes a pair of polarization camera modules 10′ (e.g., the stereo polarization camera system 10″), rather than a single polarization camera module 10′, for capturing the polarization cues of each partition of the print bed 1004, the processing circuit 100 may construct the coarse depth map based on the parallax shift between the polarization raw frames captured by the pair.

[00235] In operation 1108, the processing circuit 100 generates surface-normal images based on the coarse layer depth map and the polarization images. As described above, in some embodiments, the processing circuit 100 calculates surface normals (e.g., Zenith θ_z and Azimuth θ_a) from the polarization images by using shape from polarization (SFP) technique as discussed above with respect to equations (2), (3), (4), and (5) for each of the one or more polarization camera modules 10′. The processing circuit 100 then determines the pose of the layer being printed with respect to each of the one or more polarization camera modules 10′ using Perspective-N-Point (P*n*P), and transforms the coarse layer depth map into the camera space(s) corresponding to the one or more polarization camera modules 10′ of the surface profilometry system 1000 (as, e.g., described with reference to operation 840 of FIG. 8B). The transformed coarse layer depth map(s) corresponding to the one or more camera modules 10′ are used by the processing circuit 100 as guidance constraints to correct high frequency azimuthal π ambiguity as well as the low frequency scaling error in zenith due to refractive distortion, and thus generate the corresponding one or more corrected surface-normal images, which may be robust to noisy SFP normals in cases of fronto-parallel facets and specularity. In embodiments in which the surface profilometry system 1000 constructs the coarse depth map based on the parallax shift between the polarization raw frames captured by a pair of polarization camera modules 10′ (e.g., the stereo polarization camera system 10″), may refine the coarse layer depth model and the polarization cues (e.g., DOLP ρ and AOLP ϕ) using operations 450-490 described with reference to FIG. 4.

[00236] In operation 1110, the processing circuit 100 generates a 3D reconstruction of layer being printed based on the corrected surface-normal image(s). As described above, the corrected surface normal(s) may be (independently) integrated over a sample space (Ω) to recover the shape of the entire surface of the printed layer or a portion thereof within the corresponding partition for each camera module 10′. In examples in which the print bed 1004 is divided into two or more partitions with corresponding two or more camera modules 10′, the recovered shape of each partition may be aligned and merged with the recovered shape(s) of the adjacent partition(s) to arrive at the 3D reconstruction of the entire layer being printed. In some examples, the alignment and merging may be performed at the stage of the corrected surface-normal images, and the resulting merged surface-normal image may correspond to the entire layer being printed. This

3D reconstruction, according to some embodiments, may have a higher resolution than what can be obtained by 3D imaging techniques of the related art.

[00237] The surface recovered from such integration is expected to match the shape constrained by the CAD layer model, and differences between the surface recovered from integration and the CAD layer model may indicate defective portions of the printed layer.

[00238] _Long Range 3D face Scans_

[00239] Face recognition is an important biometric recognition process that enables authentication of the user in numerous security and surveillance applications. Among existing authentication methods, 3D face scans are potentially more robust with lower false acceptance rates (FAR) and false rejection rates (FRR). As such, there has been growing interest in the field of 3D face authentication. However, 3D face scans using traditional means may not be easy to implement robustly. For instance, in mobile systems, active illumination is used to project a pattern on a user's face and then the projected pattern is sensed to retrieve the depth of key feature points. This technology is widely used but has a number of problems, which include: the added cost and power consumption resulting from using active illumination, the limited range of active illumination, as well as the inability to work reliably in all environmental conditions (e.g., lighting conditions).

[00240] Some aspects of embodiments of the present disclosure relate to a 3D imaging system that leverages light polarization and neural networks to 3D scan of a face.

[00241] FIG. 11A illustrates the 3D imaging system 1100, according to some embodiments of the present disclosure.

[00242] In some embodiments, the 3D imaging system 1100 includes the polarization camera module 10′, the feature extractor 700, and a facial reconstruction neural network 1104. The polarization camera module 10′, which is placed at a distance from the observed face 1102, images the face 1102 and produces a corresponding polarization raw image 18′. The distance between the observed face 1102 and the polarization camera module 10′ is not particularly limited and may be any suitable distance as long as the ability to capture the face 1102 is not limited by the optics of the polarization camera module 10′. The feature extractor 700 computes the feature maps 50, which include surface normals (e.g., estimated surface normals) 58, based on the polarization raw image 18′ via the process described above with respect to FIGS. 7A-7B. According to some embodiments, the facial reconstruction neural network 1104 generates a detailed 3D reconstruction 1106 of the observed face 1102 based on (e.g., based solely on) the feature maps 50 (e.g., the surface normals 58). In some embodiments, the facial reconstruction neural network 1104 is trained to compute corrected surface normals ($Corrected\_N_{Pol_{C1}}$ and $Corrected\_N_{Pol_{C2}}$ as described above with respect to FIGS. 8B and 9B, for example) based on the estimated surface normals 58, and to generate the 3D reconstruction 1106 of the face 1102 based on the corrected surface normals.

[00243] According to some embodiments, the facial reconstruction neural network 1104, which may be a polarized CNN, is trained to correlate a wide variety of feature maps (e.g., surface normals) with detailed 3D reconstructions of faces from which the feature maps were generated.

[00244] Accordingly, aspects of some embodiments of the present disclosure are related to a 3D imaging system capable of generating a plurality of enhanced/detailed 3D facial reconstructions based on faces captured by one or more polarization camera module 10′.

[00245] FIG. 11B illustrates the 3D imaging system 1110, according to some embodiments of the present disclosure.

[00246] In some embodiments, the 3D imaging system 1110 includes one or more polarization camera modules 10′ for capturing one or more polarization raw frames/images 18, and a processing system 100 for generating a 3D reconstruction 1106′ of the observed face 1102′ based on the one or more polarization raw images 18.

[00247] In some embodiments, the one or more polarization camera modules 10′ include a least first and second polarization camera modules 10-1′ and 10-2′ (such as the stereo polarization camera system 10″), which capture polarization raw frames 18 from at least two viewpoints, and the processing system 100 generates its own coarse depth map based on the polarization raw frames 18. In some embodiments, the processing system 100 computes an initial estimate of DOLP ρ and AOLP ϕ for each of the at least two viewpoints, computes estimated surface normals from the initial estimate of DOLP ρ and AOLP ϕ for each of the at least two view points, and estimates the face geometry based on the polarization raw frames 18 (as in, e.g., operations 410 and 430 of FIG. 4). Then for each view point, the processing system 100 may proceed to refine this initial coarse face model from the perspective of that view point and the corresponding polarization cues to arrive at refined/corrected polarization cues (e.g., in accordance with the operations 450, 470, and 490 of FIG. 4). The processing system 100 then computes refined/corrected surface normals from the polarization cues for each of the at least two view points. In some embodiments, the processing system 100 independently integrates the surface normals from each view point over sample space (Ω) to synthesize a 3D reconstruction of the face 1102′ from that perspective.

[00248] Here, each pair of estimated surface normals and the corresponding 3D reconstruction associated with a particular view point form one set of training data for the 3D imaging system 1100 of FIG. 11A. By repeating this process for different faces (e.g., a large number of faces), sufficient training data may be collected to train the facial reconstruction neural network 1104 to generate 3D facial reconstructions based on estimated surface normals from a single polarization camera module 10′.

[00249] While the 3D imaging system 1110 may generate the initial coarse depth map of the observed face 1102′ based on triangulation (i.e., the parallax shift between two or more polarization camera modules 10′), embodiments of the present invention are not limited thereto.

[00250] For instance, in some embodiments, the 3D imaging system 1110 further includes a coarse depth map generator 1112 that may generate the initial coarse face model based on a monocular 2D color image (e.g., RGB or RGBG image) of the observed face 1102′. This 2D color image may be captured by one of the polarization camera modules 10′ that is capable of capturing color information in addition to polarization data, or may be captured by a separate conventional color camera (e.g., RGB camera) that has a field of view similar to that of the one or more polarization camera modules 10′. The coarse depth map generator 1112 may utilize an algorithm for depth estimation, or a neural network trained to estimate depth information, based on a 2-dimensional image (e.g., using the inverse square law), and thus obtains an initial coarse depth map or model of the observed face 1102′. In some other embodiments, 3D imaging system 1110 may further include a depth sensor configured to measure depth from focus/defocus, motion, etc., and the coarse depth map generator 1112 may generate the initial coarse face model based on input from the depth sensor . According to some further embodiments, the coarse depth map generator 1112 may provide a model of a generic human face (e.g., an average human face or a 3D morphable model (3DMM) of a face) to the processing system 100 to be used as the initial coarse face model that is then refined as described above. In such embodiments, the processing system 100 may align the model of the generic face provided by the coarse depth map generator 1112 to the image spaces corresponding to the viewpoint(s) of the one or more polarization camera modules 10′, as described above with reference to operation 840 of FIG. 8B.

[00251] *Slope-Net: Using Surface normal information for creation of object signatures*

[00252] *and face signatures.*

[00253] Object and face understanding are important problems to solve. The applications of object recognition and understanding are numerous, such as fashion design, product recognition, sorting, and more. Face emotion detection and face recognition may be important for applications such as driver monitoring, user authentication (e.g., in the areas of automatic payments, security, etc.), general surveillance, and the like. Both of object recognition and facial understanding rely on being able to take pixels from an image, and converting them into a signature, or representation, that doesn't change from image to image. The field of representation learning in machine learning is currently dedicated to the task of predicting complex inputs onto low dimensional manifolds. Generally, the lower the dimension of the data manifold, the easier it is to learn these representations. Standard RGB images, which do not contain surface normal information, may include a great deal of scene-specific information about texture, lighting, etc. The face/object recognition algorithms of the related art learn this manifold through learning an embedding space, i.e., a relatively low-dimensional space into which each image (which is a high-dimensional vector) is translated to a low dimensional vector. This training process may be very slow and prone to divergence.

[00254] In contrast, an image with just the surface normals does not contain any texture or lighting information, which lowers the complexity of the manifold that must be learnt. This may allow for quicker and more accurate learning.

[00255] Some aspects of embodiments of the present disclosure relate to generating surface-normal images of various objects and/or faces and compiling one or more libraries of such surface-normal images for later processing. According to some embodiments, in the surface-normal image, the red, green, and blue values of each of the pixels encode surface normal information at that pixel. In some embodiments, the first, second, and third values of the RGB values for a pixel respectively represent the x-axis component, the y-axis component, and the z-axis component of the surface normal at that pixel. In some examples, the first, second, and third values may be red, green, and blue values, respectively; however, embodiments of the present invention are not limited thereto, and the first, second, and third values may be blue, red, and green values, respectively, or any other ordering of these values. To achieve consistency across the library of surface-normal images and to enable ease of use in post-processing, the x-,y-, and z-axis components of the surface normals are mapped to the same ones of the RGB values in all surface-normal images of the library (this may be referred to as mapping consistency).

[00256] Each surface-normal image within the library may be associated with a label or tag that identifies and/or characterizes the object captured by the image. The surface-normal images within a library may all be associated with a particular class of objects (e.g., vehicles, bikes, animals, traffic lights, pedestrians, faces, etc.).

For example, a library may contain surface-normal image of various types of vehicles, such as trucks, sport utility vehicles (SUVs), vans, sedans, etc, and each surface-normal image within this library may be labeled to identify the vehicle type. In another example, a library may include surface-normal image of various human facial expressions, each being labeled with the associated captured expression (e.g., laugh, smile, frown, surprise, disgust, etc.).

[00257] The labeling of the surface-normal images may be done manually (e.g., labeled by a human) or may be performed by a machine (e.g., a computer) when synthesized images are used to generate the surface-normal images.

[00258] Each surface-normal image may be generated using the techniques described above. For example, an object may be imaged via a polarization camera module 10′ or the stereo polarization camera system 10″ and the feature extractor 700 may generate the corresponding surface-normal image, as described above with respect to FIGS. 6A, 7A and 7B. In some embodiments, the surface normal for each pixel are encoded using Cartesian coordinates (e.g., three-dimensional vectors indicating a direction of the surface normal) into the RGB values.

[00259] According to some embodiments, the one or more libraries are used to train different machine learning algorithms (e.g., predictor/classifiers). In some examples, a library of surface-normal images corresponding to human faces may be used to train a classifier (e.g., facial expression classifier) to identify different emotional states. In so doing, a classifier, such as the predictor 710, is trained to correlate the facial normal images with the corresponding labels, which may indicate the emotional state captured by the image. This classifier may then be used to identify emotional states of a captured face based solely on the surface normals of an observed face. Similarly, the classifier may be trained to recognize intent, gaze, physical wellness, gait, and/or the like based on normal images of human faces or human bodies.

[00260] In other examples, the classifier (e.g., the predictor 710) may be trained to identify different types of vehicles, such as trucks, SUVs, vans, sedans, etc., by using surface-normal images from a library containing labeled normal images of various vehicle types.

[00261] FIG. 12 is a block diagram of a predictor 710, according to one embodiment of the present invention.

[00262] According to some embodiments, the predictor 710 includes a trained classifier that classifies the input surface-normal image 50’ input into one or more categories. For example, the trained classifier may compute a characterization output 20 that includes a vector (e.g., a probability vector) having a length equal to the number of different possible image characteristics (e.g., facial expressions or vehicle types) that the classifier is trained to detect, where each value in the vector corresponds to a confidence that the input surface-normal image 50’ depicts the corresponding image characteristic. In some embodiments, the predictor 710 includes a plurality of statistical models (e.g., 712, 714 ... 716, etc.) associated with different types of image characteristics, and each model provides a confidence level that the input surface-normal image 50’ matches the associated image characteristic. For example, when the image characteristics are facial expressions, each of the models of the predictor 710 may be associated with corresponding one of laugh, smile, frown, surprise, etc, and the predictor 710 may output a vector, where each element in the vector is a confidence/probability value corresponding to one of a laugh, smile, frown, surprise, etc.

[00263] The classifier may be trained to input surface-normal images 50’ of a fixed size and a particular mapping consistency. For example, the first, second, and third values of the RGB values for a pixel of the surface-normal image 50’ respectively represent the x-axis component, the y-axis component, and the z-axis component of the surface normal at that pixel. The classifier may include, for example, a support vector machine, a deep neural network (e.g., a deep fully connected neural network), and the like.

[00264] In embodiments in which the classifier utilizes a neural network such as a convolutional neural network (e.g., a Polarization Mask R-CNN), the training process may include updating the weights of connections between neurons of various layers of the neural network in accordance with a backpropagation algorithm and the use of gradient descent to iteratively adjust the weights to minimize an error (or loss) between the output of the neural network and the labeled training data.

