

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

(19) World Intellectual Property Organization

International Bureau



(10) International Publication Number

WO 2013/112237 A1

(43) International Publication Date

1 August 2013 (01.08.2013)

(51) International Patent Classification:

G06T 7/00 (2006.01)

(21) International Application Number:

PCT/US2012/069088

(22) International Filing Date:

12 December 2012 (12.12.2012)

(25) Filing Language:

English

(26) Publication Language:

English

(30) Priority Data:

61/591,196 26 January 2012 (26.01.2012) US  
13/673,681 9 November 2012 (09.11.2012) US

(71) Applicant: QUALCOMM INCORPORATED [US/US];  
ATTN: International IP Administration, 5775 Morehouse Drive, San Diego, California 92121-1714 (US).

(72) Inventors: QI, Yingyong; 5775 Morehouse Drive, San Diego, California 92121-1714 (US). YANG, Ruiduo; 5775 Morehouse Drive, San Diego, California 92121-1714 (US). SUNDARESAN, Sairam; 5775 Morehouse Drive, San Diego, California 92121-1714 (US).

(74) Agent: TOLER, JEFFREY G.; 8500 Bluffstone Cove, Suite A201, Austin, Texas 78759 (US).

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

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

Published:

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

(54) Title: MOBILE DEVICE CONFIGURED TO COMPUTE 3D MODELS BASED ON MOTION SENSOR DATA

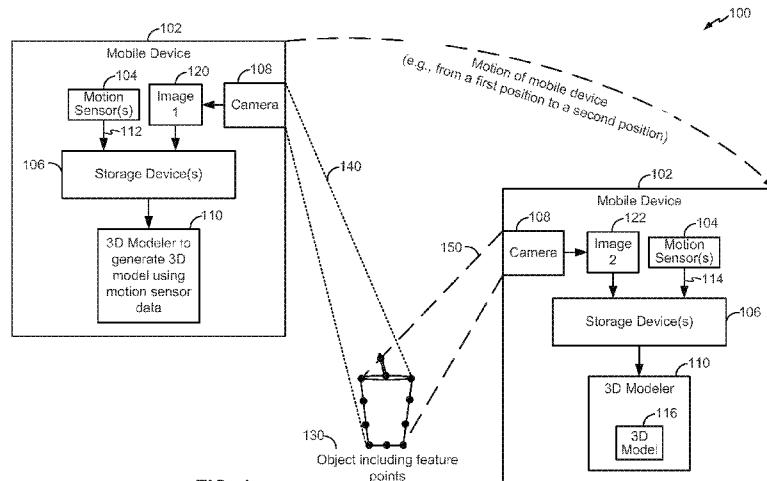


FIG. 1

(57) Abstract: A particular method includes determining, based on data received from at least one motion sensor, a movement of a mobile device from a first position to a second position. The method also includes computing a three-dimensional (3D) model of an object based on a first image of the object corresponding to a first view of the object from the first position of the mobile device, a second image of the object corresponding to a second view of the object from the second position of the mobile device, and the movement of the mobile device.

## MOBILE DEVICE CONFIGURED TO COMPUTE 3D MODELS BASED ON MOTION SENSOR DATA

### *I. Cross-Reference to Related Applications*

[0001] The present application claims the benefit of U.S. Provisional Patent Application No. 61/591,196, filed January 26, 2012 and U.S. Non-Provisional Patent Application No. 13/673,681, filed November 9, 2012, the disclosure of which is incorporated by reference herein in its entirety.

### *II. Field*

[0002] The present disclosure is generally related to generating three-dimensional (3D) models of objects.

### *III. Description of Related Art*

[0003] Advances in technology have resulted in smaller and more powerful computing devices. For example, there currently exist a variety of portable personal computing devices, including wireless computing devices, such as portable wireless telephones, personal digital assistants (PDAs), and paging devices that are small, lightweight, and easily carried by users. More specifically, portable wireless telephones, such as cellular telephones and internet protocol (IP) telephones, can communicate voice and data packets over wireless networks. Also, such wireless telephones can process executable instructions, including software applications, such as a web browser application, that can be used to access the Internet. As such, these wireless telephones can include significant computing capabilities and may be capable of graphics processing of three-dimensional (3D) objects.

[0004] A three-dimensional (3D) model of an object may be generated based on multi-view images of an object (e.g., multiple views of the object taken by one or more cameras from different angles and/or positions). The 3D models may provide useful and interesting information to a user of a wireless device and may improve an overall

user experience. 3D reconstruction may be used in security, surveillance, and robotics. In addition, 3D reconstruction may be used for developing models of buildings for generating virtual tours or for navigational purposes. Furthermore, objects that may have been destroyed may be reconstructed and preserved (e.g., in museums).

#### *IV. Summary*

[0005] Systems and methods of motion regularized object reconstruction in multi-view processing of a wireless device are disclosed. The disclosed embodiments utilize motion information of the wireless device provided by one or more motion sensors to reduce the complexity of generating a three-dimensional (3D) model (e.g., a 3D structure) in connection with one or more 3D reconstruction techniques.

[0006] According to a particular example, a feature point-based reconstruction technique (e.g., a bundle adjustment technique) is used in connection with motion sensor data to calculate a 3D model. According to some approaches, bundle adjustment is a “maximum likelihood solution” for the following expression

$$\min_{P^i X_j} \sum_{i,j} d(P^i X_j, x_j^i)^2$$
, where  $P^i$  is an i-th projection matrix of the camera (i.e., a representation of feature points of the object taken by the camera and projected onto an i-th image), where  $X_j$  is a j-th feature point of the object, where  $x_j^i$  are coordinates of the j-th feature point in the i-th image, and where  $d(\cdot)$  represents a distance function, as explained further below. Feature points of an object include parts of the object that may be important for generating a 3D model of the object (e.g., corners in the object, curves in the object, edges of the object, etc). For example, feature points may include the handle of a mug, corners of a graphics pattern of the mug, portions of the mug that include text, and the base of the mug.

[0007] According to at least one use of the bundle adjustment technique, initial (e.g.,  $i = 1$ ) and subsequent (e.g.,  $i = 2, 3, \dots, n$ ) projection matrices  $P^i$  estimating position and orientation of the camera from the initial position to subsequent positions of the camera are estimated. The feature points  $X_j$  of the object are estimated, and coordinates of the feature points  $x_j^i$  in multiple images of the object captured by the camera are also estimated. The bundle adjustment solution may result in a large and computationally

intensive problem because of the large number of parameters to be estimated. Thus, the bundle adjustment solution may be computationally demanding and may place a burden on processor resources.

[0008] According to at least one embodiment of the present disclosure, complexity of the bundle adjustment technique may be reduced by using motion sensor data to determine  $P^{i+1}$  (or an estimate of  $P^{i+1}$ ). For example, bundle adjustment complexity may be reduced according to the expression:

$$\min_{P^i, X_j} \sum_{i,j} d(P^i X_j, x_j^i)^2, \text{ subject to } P^{i+1} = [R_i \mid T_i] P^i, \text{ where } R_i \text{ and } T_i \text{ are rotational and}$$

translation matrices, respectively, and  $[R_i \mid T_i]$  is a relative pose matrix associated with the wireless device that is based on measurements made by the one or more sensors of the wireless device. Thus, at least one  $i$ -th+1 projection matrix may be derived from the motion information of the wireless device provided by the one or more motion sensors, and need not be iteratively estimated via the bundle adjustment technique, resulting in reduced overall complexity.

[0009] According to another particular example, motion sensor measurements are used to derive camera location information to generate a silhouette-based 3D object (e.g., a 3D object generated using a visual hull technique). By using the motion sensor measurements, silhouette-based 3D models can be generated effectively by a mobile device (e.g., a device that is not associated with a “fixed” or predetermined location).

[0010] In a particular embodiment, a method includes capturing, at a camera of a mobile device, a first image of an object while the mobile device is at a first position and capturing a second image of the object while the mobile device is at a second position. The method includes determining, based on data output by at least one motion sensor of the mobile device, a movement of the mobile device from the first position to the second position. The method also includes computing a three-dimensional (3D) model of the object based on the first image, the second image, and the determined movement of the mobile device.

[0011] In a particular embodiment, an apparatus includes a camera configured to capture a first image of an object while at a first position and to capture a second image

of the object while at a second position. The apparatus includes at least one motion sensor, a processor, and a memory storing instructions executable by the processor. The instructions are executable to determine, based on data output by the at least one motion sensor, a movement of the camera from the first position to the second position and to compute a three-dimensional (3D) model of the object based on the first image, the second image, and the determined movement of the camera.

[0012] In a particular embodiment, a method includes determining, based on data received from at least one motion sensor, a movement of a mobile device from a first position to a second position. The method also includes computing a three-dimensional (3D) model of an object based on a first image of the object corresponding to a first view of the object from the first position of the mobile device, based further on a second image of the object corresponding to a second view of the object from the second position of the mobile device, and based further on the movement of the mobile device.

[0013] In a particular embodiment, an apparatus includes means for determining, based on data received from at least one means for sensing motion, a movement of a mobile device from a first position to a second position. The apparatus also includes means for computing a three-dimensional (3D) model of an object based on a first image of the object corresponding to a first view of the object from the first position of the mobile device, based further on a second image of the object corresponding to a second view of the object from the second position of the mobile device, and based further on the movement of the mobile device.