[00265] The operations performed by the constituent components of the surface profilometry system and the various 3D imaging systems of the present disclosure may be performed by a “processing circuit” or “processor” that may include any combination of hardware, firmware, and software, employed to process data or digital signals. Processing circuit hardware may include, for example, application specific integrated circuits (ASICs), general purpose or special purpose central processing units (CPUs), digital signal processors (DSPs), graphics processing units (GPUs), and programmable logic devices such as field programmable gate arrays (FPGAs). In a processing circuit, as used herein, each function is performed either by hardware configured, i.e., hard-wired, to perform that function, or by more general-purpose hardware, such as a CPU, configured to execute instructions stored in a non-transitory storage medium. A processing circuit may be fabricated on a single printed wiring board (PWB) or distributed over several interconnected PWBs. A processing circuit may contain other processing circuits; for example, a processing circuit may include two processing circuits, an FPGA and a CPU, interconnected on a PWB.

[00266] It will be understood that, although the terms "first", "second", "third", etc., may be used herein to describe various elements, components, regions, layers, and/or sections, these elements, components, regions, layers, and/or sections should not be limited by these terms. These terms are used to distinguish one element, component, region, layer, or section from another element, component, region, layer, or section. Thus, a first element, component, region, layer, or section discussed below could be termed a second element, component, region, layer, or section, without departing from the scope of the inventive concept.

[00267] The terminology used herein is for the purpose of describing particular embodiments and is not intended to be limiting of the inventive concept. As used herein, the singular forms "a" and "an" are intended to include the plural forms as well, unless the context clearly indicates otherwise. It will be further understood that the terms "include", "including", "comprises", and/or "comprising", when used in this specification, specify the presence of stated features, integers, steps, operations, elements, and/or components, but do not preclude the presence or addition of one or more other features, integers, steps, operations, elements, components, and/or groups thereof. As used herein, the term "and/or" includes any and all combinations of one or more of the associated listed items. Further, the use of "may" when describing embodiments of the inventive concept refers to "one or more embodiments of the inventive concept". Also, the term "exemplary" is intended to refer to an example or illustration.

[00268] As used herein, the terms "use", "using", and "used" may be considered synonymous with the terms "utilize", "utilizing", and "utilized", respectively.

[00269] Further, the use of “may” when describing embodiments of the inventive concept refers to “one or more embodiments of the inventive concept.” Also, the term “exemplary” is intended to refer to an example or illustration.

[00270] 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 method of performing surface profilometry, the method comprising:

receiving one or more polarization raw frames of a printed layer of a physical object undergoing additive manufacturing, the one or more polarization raw frames being captured at different polarizations by one or more polarization cameras;

extracting one or more polarization feature maps in one or more polarization representation spaces from the one or more polarization raw frames;

obtaining a coarse layer depth map of the printed layer;

generating one or more surface-normal images based on the coarse layer depth map and the one or more polarization feature maps; and

generating a 3D reconstruction of the printed layer based on the one or more surface-normal images.

2. The method of claim 1, wherein the one or more polarization feature maps in the one or more polarization representation spaces comprise:

a degree of linear polarization (DOLP) image in a DOLP representation space; and

an angle of linear polarization (AOLP) image in an AOLP representation space.

3. The method of claim 1, wherein the one or more polarization cameras comprise:

a first polarization camera configured to capture a first partition of a print bed on which the printed layer of the physical object resides; and

a second polarization camera at a distance from the first polarization camera and configured to capture a second partition of the print bed.

4. The method of claim 3, wherein the obtaining the coarse layer depth map of the printed layer comprises:

constructing the coarse layer depth map based on a parallax shift between polarization raw frames captured by the first and second polarization cameras.

5. The method of claim 1, wherein the obtaining the coarse layer depth map of the printed layer comprises:

receiving a computer-aided-design (CAD) layer model corresponding to the printed layer of the physical object.

6. The method of claim 1, wherein the generating the one or more surface-normal images comprises:

determining a pose of the printed layer with respect to each of the one or more polarization cameras;

transforming the coarse layer depth map into one or more camera spaces corresponding to the one or more polarization cameras to generate one or more transformed coarse layer depth maps; and

correcting the one or more surface-normal images based on the one or more transformed coarse layer depth maps.

7. The method of claim 1, wherein the generating the 3D reconstruction of the printed layer comprises:

integrating surface normals of the one or more surface-normal images over a sample space to determine a shape of a surface of the printed layer.

8. A surface profilometry system comprising:

one or more polarization cameras comprising a polarizing filter, the one or more polarization cameras being configured to capture polarization raw frames at different polarizations; and

a processing system comprising a processor and memory storing instructions that, when executed by the processor, cause the processor to perform:

receiving one or more polarization raw frames of a printed layer of a physical object undergoing additive manufacturing, the one or more polarization raw frames being captured at different polarizations by the one or more polarization cameras;

extracting one or more polarization feature maps in one or more polarization representation spaces from the one or more polarization raw frames;

obtaining a coarse layer depth map of the printed layer;

generating one or more surface-normal images based on the coarse layer depth map and the one or more polarization feature maps; and

generating a 3D reconstruction of the printed layer based on the one or more surface-normal images.

9. The surface profilometry system of claim 8, wherein the one or more polarization feature maps in the one or more polarization representation spaces comprise:

a degree of linear polarization (DOLP) image in a DOLP representation space; and an angle of linear polarization (AOLP) image in an AOLP representation space.

10. The surface profilometry system of claim 8, wherein the one or more polarization cameras comprise:

a first polarization camera configured to capture a first partition of a print bed on which the printed layer of the physical object resides; and

a second polarization camera at a distance from the first polarization camera and configured to capture a second partition of the print bed.

11. The surface profilometry system of claim 9, wherein the obtaining the coarse layer depth map of the printed layer comprises:

constructing the coarse layer depth map based on a parallax shift between polarization raw frames captured by the first and second polarization cameras.

12. The surface profilometry system of claim 8, wherein the obtaining the coarse layer depth map of the printed layer comprises:

receiving a computer-aided-design (CAD) layer model corresponding to the printed layer of the physical object.

13. The surface profilometry system of claim 8, wherein the generating the one or more surface-normal images comprises:

determining a pose of the printed layer with respect to each of the one or more polarization cameras;

transforming the coarse layer depth map into one or more camera spaces corresponding to the one or more polarization cameras to generate one or more transformed coarse layer depth maps; and

correcting the one or more surface-normal images based on the one or more transformed coarse layer depth maps.

14. The surface profilometry system of claim 8, wherein the generating the 3D reconstruction of the printed layer comprises:

integrating surface normals of the one or more surface-normal images over a sample space to determine a shape of a surface of the printed layer.

15. The surface profilometry system of claim 8, wherein the memory further stores instructions that, when executed by the processor, cause the processor to further perform:

providing the 3D reconstruction of the printed layer to a 3D printing system as a control feedback,

wherein the 3D printing system being configured to additively manufacture the physical object layer by layer.

16. The surface profilometry system of claim 15, wherein operations of the one or more polarization cameras, the processing system, and the 3D printing system are synchronized via a synchronization signal.

17. A method of capturing a 3D image of a face, the method comprising:

receiving one or more polarization raw frames of the face, the one or more polarization raw frames being captured at different polarizations by a polarization camera at a distance from the face;

extracting one or more polarization feature maps in one or more polarization representation spaces from the one or more polarization raw frames; and

generating a 3D reconstruction of the face based on the one or more polarization feature maps via a facial reconstruction neural network.

18. The method of claim 17, wherein the one or more polarization feature maps in the one or more polarization representation spaces comprise:

a degree of linear polarization (DOLP) image in a DOLP representation space; and

an angle of linear polarization (AOLP) image in an AOLP representation space.

19. The method of claim 17, wherein the one or more polarization feature maps comprise estimated surface normals, and

wherein the extracting the one or more polarization feature maps comprises:

generating the estimated surface normals based on the one or more polarization raw frames.

20. The method of claim 19, wherein the facial reconstruction neural network is trained to compute corrected surface normals based on the estimated surface normals, and to generate the 3D reconstruction of the face based on the corrected surface normals.

21. The method of claim 17, wherein the facial reconstruction neural network comprises a trained polarized convolutional neural network (CNN).

22. A 3D imaging system for capturing a 3D image of a face, the 3D imaging system comprising:

a polarization camera comprising a polarizing filter and configured to capture one or more polarization raw frames at different polarizations; and

a processing system comprising a processor and memory storing instructions that, when executed by the processor, cause the processor to perform:

receiving one or more polarization raw frames of the face, the one or more polarization raw frames being captured at different polarizations by the polarization camera;

extracting one or more polarization feature maps in one or more polarization representation spaces from the one or more polarization raw frames; and

generating a 3D reconstruction of the face based on the one or more polarization feature maps via a facial reconstruction neural network.

23. The 3D imaging system of claim 22, wherein the one or more polarization feature maps in the one or more polarization representation spaces comprise:

a degree of linear polarization (DOLP) image in a DOLP representation space; and

an angle of linear polarization (AOLP) image in an AOLP representation space.

24. The 3D imaging system of claim 22, wherein the one or more polarization feature maps comprise estimated surface normals, and

wherein the extracting the one or more polarization feature maps comprises:

generating the estimated surface normals based on the one or more polarization raw frames.

25. The 3D imaging system of claim 8, wherein the facial reconstruction neural network is trained to compute corrected surface normals based on the estimated surface normals, and to generate the 3D reconstruction of the face based on the corrected surface normals.

26. The 3D imaging system of claim 22, wherein the facial reconstruction neural network comprises a trained polarized convolutional neural network (CNN).

27. A method of capturing a 3D image of a face, the method comprising:

receiving one or more polarization raw frames of the face, the one or more polarization raw frames being captured at different polarizations by a polarization camera;

extracting estimated polarization cues from the one or more polarization raw frames;

generating estimated surface normals based on the estimated polarization cues

generating an initial coarse depth map of the face;

refining the estimated polarization cues and the initial coarse depth map to generate refined polarization cues and a refined depth map;

generating corrected surface normals based on the refined polarization cues and the refined depth map; and

generating a 3D reconstruction of the face based on the corrected surface normals.

28. The method of claim 27, wherein the generating the initial coarse depth map of the face comprises:

receiving a 2D color image of the face; and

computing the initial coarse depth map based on the 2D color image of the face.

29. The method of claim 27, wherein the generating the initial coarse depth map of the face comprises:

receiving a 3D model of a generic human face; and

generating the initial coarse depth map based on the 3D model of the generic human face.

30. The method of claim 27, further comprising:

providing the estimated surface normals and the corrected surface normals to a facial reconstruction neural network as a set of training data.

31. A method of computing a prediction, the method comprising:

receiving a surface-normal image corresponding to an object, the surface-normal image comprising surface-normal information for each pixel of the image; and

computing the prediction based on the surface-normal image.

32. The method of claim 31, wherein a red color value, a green color value, and blue color values of a pixel of the surface normal image encode an x-axis component, a y-axis component, and a z-axis component of a surface normal of the object at the pixel.

33. The method of claim 32, wherein the red, green, and blue color values of the pixel are respectively the x-axis component, the y-axis component, and the z-axis component of the surface normal of the object at the pixel.

34. The method of claim 31, wherein the prediction is a probability vector,

wherein the computing the prediction comprises supplying the surface-normal image to a trained classifier, and

wherein the trained classifier is configured to identify image characteristics of the surface-normal image and to output the probability vector, each element of the probability vector being a probability value corresponding to one of possible image characteristics.

35. The method of claim 34, wherein the trained classifier comprises a plurality of statistical models corresponding to the possible image characteristics.