[0014] In a particular embodiment, a computer-readable non-transitory medium includes instructions that, when executed by a processor, cause the processor to determine, based on data received from at least one motion sensor, a movement of a mobile device from a first position to a second position. The instructions are further executable to compute a three-dimensional (3D) model of an object based on a first image of the object corresponding to a first view of the object from the first position of the mobile device, based further on a second image of the object corresponding to a second view of the object from the second position of the mobile device, and based further on the movement of the mobile device.

[0015] Other aspects, advantages, and features of the present disclosure will become apparent after review of the entire application, including the following sections: Brief Description of the Drawings, Detailed Description, and the Claims.

*V. Brief Description of the Drawings*

[0016] FIG. 1 is a diagram of a particular embodiment of a system that includes a mobile device configured to compute 3D models based on motion sensor data;

[0017] FIG. 2 is a diagram of a particular illustrative embodiment of the mobile device of FIG. 1;

[0018] FIG. 3 is a diagram of another particular illustrative embodiment of the mobile device of FIG. 1;

[0019] FIG. 4 is a diagram of another particular illustrative embodiment of the mobile device of FIG. 1;

[0020] FIG. 5 is a flow diagram illustrating a particular embodiment of a method of computing a 3D model based on motion sensor data;

[0021] FIG. 6 is a flow diagram illustrating another particular embodiment of a method of computing a 3D model based on motion sensor data;

[0022] FIG. 7 is a flow diagram illustrating a particular embodiment of a method of computing a 3D model according to a feature point-based reconstruction technique (e.g., a bundle adjustment technique) based on motion sensor data;

[0023] FIG. 8 is a flow diagram illustrating a particular embodiment of a method of computing a 3D model according to a silhouette-based reconstruction technique (e.g., a visual hull technique) based on motion sensor data;

[0024] FIG. 9 is a flow diagram illustrating another particular embodiment of a method of computing a 3D model according to a silhouette-based reconstruction technique (e.g., a visual hull technique) based on motion sensor data;

[0025] FIG. 10 is a diagram of another particular illustrative embodiment of the mobile device of FIG. 1, where the mobile device is configured to compute 3D models

according to a feature point-based reconstruction technique (e.g., a bundle adjustment technique) based on motion sensor data; and

[0026] FIG. 11 is a diagram of another particular illustrative embodiment of the mobile device of FIG. 1, where the mobile device is configured to compute 3D models according to a silhouette-based reconstruction technique (e.g., a visual hull technique) based on motion sensor data.

#### *VI. Detailed Description*

[0027] Referring to FIG. 1, a diagram of a particular embodiment of a system is disclosed and generally designated 100. The system 100 includes a mobile device 102 including a camera 108 configured to capture one or more images (e.g., a first image 120 and a second image 122) of one or more objects, such as an object 130. The mobile device 102 further includes one or more motion sensors 104, one or more storage devices 106, and a 3D modeler 110. The 3D modeler 110 may be implemented using hardware, using processor-executable instructions stored at a computer-readable tangible medium, or a combination thereof, as described further below. The one or more motion sensors 104 may include one or more electrical, mechanical, electromechanical, or other sensors, such as a gyroscope, an accelerometer, other motion sensor devices, or a combination thereof.

[0028] The object 130 may include one or more feature points, such as apexes of the object 130, a base of the object 130, a handle of the object 130, text or graphics on the object 130, curves of the object 130, one or more other distinct portions of the object 130, or any combination thereof. Although FIG. 1 shows two images 120, 122 of the object 130 captured by the camera 108 when the mobile device 102 is at two positions, it should be noted that multiple images of the object 130 corresponding to multiple positions of the mobile device 102 may be captured by the camera 108.

[0029] The camera 108 may be configured to capture images of the object 130 at multiple positions of the mobile device 102 (e.g., at different angles and/or rotations with respect to a coordinate system or with respect to an initial position). For example, the camera 108 may capture a first image 120 of the object 130 when the mobile device 102 (and the camera 108) is at a first position (e.g., a first view 140 of the object 130)

and the camera 108 may capture a second image 122 of the object 130 when the mobile device 102 (and the camera 108) is at a second position (e.g., a second view 150 of the object 130). In at least one embodiment, image data corresponding to the first view 140 and to the second view 150 may be stored at the one or more storage devices 106.

**[0030]** The one or more motion sensors 104 may be configured to generate motion sensor data that indicates a position of the mobile device 102 (e.g., rotation and/or translation of the mobile device 102 with respect to an initial position of the mobile device 102 or with respect to a coordinate system). For example, the one or more motion sensors 104 may be configured to generate motion sensor data 112 concurrently with the camera 108 capturing the first image 120. As another example, the one or more motion sensors 104 may be configured to generate motion sensor data 114 concurrently with the camera 108 capturing the second image 122. In at least one embodiment, the mobile device 102 uses the 3D modeler 110 to generate a 3D model 116 of the object 130 using the motion sensor data 112, 114 and further using image data related to the images 120, 122, as described further below.

**[0031]** During operation, the mobile device 102 may capture the first image 120 of the object 130 via the camera 108 based on the first view 140 of the camera 108. The mobile device 102 may move to a second position (e.g., due to movement of a user of the mobile device 102) and may capture the second image 122 of the object 130 via the camera 108 based on the second view 150 of the camera 108. The one or more motion sensors 104 of the mobile device 102 may output motion sensor data 112, 114 corresponding to motion information of the mobile device 102 that enables determination of movement of the mobile device 102 from the first position to the second position. The motion sensor data 112, 114 may indicate rotational and translational motion of the mobile device 102.

**[0032]** To illustrate, when the mobile device 102 is at the first position, the 3D modeler 110 may receive data corresponding to the first image 120 and may further receive the motion sensor data 112 output by the one or more motion sensors 104. Similarly, when the mobile device 102 is at the second position, the 3D modeler 110 may receive the data corresponding to the second image 122 and may further receive the motion sensor data 114 output by the one or more motion sensors 104. The motion sensor data 114

may indicate movement information of the mobile device 102 from the first position to the second position. For example, the motion sensor data 114 may include rotational and translational matrices of the mobile device 102 associated with rotation and translation of the mobile device 102 from the first position to the second position. As another example, the motion sensor data 112, 114 may provide information that enables generation of rotational and translational matrices of the mobile device 102. The motion sensor data 112, 114 may include or indicate a “camera pose” matrix obtained from the one or more motion sensors 104, as described further below.

[0033] For purposes of illustration, a non-limiting, example operation is provided of the 3D modeler 110 using the motion sensor data 112, 114 and data corresponding to the images 120, 122 to generate the 3D model 116 of the object 130. In a particular example, the 3D modeler 110 is configured to generate the 3D model 116 by calculating the bundle adjustment expression  $\min_{P^i, X_j} \sum_{i,j} d(P^i X_j, x_j^i)^2$  based on the motion sensor data

114 provided by the one or more motion sensors 104. According to such embodiments and as described further below, the 3D modeler 110 may output 3D feature points of the object 130 based on the calculation. The output 3D feature points correspond to a 3D model of the object 130. In a particular embodiment, computing the bundle adjustment expression may include determining feature points of the object 130 based on a function of a first estimated reprojected image point, a second estimated reprojected image point, and the data output by the one or more motion sensors 104. For example, computing the bundle adjustment expression may include determining the first estimated reprojected image point of the object 130 based on a product of a first estimate of the first position (i.e.,  $P^1$ ) and one or more feature points ( $X_j$ ) of the object 130 (i.e., computing  $P^1 X_j$ ). Computing the bundle adjustment expression may also include determining a second estimated reprojected image point of the object based on a product of a second estimate of the second position (i.e.,  $P^2$ ) and the one or more feature points ( $X_j$ ) of the object 130 (i.e., computing  $P^2 X_j$ ). Computing the bundle adjustment expression may also include determining a square of a first Euclidean distance between the first estimated reprojected image point and coordinates of the object 130 in the first image 120 (i.e., computing  $d(P^1 X_j, x_j^1)^2$ ) and determining a square of a second

Euclidean distance between the second estimated reprojected image point and coordinates of the object 130 in the second image 122 (i.e., computing  $d(P^2 X_j, x_j^2)$ ). Computing the bundle adjustment expression may include modifying an estimate of the 3D model 116 to reduce or to “minimize” a sum of the square of the first Euclidean distance and the square of the second Euclidean distance (i.e., “minimizing”  $d(P^1 X_j, x_j^1)^2 + d(P^2 X_j, x_j^2)^2$ ). The first estimate of the first position  $P^1$  may also be modified to further reduce the sum of the squares. However, the second estimate of the second position  $P^2$  is constrained such that  $P^{i+1} = \beta P^i$ , where  $\beta$  indicates camera pose information, such as a rotation matrix  $R$  and/or a transformation matrix  $T$ . For example, in at least one embodiment,  $\beta = [R_i | T_i]$ . By constraining  $P^2 \dots P^n$ , a complexity of generating the feature points may be significantly reduced. It should be noted that although a Euclidean distance function is shown in connection with the foregoing example, other distance functions may be utilized. For example,  $d(\cdot)$  may represent a distance function such as a geometric distance or any other metric that measures the distance or difference between the projected feature coordinates and the measured feature coordinates in the images.