36. The method of claim 34, wherein the image characteristics comprise facial expressions or object types.

37. The method of claim 31, wherein the object comprises a vehicle, a face, or a body.

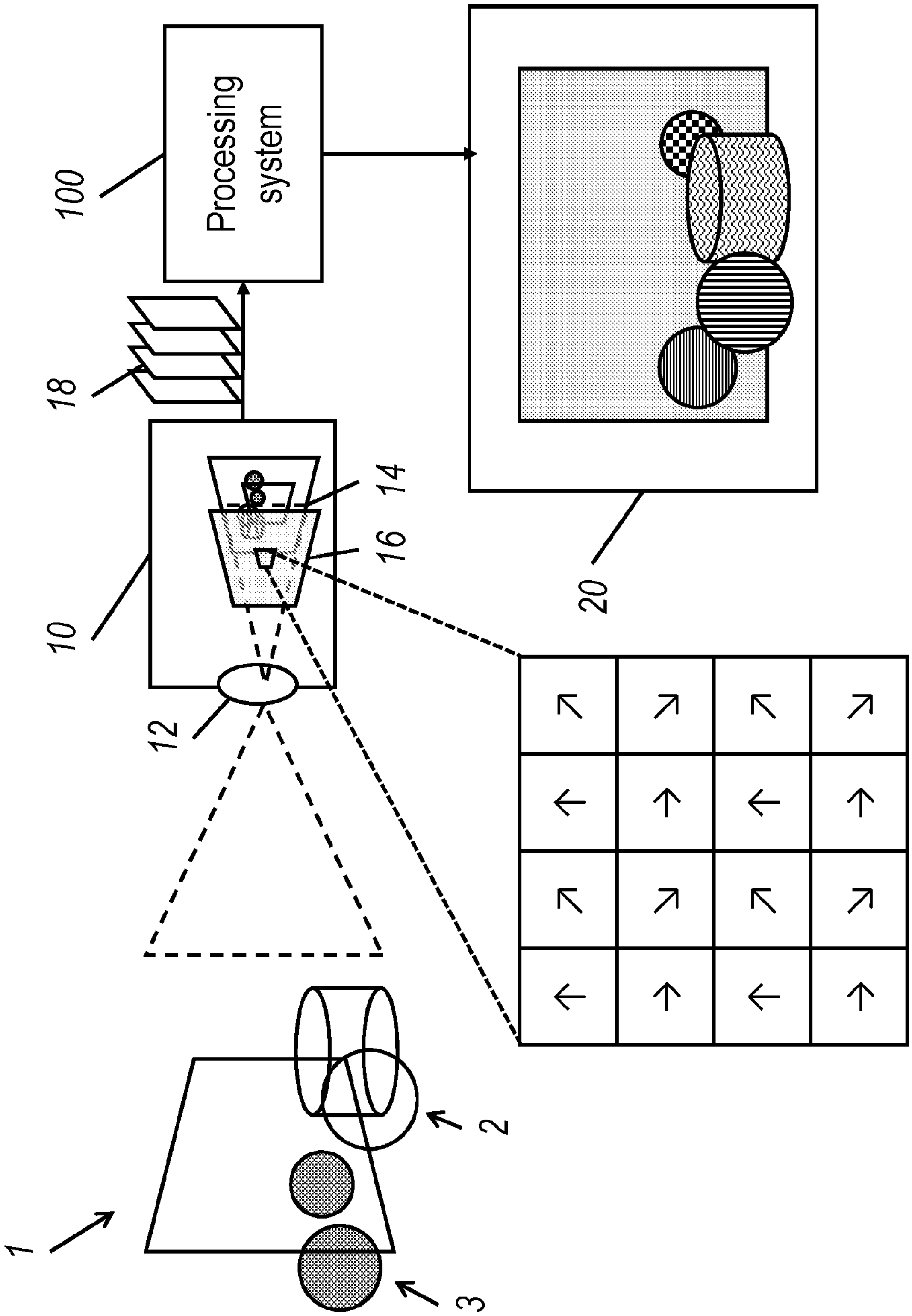

FIG. 1A

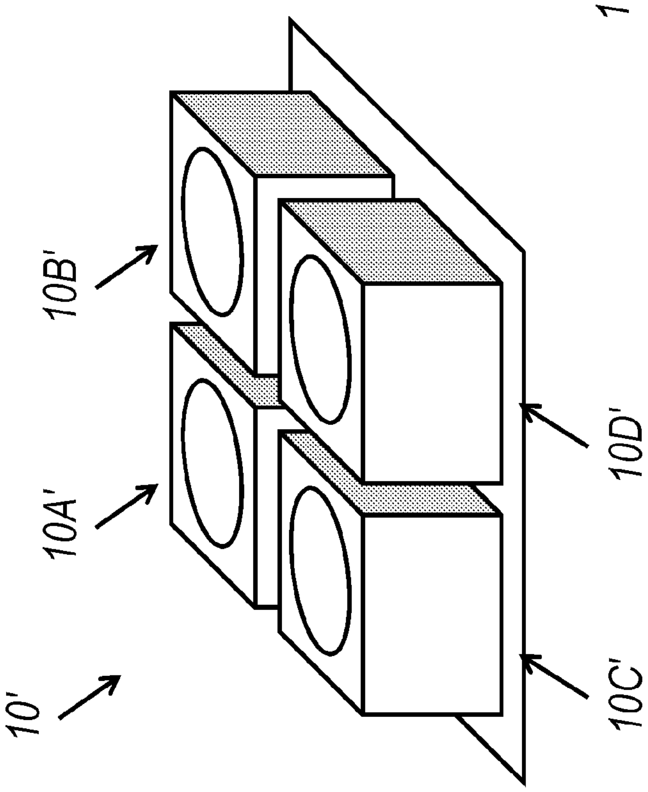

FIG. 1B

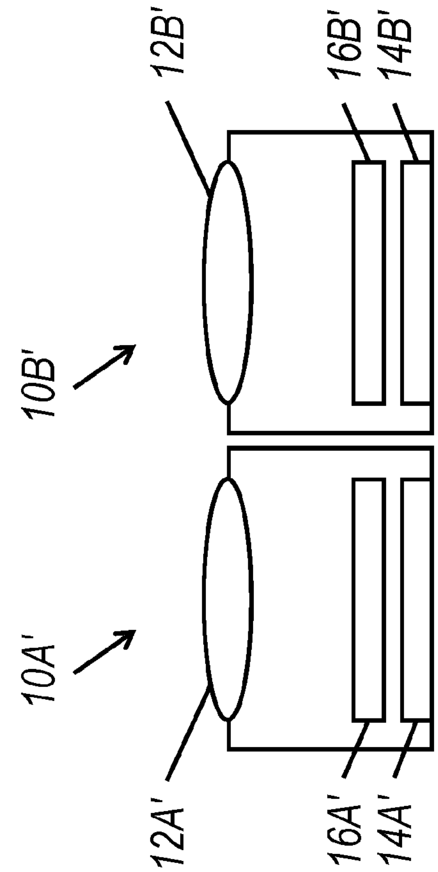

FIG. 1C

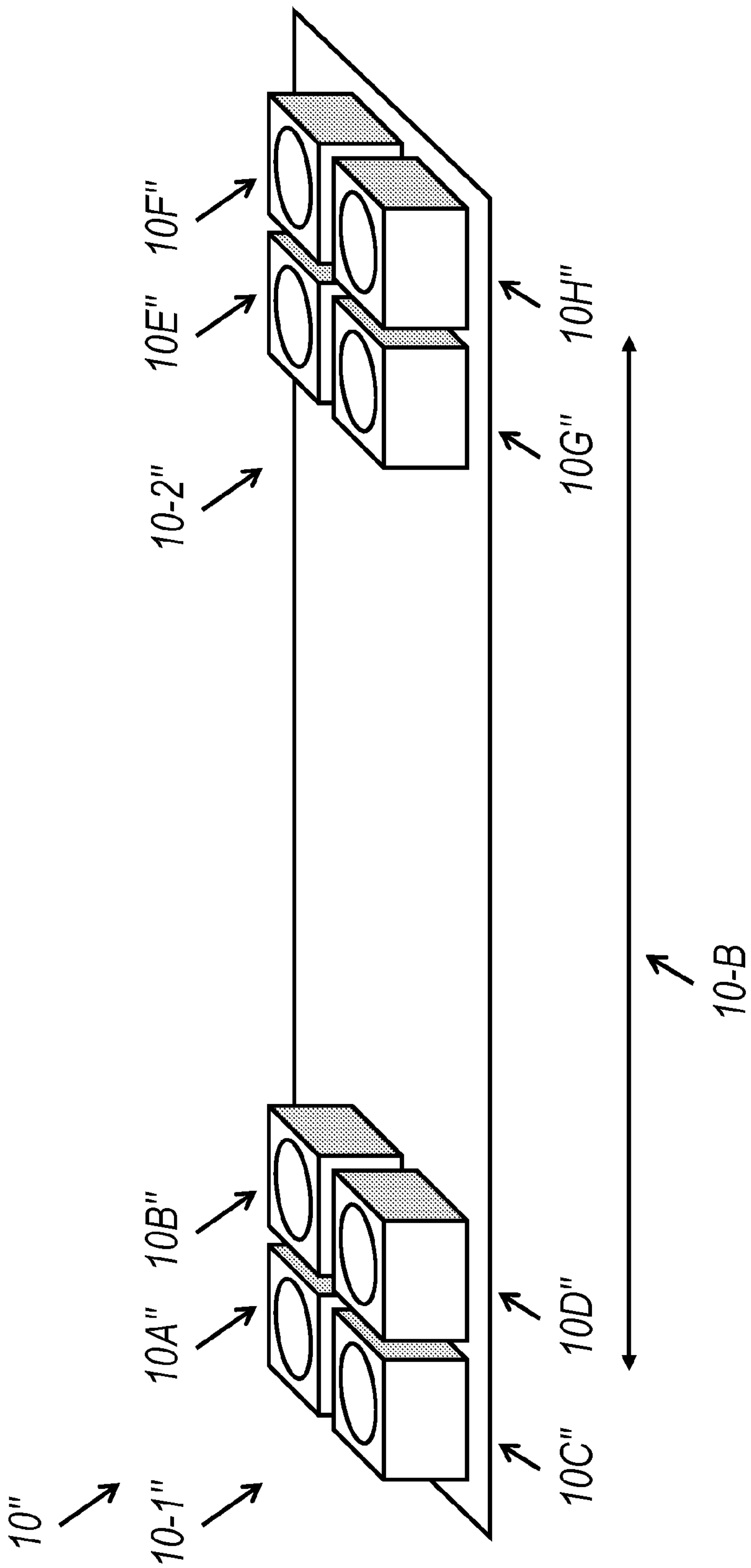

FIG. 1D

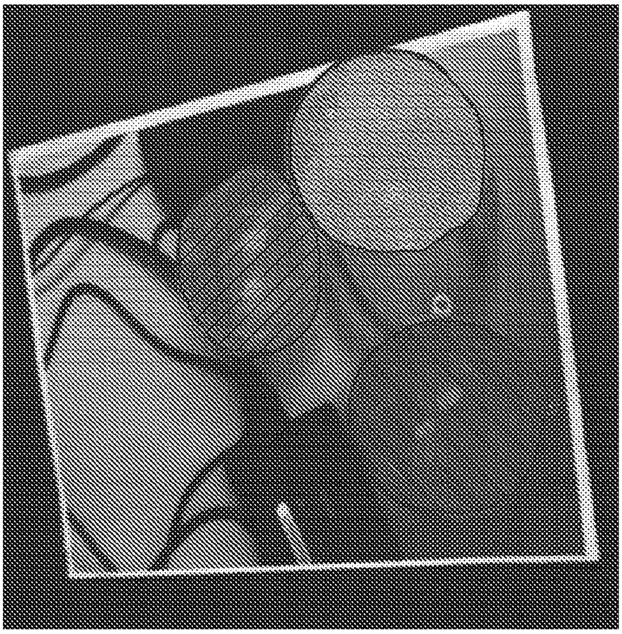

FIG. 2B

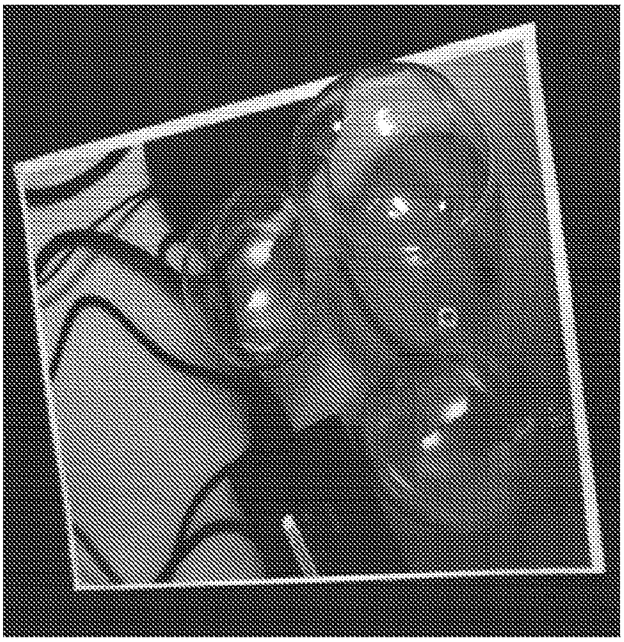

FIG. 2A

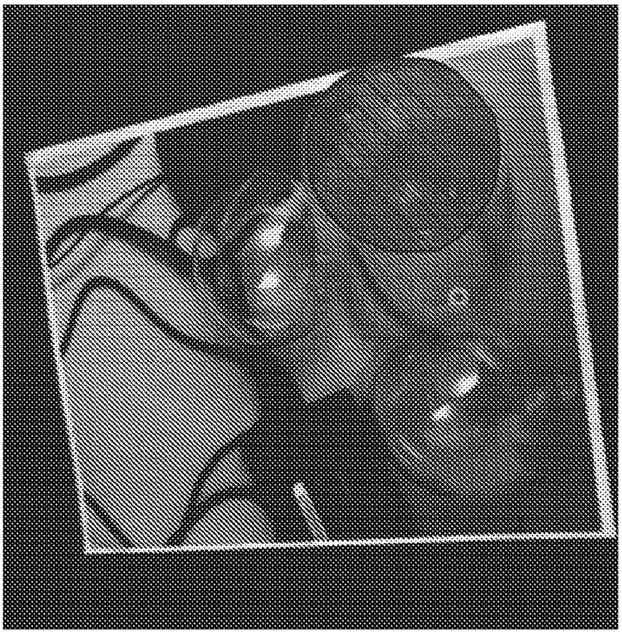

FIG. 2D

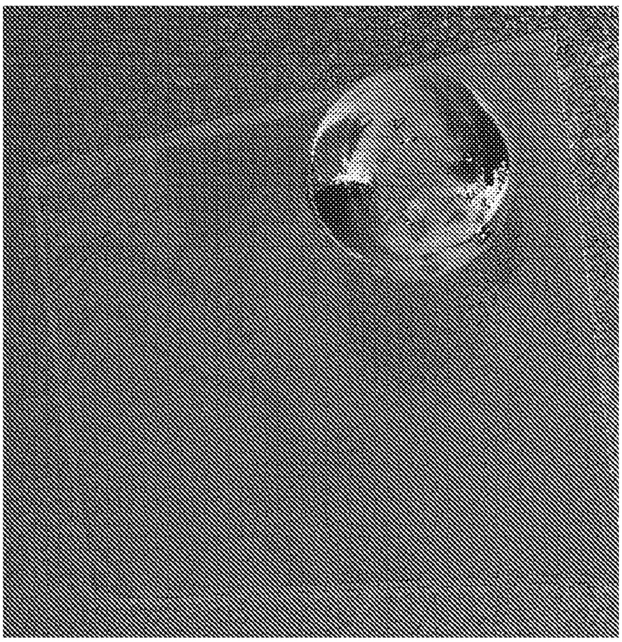

FIG. 2C

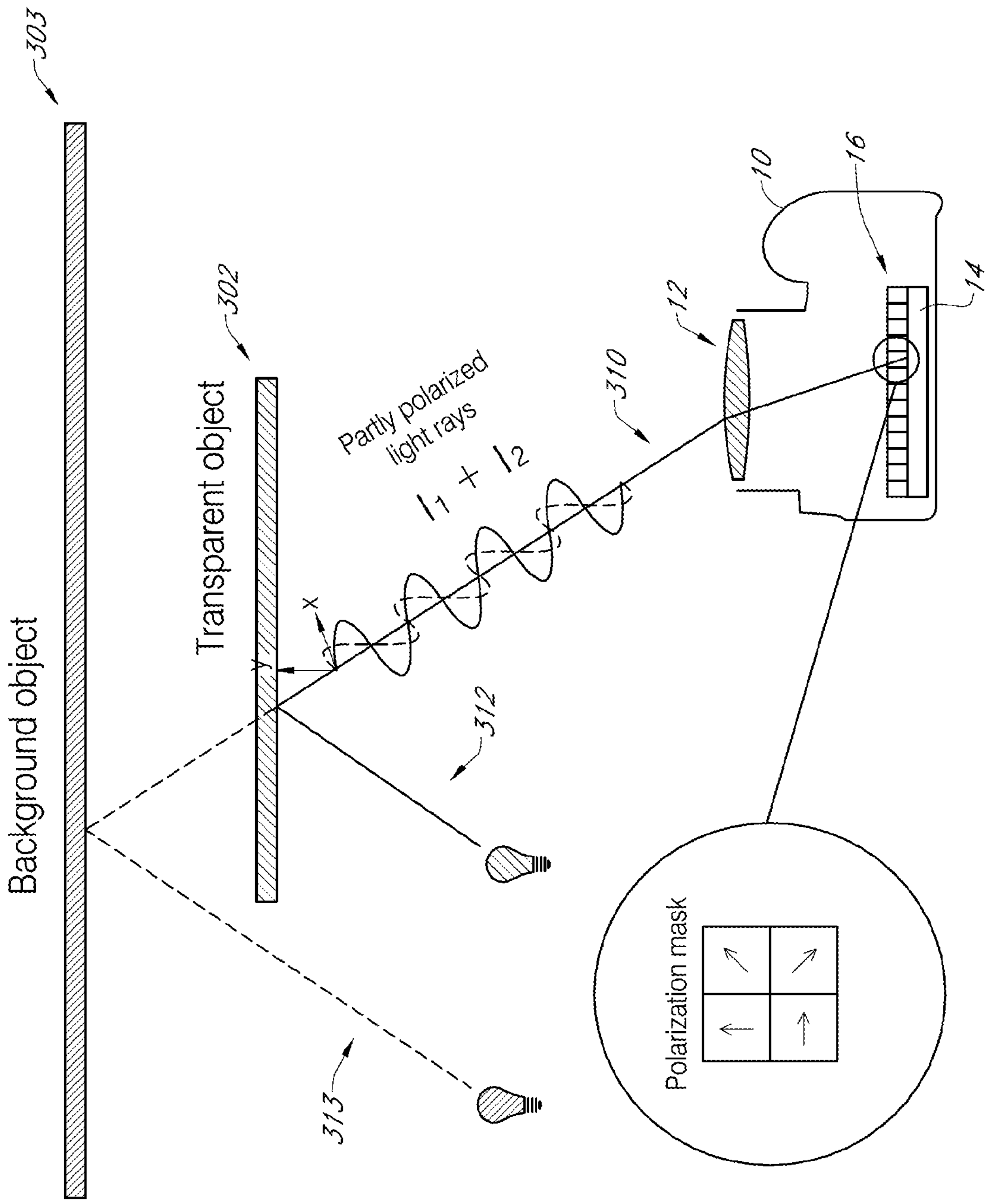
303
Background object
302
Transparent object
Partly polarized light rays
$I_1 + I_2$
x
y
310
312
313
12
10
16
14
Polarization mask

FIG. 3

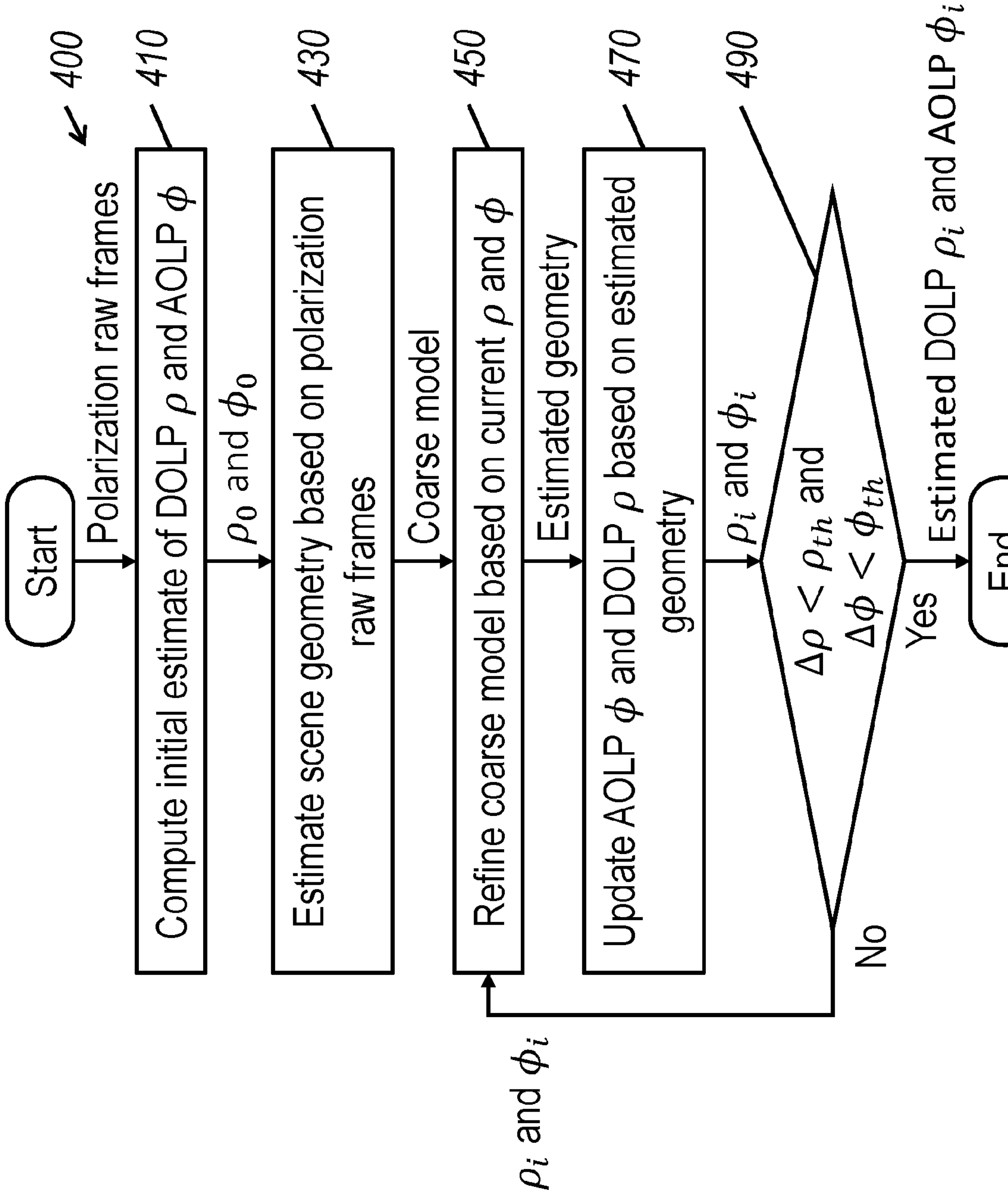