[0034] The camera pose information may be determined according to a variety of techniques. To illustrate, suppose at least one of the one or more feature points  $X_j$  is “matched” (e.g., identified) in  $N$  images (e.g., the first image 120 and the second image 122) captured by the camera 108. The matched feature point may be denoted as  $m^i$ , where  $m^i = [x^i, y^i, 1]^T$  for  $(i = 1, \dots, N)$ . The matched feature point  $m^i$  is associated with a spatial position  $M = [X \ Y \ Z \ 1]^T$ . Accordingly, for a scaling factor  $\mu^i$ ,  $\mu^i [x^i, y^i, 1]^T = P^i M$  for  $i = 1, \dots, N \geq 2$ , where  $P^i = K^i [R^i | T^i]$ .  $P^i = K^i [R^i \ T^i]$  is the  $i$ th camera projection matrix and  $K^i$  is a upper triangular calibration matrix (e.g., a 3x3 upper triangular calibration matrix) of the  $i$ -th camera of the mobile device 102 for  $i = 1, \dots, N$  (e.g., for a multi-camera embodiment of the camera 108, which is omitted from FIG. 1 for clarity of illustration). The upper triangular calibration matrix  $K^i$  may be the same for images captured with a same camera (e.g., when the camera 108 is used to capture each of the  $N$  images, such as for a single-camera embodiment of the mobile device 102).  $[R^i | T^i]$  may be the

relative pose matrix associated with the  $i$ -th camera. Thus,

$$AM = 0 \quad \text{where } A = \begin{bmatrix} p_1^i - x^i p_1^i \\ p_1^i - y^i p_1^i \\ p_1^i - z^i p_1^i \\ p_2^i - x^i p_2^i \\ p_2^i - y^i p_2^i \\ \vdots \\ p_N^i - x^i p_N^i \\ p_N^i - y^i p_N^i \end{bmatrix}$$

[0035] Here,  $p_j^i$  is the  $j$ -th row of the  $i$ -th camera's projection matrix and  $AM = 0$  may be a set of  $2N$  homogeneous equations with three unknowns (X, Y, Z). Because all camera projection matrixes may be known and  $N$  may be equal to or greater than 2, the equation may have a least squares solution, which may enable "minimization" of algebraic reprojection error.

[0036] It will be appreciated that a bundle adjustment technique in accordance with the foregoing description may utilize camera pose data (e.g., the motion sensor data 112, 114) to enable simplified and computationally efficient determination of the 3D model 116. Conventional approaches (e.g., bundle adjustment calculations that do not utilize motion sensor information) may be subject to a computationally intensive "least squares solution," such as by obtaining a nonlinear solution to "minimize" geometric error:

$$[0037] \quad \min \sum_{i=1}^N [(x^i - \frac{p_1^i N}{p_3^i})^2 + (y^i - \frac{p_2^i N}{p_3^i})^2]$$

[0038] FIG. 2 depicts a particular illustrative embodiment of the mobile device 102. The mobile device 102 of FIG. 2 includes the one or more motion sensors 104, the one or more storage devices 106, the camera 108, and the 3D modeler 110.

[0039] In the particular embodiment of FIG. 2, the one or more storage devices 106 store 3D reconstruction data 202 to be used by the 3D modeler 110 to generate the 3D model 116. The 3D reconstruction data 202 may include the motion sensor data 112, 114 and may further include first image data 220 and second image data 222 corresponding to the first image 120 of FIG. 1 and to the second image 122 of FIG. 1,

respectively. The 3D model 116 may be displayed on a display of the mobile device 102 (not shown in FIG. 2).

[0040] According to a particular illustrative embodiment, the image data 220, 222 of FIG. 2 corresponds to a sequence of pictures, each picture individually captured in response to user input (e.g., “still” or “snapshot” pictures that are “manually” captured by a user and not part of a video sequence). In at least one embodiment, the example of FIG. 2 may be used in connection with an application of the mobile device 102 that enables users to generate 3D models, such as the 3D model 116, based on still pictures captured by users using the camera 108. According to additional embodiments and as discussed further below, the image data 220, 222 may correspond to video frames of a video frame sequence (e.g., sequential still images that are “automatically” periodically captured).

[0041] FIG. 3 depicts another particular illustrative embodiment of the mobile device 102 described with reference to FIG. 1. The mobile device 102 of FIG. 3 includes the one or more motion sensors 104, the one or more storage devices 106, the camera 108, the 3D modeler 110, and a display 320. The 3D modeler 110 of FIG. 3 includes a synchronizer 308. The one or more storage devices 106 of FIG. 3 include a motion sensor data buffer 306 and a video frame buffer 304. Optionally, the mobile device 102 of FIG. 3 may include a main memory 310.

[0042] In operation, the camera 108 may capture a video frame sequence 302. The video frame buffer 304 may be responsive to the camera 108 and configured to store one or more video frames of the video frame sequence 302. For example, in the example of FIG. 3, the image data 220, 222 correspond to video frames of the video frame sequence 302. The motion sensor data buffer 306 may be responsive to the one or more motion sensors 104. The motion sensor data buffer 306 may store the motion sensor data 112, 114.

[0043] The synchronizer 308 may synchronize data from the video frame buffer 304 with respect to data from the motion sensor data buffer 306. For example, the synchronizer 308 may synchronize the first motion sensor data 112 with the first image data 220 and may further synchronize the second motion sensor data 114 with the second image data 222. According to at least a first technique, the synchronizer 308

synchronizes data from the video frame buffer 304 with respect to data from the motion sensor data buffer 306 by comparing timestamps, such as by comparing timestamps associated with the motion sensor data 112, 114 to timestamps associated with the image data 220, 222. According to at least a second technique, the synchronizer 308 periodically accesses data from the motion sensor data buffer 306 and from the video frame buffer 304. According to at least a third technique, the synchronizer 308 is responsive to the movement of the mobile device 102 (e.g., is responsive to the one or more motion sensors 104) and pulls data from the motion sensor data buffer 306 and from the video frame buffer 304 in response to the detected motion. Further techniques may utilize a combination of one or more of the first technique, the second technique, and the third technique. Optionally, the data output by the video frame buffer 304 may be stored at the main memory 310 as a video file 312 that includes the video frame sequence 302.

[0044] As will be appreciated, by pulling data from “temporary” storage (e.g., the video frame buffer 304 and the motion sensor data buffer 306 of FIG. 3) instead of the main memory 310, the 3D modeler 110 may operate “on the fly.” For example, the 3D modeler 110 may generate the 3D model 116 and the mobile device 102 may display the 3D model 116 on the display 320 concurrently or substantially concurrently with the video frame sequence 302 being captured. As used herein, “on the fly,” “concurrently,” and “substantially concurrently” refer to a relationship between a view of a captured object and generation of a 3D model of the object that a user may perceive as being simultaneous.

[0045] Accordingly, the example illustrated in connection with FIG. 3 may correspond to an “on the fly” generation of the 3D model 116. For example, as a user moves the mobile device 102 about an object (such as the object 130 of FIG. 1), the 3D modeler 110 may “automatically” render the 3D model 116, which is displayed on the display 320 concurrently with the user movement. According to additional embodiments and as discussed further below, one or more embodiments may enable non-concurrent generation of 3D models based on motion sensor data and a video frame sequence.

[0046] FIG. 4 depicts another particular illustrative embodiment of the mobile device 102 described with reference to FIG. 1. The mobile device 102 of FIG. 4 includes the

one or more motion sensors 104, the one or more storage devices 106, the camera 108, the 3D modeler 110, the video frame buffer 304, and the display 320.

[0047] In operation, the one or more motion sensors 104 may generate the motion sensor data 112, 114, which may be stored at the one or more storage devices 106. The camera 108 may capture the video frame sequence 302, which may be stored temporarily at the video frame buffer 304 and displayed at the display 320, such as in connection with a “preview” feature associated with the video frame sequence 302. For example, the video frame sequence 302 may be displayed at the display 320 prior to the video frame sequence 302 being stored at the one or more storage devices 106 (e.g., as the image data 220, 222) and/or prior to the 3D modeler 110 generating the 3D model 116.

[0048] Accordingly, in the particular embodiment of FIG. 4, the 3D modeler 110 may generate the 3D model 116 according to a “post-processing” technique. For example, because the 3D model 116 may not be generated and displayed concurrently with the video frame sequence 302 being captured by the camera 108, the display 320 can instead show a “preview” of the images being captured by the camera 108. Further, the display 320 may display user instructions for camera positioning (e.g., angles and/or positions at which the camera is to be held with respect to object being viewed). The 3D model 116 can thereafter be generated by the 3D modeler 110 by accessing the one or more storage devices 106, which may be a “main memory,” such as the main memory 310 described in connection with FIG. 3.

[0049] Referring to FIG. 5, a diagram illustrating a particular embodiment of a method of computing a 3D model based on motion sensor data is disclosed and generally designated 500. The method 500 includes capturing, at a camera of a mobile device, a first image of an object while the mobile device is at a first position, at 502, and capturing a second image of the object while the mobile device is at a second position, at 504. For example, the camera 108 may capture the first image 120 of the object 130 while the mobile device 102 is at the first position and the camera 108 may capture the second image 122 of the object 130 while the mobile device 102 is at the second position. The first image 120 may correspond to the first view 140 of the object 130 as captured by the camera 108 while at the first position and the second image 122 may

correspond to the second view 150 of the object 130 as captured by the camera 108 while at the second position.