400
Start
Polarization raw frames
410
Compute initial estimate of DOLP ρ and AOLP ϕ
ρ_0 and ϕ_0
430
Estimate scene geometry based on polarization raw frames
Coarse model
450
Refine coarse model based on current ρ and ϕ
Estimated geometry
470
Update AOLP ϕ and DOLP ρ based on estimated geometry
ρ_i and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

FIG. 4

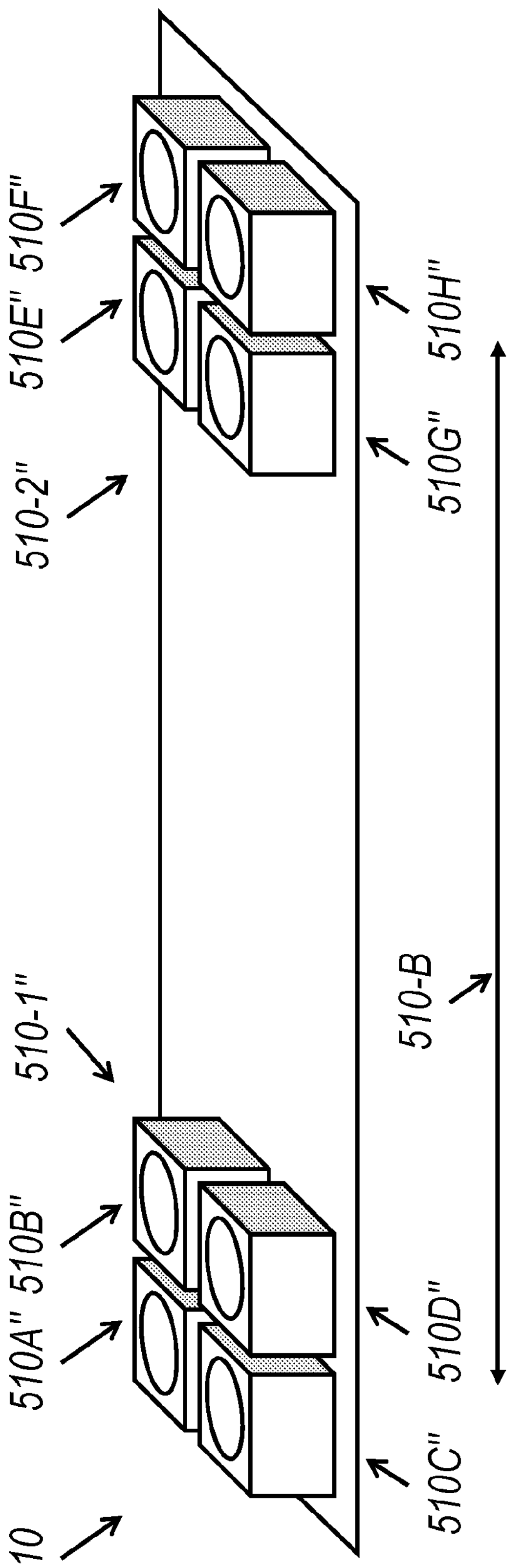
510
510A" 510B"
510C"
510D"
510-1"
510-B
510-2"
510E" 510F"
510G"
510H"

FIG. 5A

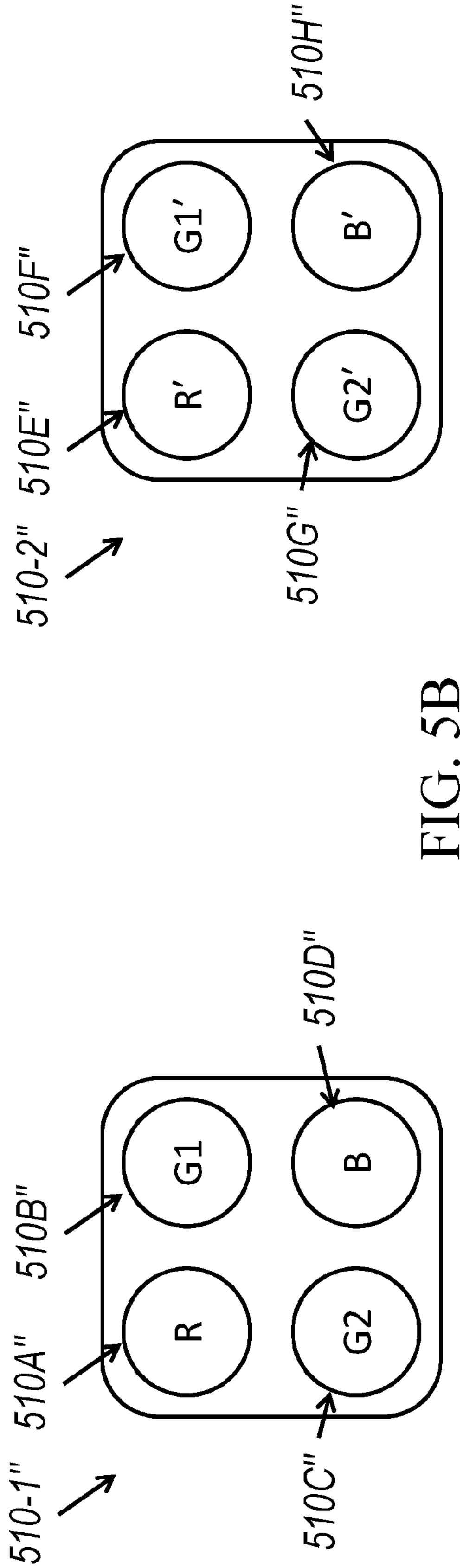
510-1"
510A"
510B"
510C"
510D"
R
G1
G2
B
510-2"
510E"
510F"
510G"
510H"
R′
G1′
G2′
B′

FIG. 5B

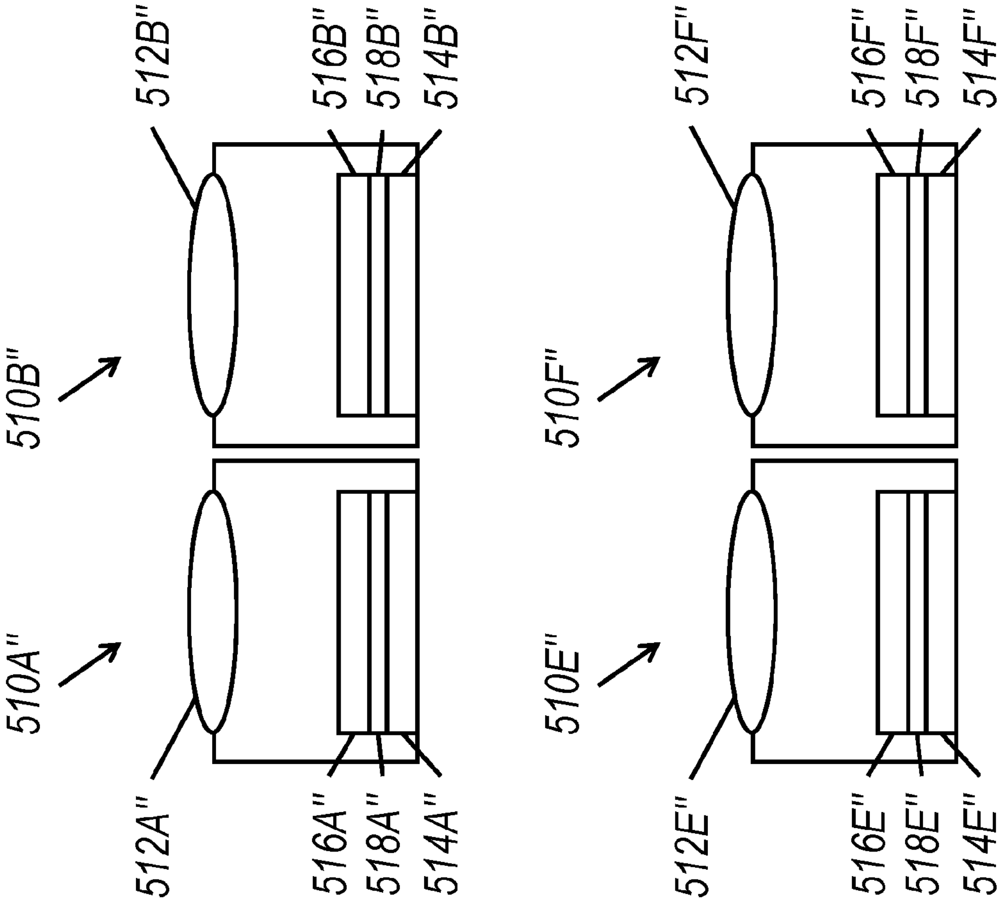
510A''
510B''
512A''
512B''
516A''
518A''
514A''
516B''
518B''
514B''
510E''
510F''
512E''
512F''
516E''
518E''
514E''
516F''
518F''
514F''

FIG. 5C

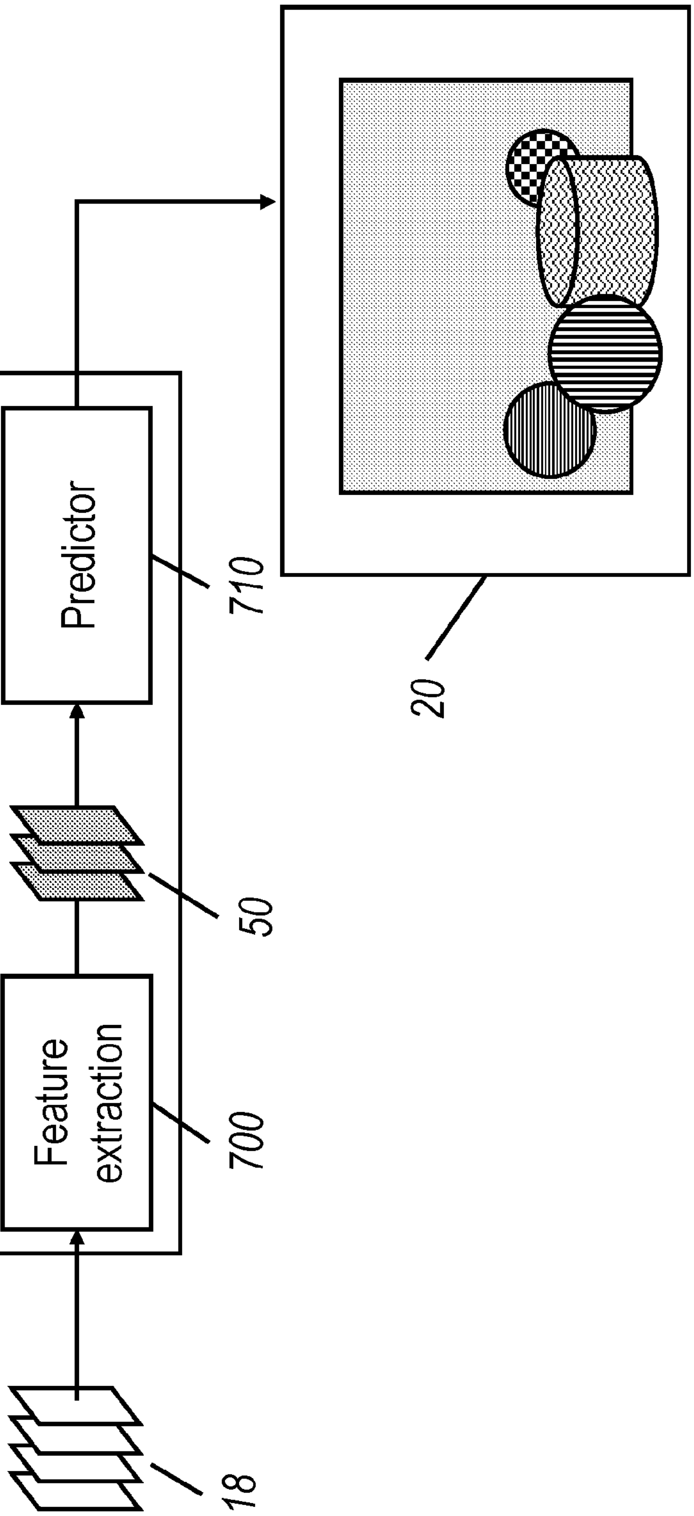

FIG. 6A

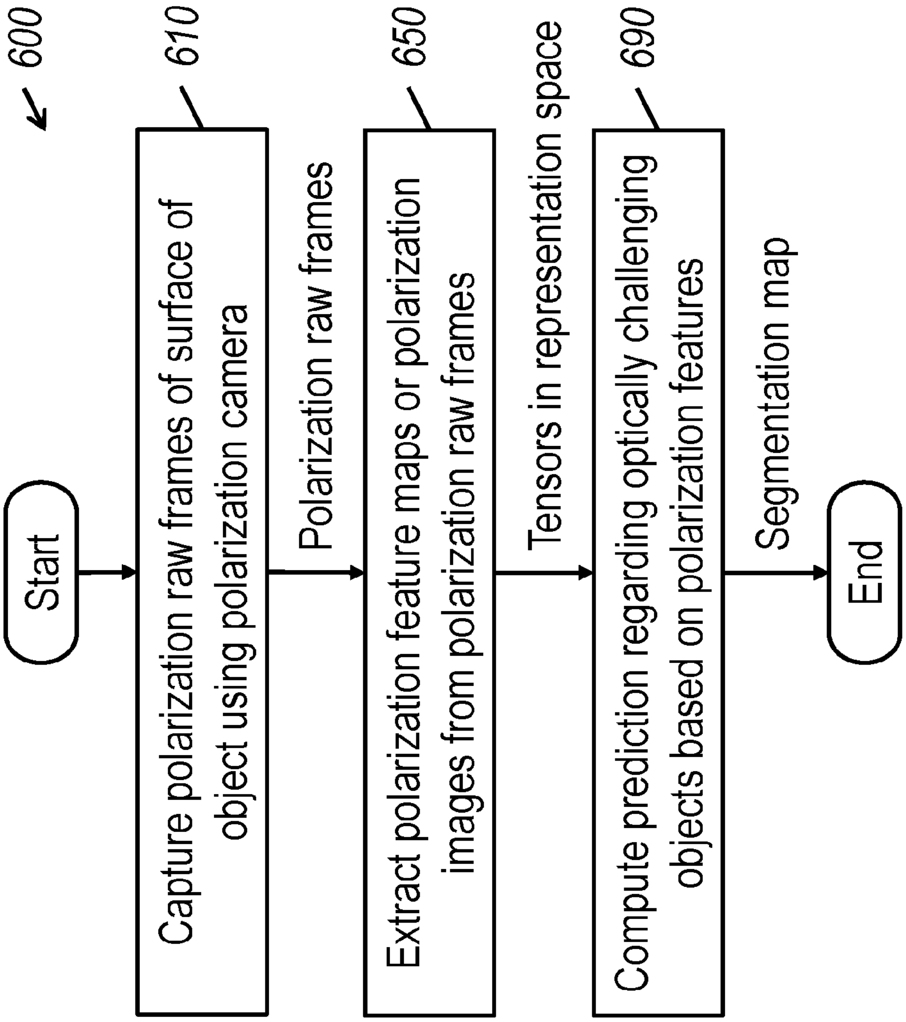
600
Start
610
Capture polarization raw frames of surface of object using polarization camera
Polarization raw frames
650
Extract polarization feature maps or polarization images from polarization raw frames
Tensors in representation space
690
Compute prediction regarding optically challenging objects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

FIG. 6B

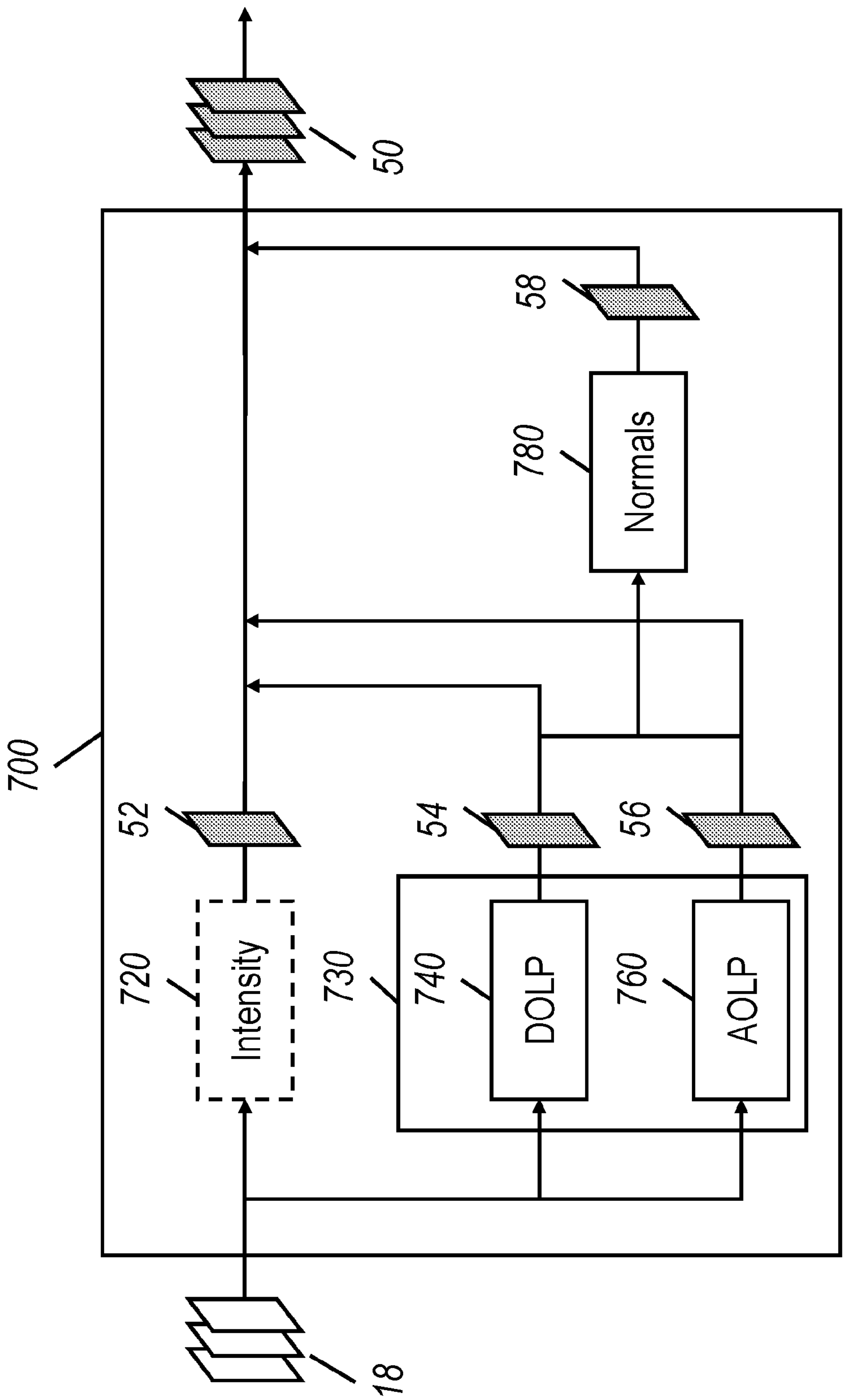
50
58
780
Normals
700
52
54
56
720
Intensity
730
740
DOLP
760
AOLP
18

FIG. 7A

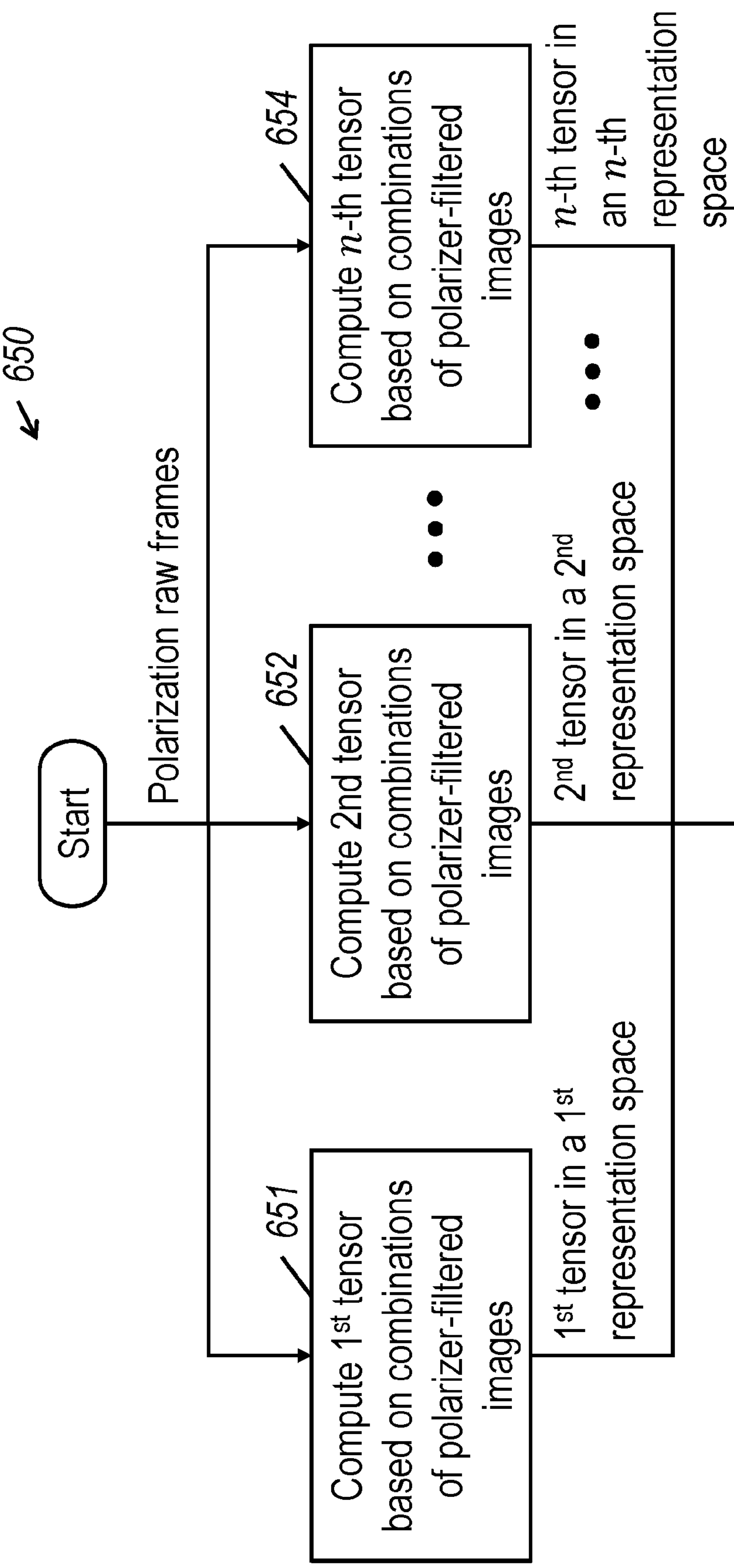
650
Start
Polarization raw frames
651
Compute 1st tensor based on combinations of polarizer-filtered images
652
Compute 2nd tensor based on combinations of polarizer-filtered images
654
Compute n-th tensor based on combinations of polarizer-filtered images
1st tensor in a 1st representation space
2nd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