[0050] The method 500 further includes determining, based on data output by at least one motion sensor of the mobile device, a movement of the mobile device from the first position to the second position, at 506. For example, the one or more motion sensors 104 may output motion sensor data 112, 114 corresponding to motion information of the mobile device 102. In a particular embodiment, the motion information may include rotation and translation of the mobile device 102 from the first position to the second position. To illustrate, the one or more motion sensors may provide first rotational and/or translational matrices of the mobile device 102 corresponding to the first position and second rotational and/or translational matrices of the mobile device 102 corresponding to the second position. In at least one embodiment, the first position of the mobile device is used as a “reference” position, and the second position accordingly indicates a change in position of the mobile device. According to other embodiments, another position (e.g., a previous position, or a position associated with a predetermined coordinate system) is used as the “reference” position, and both the first position and the second position each indicate a change in position of the mobile device.

[0051] The method 500 also includes computing a 3D model of the object based on the first image, the second image, and the determined movement of the mobile device, at 508. For example, according to a feature point-based reconstruction technique and as described with reference to at least FIG. 1, the 3D modeler 110 may compute feature points of the object 130 based on the first image 120 and the second image 122. In a particular embodiment, the 3D modeler 110 may be configured to calculate the bundle adjustment expression  $\min_{P^i, X_j} \sum_{i,j} d(P^i X_j, x'_j)^2$  based on the motion sensor data 114

provided by the one or more motion sensors 104, the first image 120, and the second image 122. According to further embodiments and as described further below, the 3D model may be computed according to a silhouette-based technique (e.g., a visual hull technique).

[0052] Referring to FIG. 6, a diagram illustrating another particular embodiment of a method of computing a three-dimensional (3D) model based on motion sensor data is disclosed and generally designated 600. The method 600 includes determining, based

on data received from at least one motion sensor, a movement of a mobile device from a first position to a second position, at 602. The data, the at least one motion sensor, and the mobile device may correspond to the motion sensor data 112, 114, the one or more motion sensors 104, and the mobile device 102, respectively.

[0053] The method 600 further includes computing a 3D model of an object based on a first image of the object corresponding to a first view of the object from the first position of the mobile device, based further on a second image of the object corresponding to a second view of the object from the second position of the mobile device, and based further on the movement of the mobile device, at 604. For example, the camera 108 may capture the first image 120 of the object 130 while the mobile device 102 is at the first position and the camera 108 may capture the second image 122 of the object 130 while the mobile device 102 is at the second position. The first image 120 may correspond to the first view 140 of the object 130 as captured by the camera 108 while at the first position and the second image 122 may correspond to the second view 150 of the object 130 as captured by the camera 108 while at the second position.

[0054] Referring to FIG. 7, a flow diagram illustrating a particular embodiment of a method of computing a 3D model according to a feature point-based reconstruction technique (e.g., a bundle adjustment technique) based on motion sensor data is depicted and generally designated 700. The method 700 includes capturing images (e.g., the images 120, 122) of an object (e.g., the object 130) using a camera (e.g., the camera 108) of a mobile device (e.g., the mobile device 102), at 702.

[0055] At 704, the method 700 includes determining motion sensor data (e.g., the motion sensor data 112, 114) using one or more motion sensors (e.g., the one or more motion sensors 104). At 706, one or more feature points of the images are matched, and at 708, a 3D location for each of the one or more matched features is determined. Determining the one or more feature points and the 3D locations for each of the one or more matched feature points can be performed using any suitable technique, such as the bundle adjustment technique described above with reference to at least FIG. 1.

[0056] At 710, a 3D model of the object is generated using the one or more matched features (e.g., according to the bundle adjustment technique). In at least one

embodiment, the method 700 of FIG. 7 is performed using a processor of the mobile device 102, as described further with reference to at least FIG. 10.

[0057] Referring to FIG. 8, a flow diagram illustrating a particular embodiment of a method of computing a 3D model according to a silhouette-based reconstruction technique (e.g., a visual hull technique) based on motion sensor data is depicted and generally designated 800. The method 800 includes capturing images (e.g., the images 120, 122) of an object (e.g., the object 130) using a camera (e.g., the camera 108) of a mobile device (e.g., the mobile device 102), at 802.

[0058] At 804, motion sensor data (e.g., the motion sensor data 112, 114) is determined using one or more motion sensors (e.g., the one or more motion sensors 104) of the mobile device. A silhouette is generated of the object in each of the images, at 806. The method 800 further includes generating a silhouette-based model of the object, at 808, and generating a 3D model based on the silhouette-based model, at 810. A particular example of a silhouette-based reconstruction technique is described with reference to FIG 9.

[0059] Referring to FIG. 9, a flow diagram illustrating another particular embodiment of a method of computing a 3D model according to a silhouette-based reconstruction technique (e.g., a visual hull technique) based on motion sensor data is depicted and generally designated 900. The method 900 includes scanning by a mobile device (e.g., the mobile device 102) each pixel of a first view and projecting a ray  $R$  based on a view direction, at 902. At 904, the ray  $R$  is projected to one or more reference images and an interval  $I$  is found such that the interval  $I$  intersects a silhouette of each of the images.