FIG. 7B

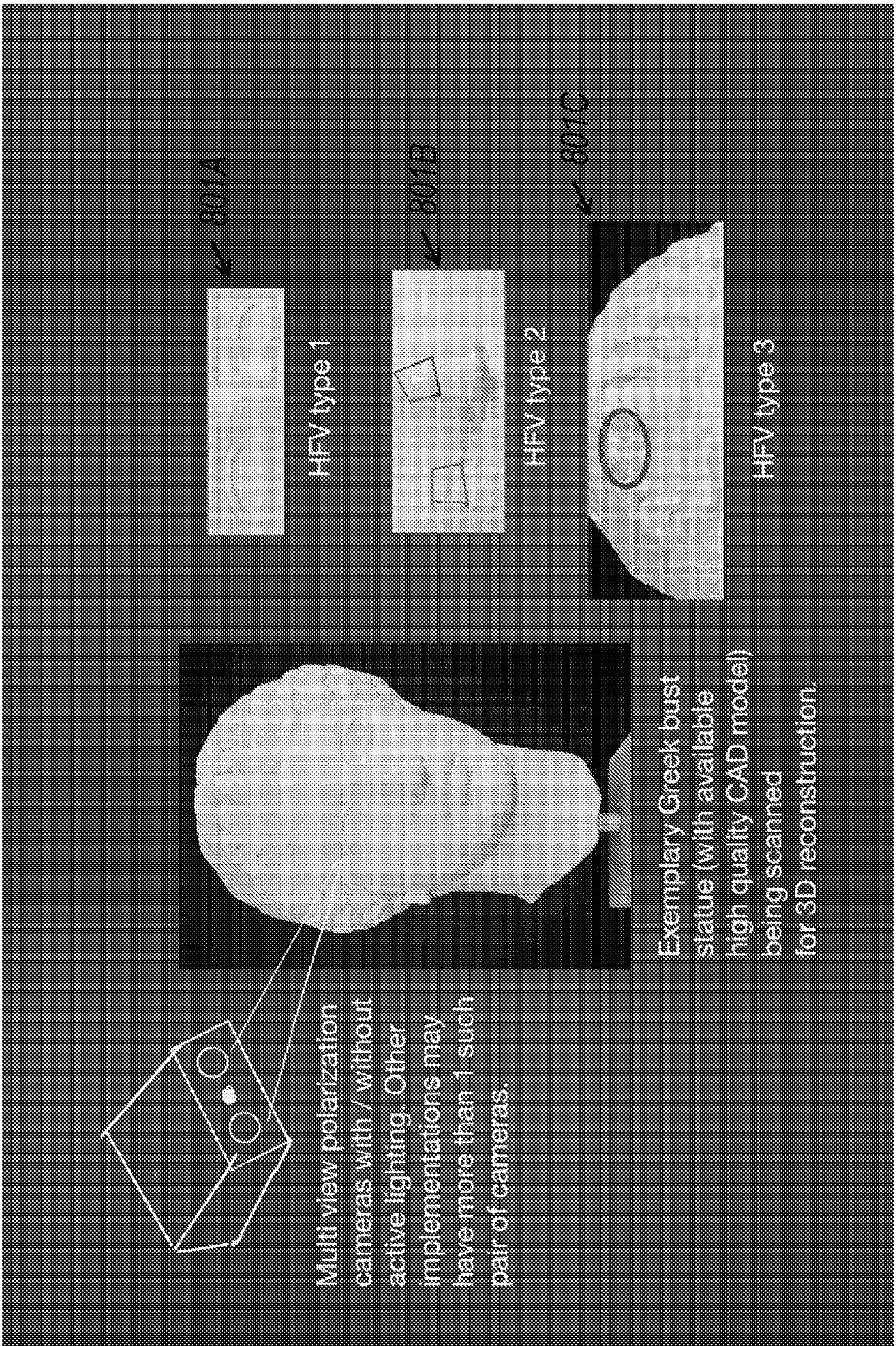
801A
801B
801C
HFV type 1
HFV type 2
HFV type 3
Multi view polarization cameras with / without active lighting. Other implementations may have more than 1 such pair of cameras.
Exemplary Greek bust statue (with available high quality CAD model) being scanned for 3D reconstruction.

FIG. 8A

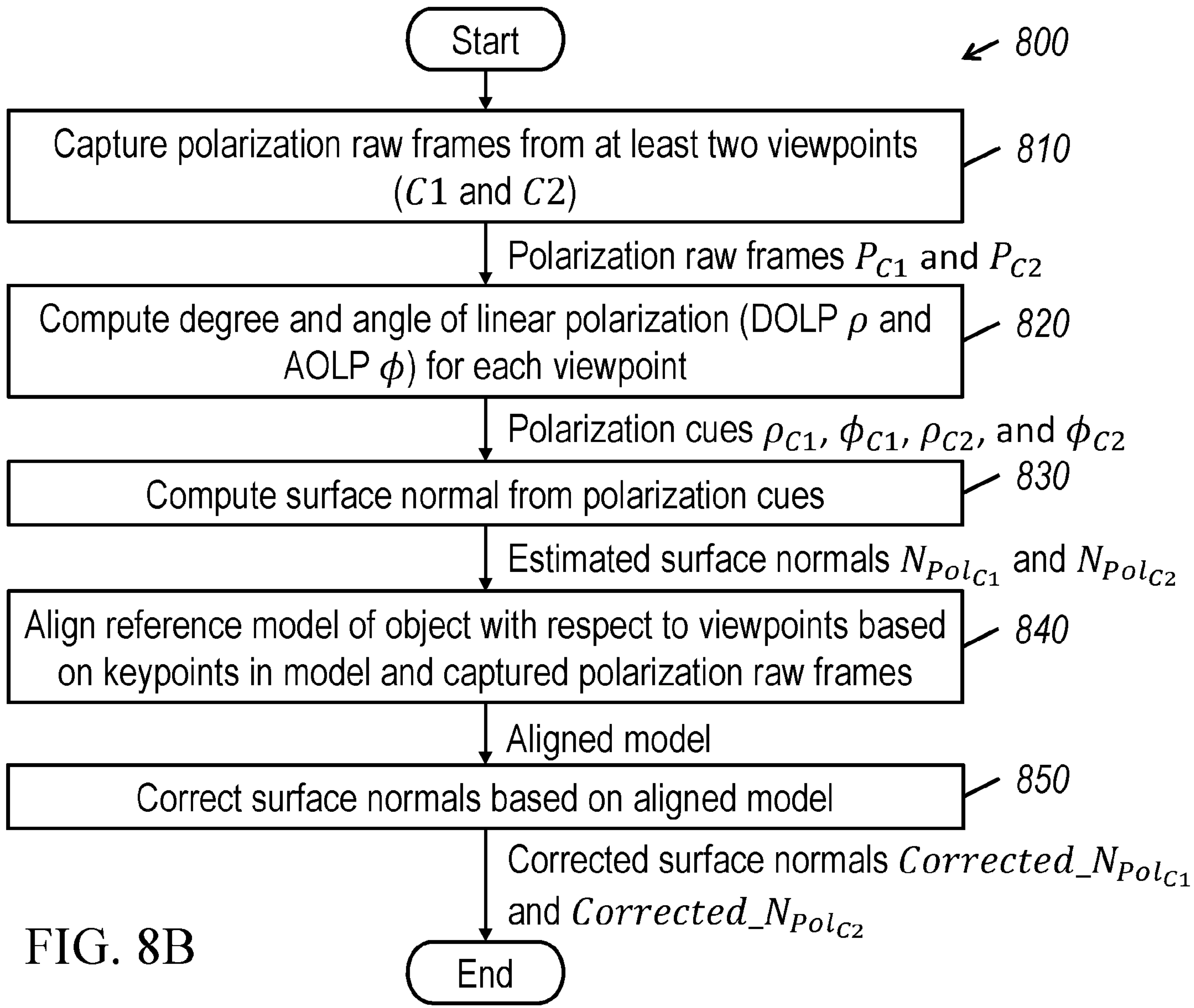
Start
800
Capture polarization raw frames from at least two viewpoints ($C1$ and $C2$)
810
Polarization raw frames P_{C1} and P_{C2}
Compute degree and angle of linear polarization (DOLP ρ and AOLP ϕ) for each viewpoint
820
Polarization cues ρ_{C1}, ϕ_{C1}, ρ_{C2}, and ϕ_{C2}
Compute surface normal from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

FIG. 8B

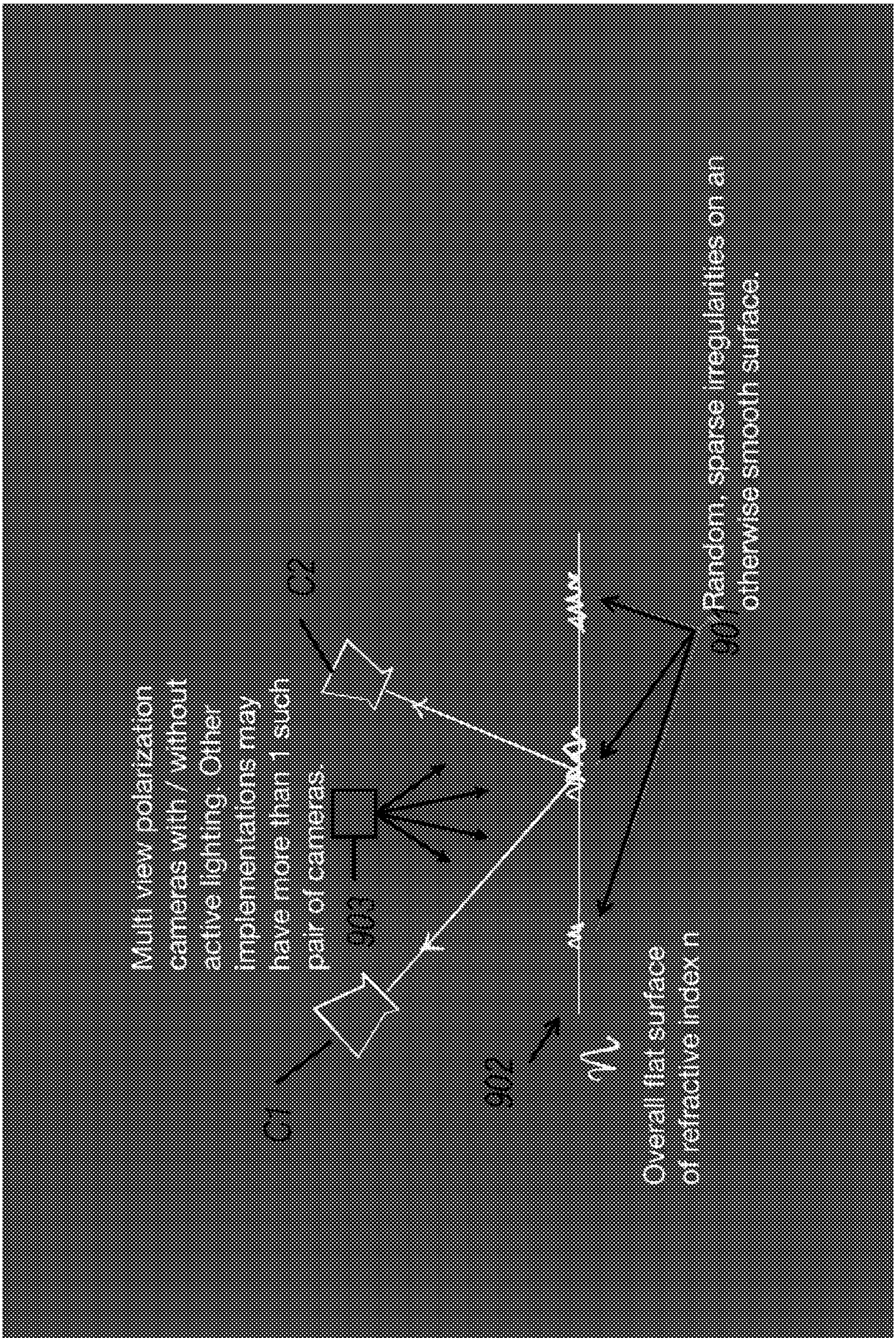
Multi view polarization cameras with / without active lighting. Other implementations may have more than 1 such pair of cameras.
C1
C2
903
902
901
Overall flat surface of refractive index n
Random, sparse irregularities on an otherwise smooth surface.

FIG. 9A

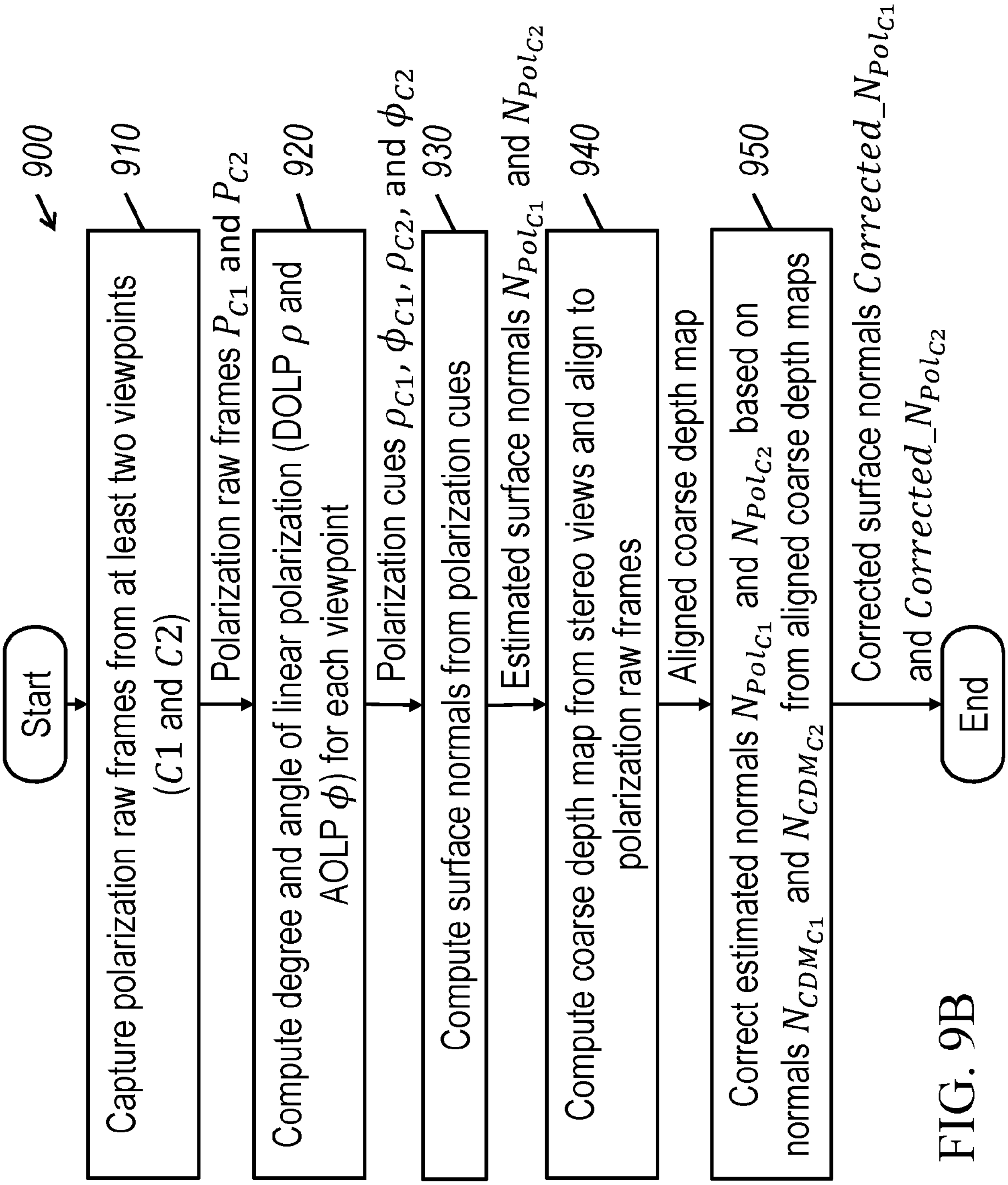
900
Start
910
Capture polarization raw frames from at least two viewpoints ($C1$ and $C2$)
Polarization raw frames P_{C1} and P_{C2}
920
Compute degree and angle of linear polarization (DOLP ρ and AOLP ϕ) for each viewpoint
Polarization cues ρ_{C1}, ϕ_{C1}, ρ_{C2}, and ϕ_{C2}
930
Compute surface normals 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

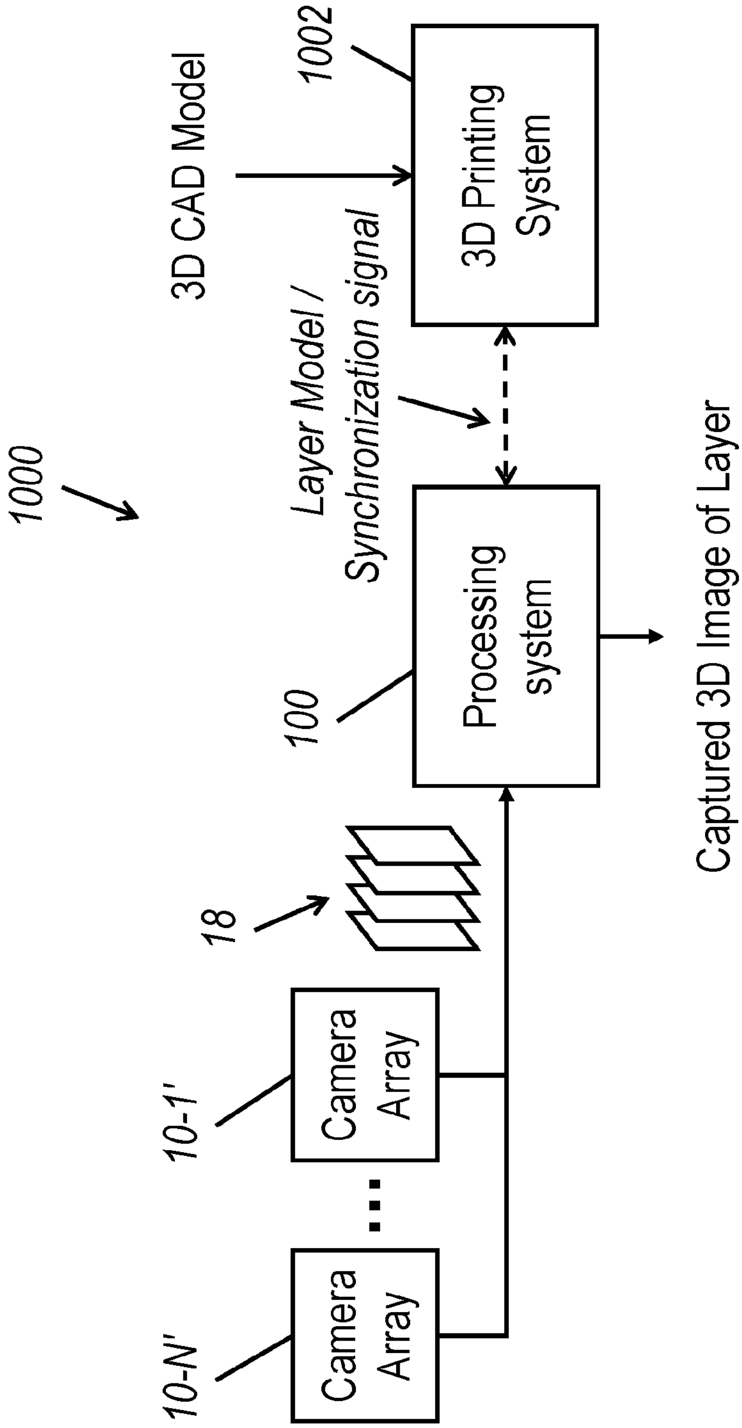
1000
3D CAD Model
1002
3D Printing System
Layer Model / Synchronization signal
Processing system
100
Captured 3D Image of Layer
18
Camera Array
10-1'
Camera Array
10-N'

FIG. 10A

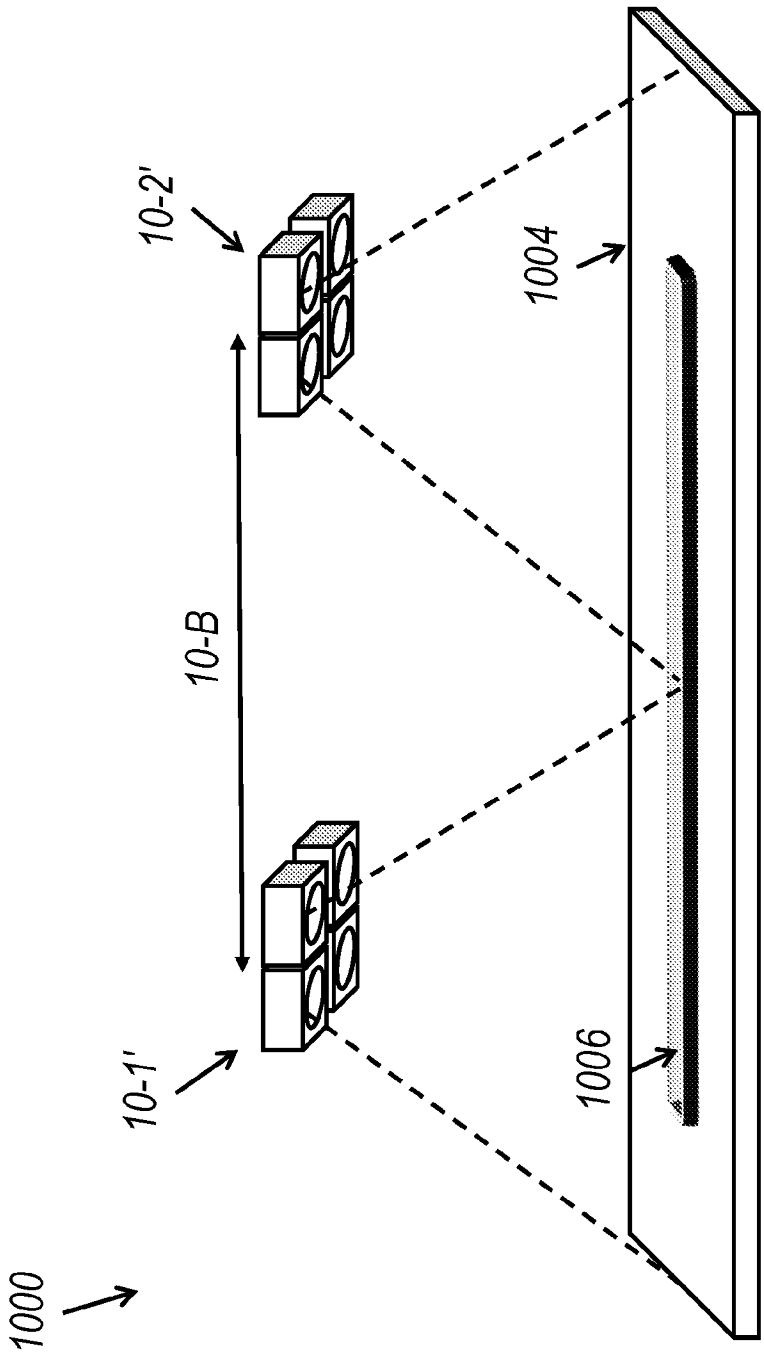

FIG. 10B

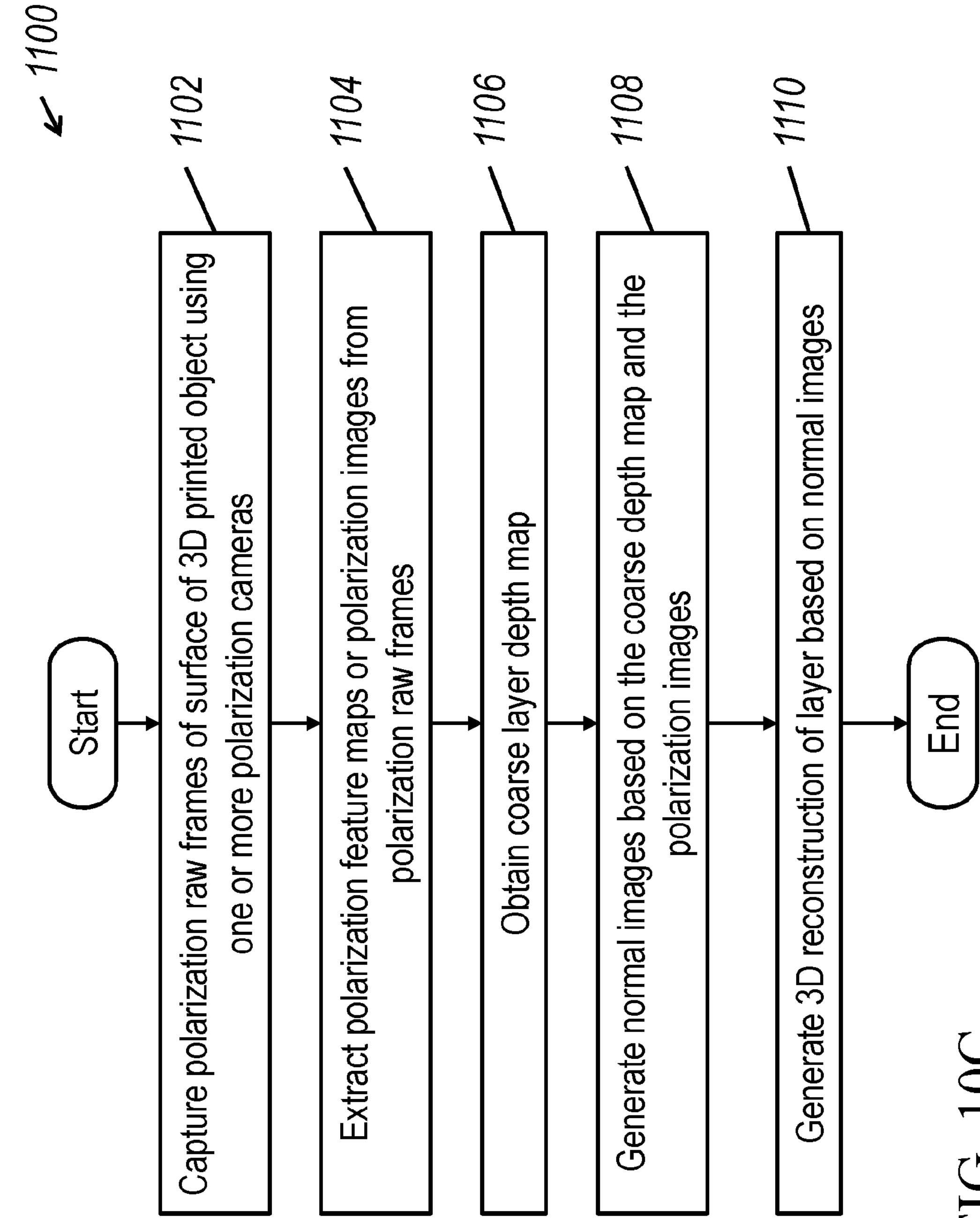
1100
Start
Capture polarization raw frames of surface of 3D printed object using one or more polarization cameras
1102
Extract polarization feature maps or polarization images from polarization raw frames
1104
Obtain coarse layer depth map
1106
Generate normal images based on the coarse depth map and the polarization images
1108
Generate 3D reconstruction of layer based on normal images
1110
End

FIG. 10C

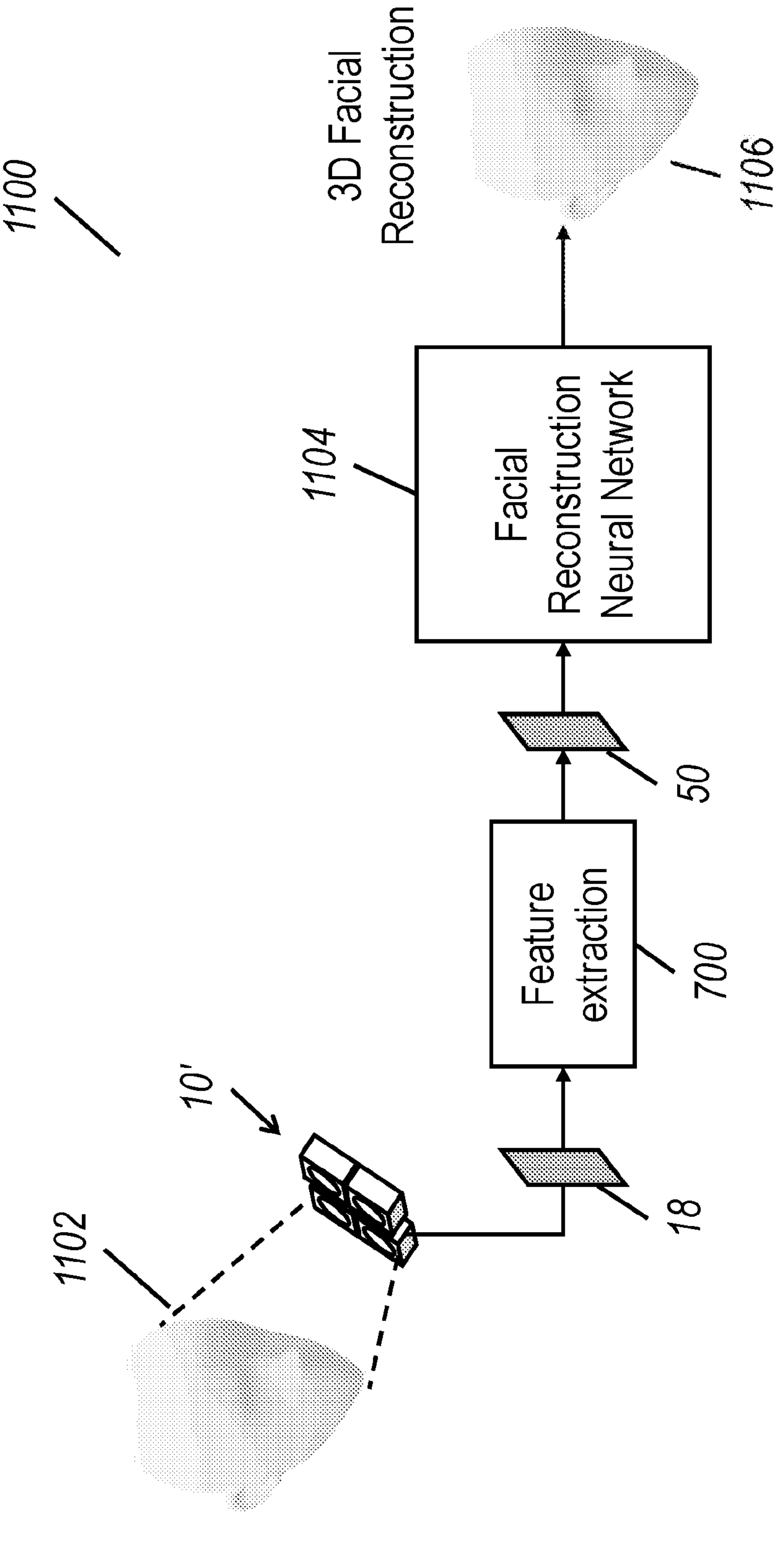
1100
3D Facial Reconstruction
1106
1104
Facial Reconstruction Neural Network
50
Feature extraction
700
10'
1102
18

FIG. 11A

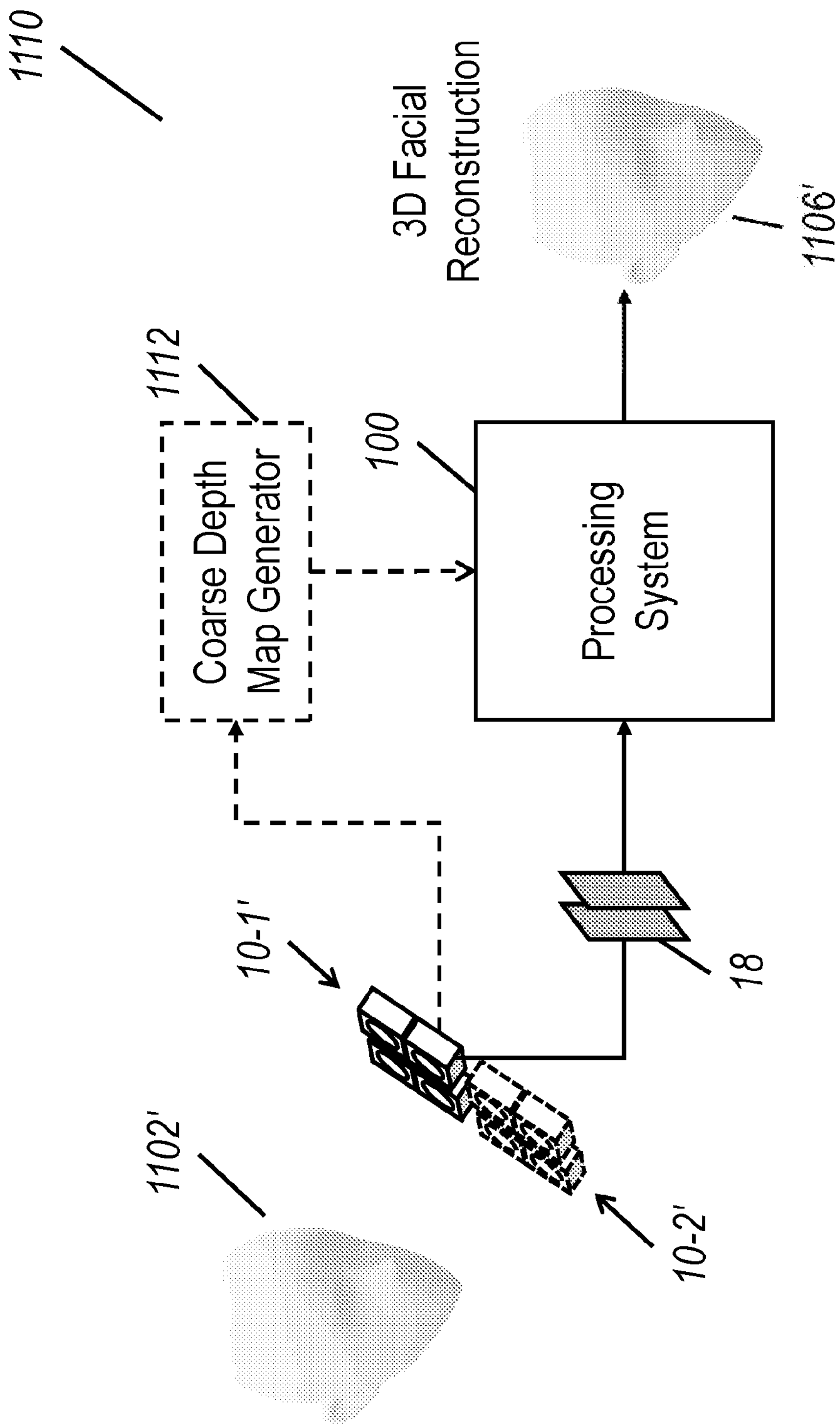
1110
1112
Coarse Depth Map Generator
100
Processing System
3D Facial Reconstruction
1106'
10-1'
1102'
10-2'
18

FIG. 11B

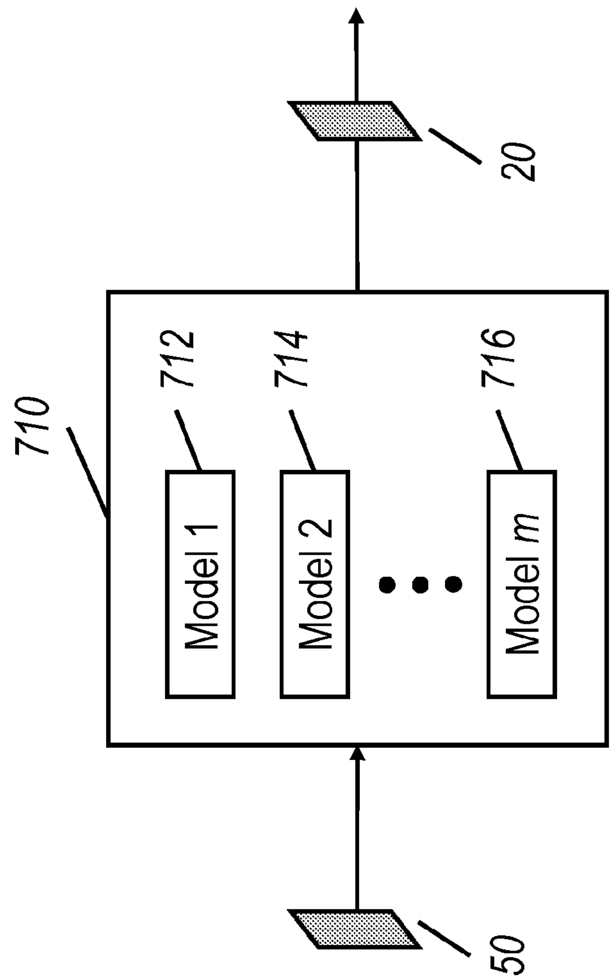

FIG. 12

INTERNATIONAL SEARCH REPORT

International application No.
PCT/US 20/54645

A. CLASSIFICATION OF SUBJECT MATTER

IPC - G01B 11/24 (2021.01)

CPC - B29C 64/386; B29C 64/393; B33Y 50/00; B33Y 50/02; G01B 11/24; G06T 2207/10028; G06T 2207/10141; G06T 7/50; G06T 7/586

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

B. FIELDS SEARCHED

Minimum documentation searched (classification system followed by classification symbols)
See Search History document

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

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

C. DOCUMENTS CONSIDERED TO BE RELEVANT

Category*	Citation of document, with indication, where appropriate, of the relevant passages	Relevant to claim No.
Y	US 2019/0186901 A1 (MASSACHUSETTS INSTITUTE OF TECHNOLOGY) 20 June 2019 (20.06.2019): abstract, [0007], [0018], [0039], [0041], [0043], [0053], [0195]	1-16
Y	US 2019/0143412 A1 (VELO3D INC) 16 May 2019 (16.05.2019): abstract, [0073]	1-16
Y	US 2016/0344948 A1 (MTU AERO ENGINES AG) 24 November 2016 (24.11.2016): abstract, [0018]	16

☐ Further documents are listed in the continuation of Box C. ☐ See patent family annex.

* Special categories of cited documents:

"A" document defining the general state of the art which is not considered to be of particular relevance

"D" document cited by the applicant in the international application

"E" earlier application or patent but published on or after the international filing date

"L" document which may throw doubts on priority claim(s) or which is cited to establish the publication date of another citation or other special reason (as specified)

"O" document referring to an oral disclosure, use, exhibition or other means

"P" document published prior to the international filing date but later than the priority date claimed

"T" later document published after the international filing date or priority date and not in conflict with the application but cited to understand the principle or theory underlying the invention

"X" document of particular relevance; the claimed invention cannot be considered novel or cannot be considered to involve an inventive step when the document is taken alone

"Y" document of particular relevance; the claimed invention cannot be considered to involve an inventive step when the document is combined with one or more other such documents, such combination being obvious to a person skilled in the art

"&" document member of the same patent family

Date of the actual completion of the international search	Date of mailing of the international search report
22 January 2021	17 FEB 2021

Name and mailing address of the ISA/US	Authorized officer
Mail Stop PCT, Attn: ISA/US, Commissioner for Patents P.O. Box 1450, Alexandria, Virginia 22313-1450 Facsimile No. 571-273-8300	Lee Young Telephone No. PCT Helpdesk: 571-272-4300

INTERNATIONAL SEARCH REPORT

International application No.
PCT/US 20/54645

Box No. II Observations where certain claims were found unsearchable (Continuation of item 2 of first sheet)

This international search report has not been established in respect of certain claims under Article 17(2)(a) for the following reasons:

1. ☐ Claims Nos.:
because they relate to subject matter not required to be searched by this Authority, namely:

2. ☐ Claims Nos.:
because they relate to parts of the international application that do not comply with the prescribed requirements to such an extent that no meaningful international search can be carried out, specifically:

3. ☐ Claims Nos.:
because they are dependent claims and are not drafted in accordance with the second and third sentences of Rule 6.4(a).

Box No. III Observations where unity of invention is lacking (Continuation of item 3 of first sheet)

This International Searching Authority found multiple inventions in this international application, as follows:

This application contains the following inventions or groups of inventions which are not so linked as to form a single general inventive concept under PCT Rule 13.1. In order for all inventions to be searched, the appropriate additional search fees must be paid.

Group I: Claims 1-16 drawn to a method of performing surface profilometry or a surface profilometry system.

Group II: Claims 17-30 drawn to a method of capturing a 3D image of a face or a 3D imaging system for capturing a 3D image of a face.

Group III: Claims 31-37 drawn to a method of computing a prediction.

--see Extra Sheet

1. ☐ As all required additional search fees were timely paid by the applicant, this international search report covers all searchable claims.

2. ☐ As all searchable claims could be searched without effort justifying additional fees, this Authority did not invite payment of additional fees.

3. ☐ As only some of the required additional search fees were timely paid by the applicant, this international search report covers only those claims for which fees were paid, specifically claims Nos.:

4. ☒ No required additional search fees were timely paid by the applicant. Consequently, this international search report is restricted to the invention first mentioned in the claims; it is covered by claims Nos.:
1-16

Remark on Protest

☐ The additional search fees were accompanied by the applicant's protest and, where applicable, the payment of a protest fee.

☐ The additional search fees were accompanied by the applicant's protest but the applicable protest fee was not paid within the time limit specified in the invitation.

☐ No protest accompanied the payment of additional search fees.

Continued from Box III -- Observations where unity of invention is lacking

The inventions listed as Groups I-III do not relate to a single general inventive concept under PCT Rule 13.1 because, under PCT Rule 13.2, they lack the same or corresponding special technical features for the following reasons:

Special Technical Features:
Group I includes special technical features of surface prolimetry applied to a physical object undergoing additive manufacturing, not included in the other groups.

Group II includes special technical features relating to of capturing a 3D image of a face, not included in the other groups.

Group III includes the special technical feature of computing a prediction based on a the surface-normal image of an object, not included in the other groups.

Common Technical Features:
The only technical features shared by Groups I-III that would otherwise unify the groups, is that Groups I-III all involve processing surface-normal images and Groups I and II both involve one or more polarization cameras capturing raw image frames and processing image data based on a coarse layer depth map.

However, these shared technical features do not represent a contribution over prior art, because the shared technical features are disclosed by US 2016/0261844 A1 to Massachusetts Institute of Technology (hereinafter 'M.I.T.').

M.I.T. discloses processing surface-normal images and one or more polarization cameras (102) capturing raw image frames and processing image data based on a coarse layer depth map from depth camera (101) (Fig 1A, abstract, para [0058]).

As the shared technical features were known in the art at the time of the invention, they cannot be considered special technical features that would otherwise unify the groups.

Therefore, Groups I-III lack unity under PCT Rule 13.