[0060] At 906, motion sensor data (e.g., the motion sensor data 112, 114) indicating a change in position of the mobile device (e.g., from a first position to a second position) is determined via at least one sensor (e.g., via the one or more motion sensors 104). In at least one embodiment, the motion sensor data determined at 906 is used to project the ray  $R$  to the one or more reference images, at 902. The method 900 further includes projecting the interval  $I$  back into 3D space using a known camera projection  $P^i$  as a 3D interval  $L$ , at 908. The known camera projection  $P^i$  may be derived from the motion sensor data. The method 900 further includes selecting an innermost (e.g., shortest) intersection  $Q$  from the overlapping  $Ls$ , where  $Q$  is a final 3D interval having end points

corresponding to surface points of a 3D model to be constructed, at 910. In response to the number of intersections  $Q_s$  being less than a threshold  $N$ , another ray or view is added, at 914. In response to the number of  $Q_s$  satisfying the threshold  $N$ , the 3D model (e.g., the 3D model 116) is generated using the intersections  $Q_s$ . The method 800 of FIG. 8, the method 900 of FIG. 9, or a combination thereof, may be performed using a processor of the mobile device 102, as described further with reference to at least FIG. 11.

[0061] FIG. 10 is a diagram of another particular illustrative embodiment of the mobile device 102, where the mobile device 102 is configured to compute 3D models according to a feature point-based reconstruction technique (e.g., a bundle adjustment technique) based on motion sensor data. The mobile device 102 of FIG. 10 includes a processor, such as a processor 1010, coupled to a memory 1032, and to the one or more motion sensors 104. The one or more motion sensors 104 may also be coupled to the memory 1032.

[0062] The memory 1032 may be a computer-readable non-transitory medium that stores data (e.g., motion data of the mobile device 102), instructions, or a combination thereof. In a particular embodiment, the memory 1032 may include instructions 1054 that may be executable by the processor 1010 to cause the processor 1010 to perform one or more functions of the mobile device 102. For example, the instructions 1054 may include user applications, an operating system, or other executable instructions, or a combination thereof. The instructions 1054 may be executable by the processor 1010 to cause the processor 1010 to perform one or more operations described herein, such as to determine movement of the mobile device 102 based on motion sensor data (e.g., the motion sensor data 112, 114) determined by the one or more motion sensors 104. In at least one embodiment, the memory 1032 is a “main memory,” such as the main memory 310 described with reference to FIG. 3. Alternatively or in addition, the memory 1032 may include the one or more storage devices 106, the motion sensor data buffer 306, the video frame buffer 304, or a combination thereof.

[0063] The processor 1010 may include the 3D modeler 110. For example, the 3D modeler 110 may be a hardware component of the processor 1010, a set of instructions (e.g., instructions stored at the memory 1032) executable by the processor 1010, or a

combination thereof. In the particular embodiment of FIG. 10, the 3D modeler 110 includes a feature point generator 1092, which may determine feature points 1094 of captured images (e.g., the first image 120 and the second image 122 of FIG. 1) of an object in connection with a feature-point based reconstruction technique. The 3D modeler 110 may generate the 3D model 116 based on the feature points 1094.

**[0064]** FIG. 10 illustrates the camera 108 coupled to a camera controller 1090 that may provide the images 120, 122 and/or video data to be processed by the 3D modeler 110. The camera controller 1090 may be coupled to the processor 1010 and to a display controller 1026. The camera 108 may be a video camera, a non-video camera, a single camera, a multi-camera (e.g., a stereo camera), or a combination thereof.

**[0065]** FIG. 10 also shows that the display controller 1026 may be coupled to the processor 1010 and to the display 320. A coder/decoder (CODEC) 1034 (e.g., an audio and/or voice CODEC) may be coupled to the processor 1010. A speaker 1036 and a microphone 1038 may be coupled to the CODEC 1034. FIG. 10 also indicates that a wireless controller 1040 may be coupled to the processor 1010 and to a transceiver 1050 that is coupled to a wireless antenna 1042. In a particular embodiment, the processor 1010, the one or more motion sensors 104, the camera controller 1090, the display controller 1026, the memory 1032, the CODEC 1034, the wireless controller 1040, and the transceiver 1050 are included in a system-in-package or system-on-chip device 1022.

**[0066]** In a particular embodiment, an input device 1030 and a power supply 1044 are coupled to the system-on-chip device 1022. Moreover, in a particular embodiment, and as illustrated in FIG. 10, the display 320, the camera 108, the input device 1030, the speaker 1036, the microphone 1038, the wireless antenna 1042, and the power supply 1044 are external to the system-on-chip device 1022. However, each of the display 320, the input device 1030, the speaker 1036, the microphone 1038, the wireless antenna 1042, and the power supply 1044 can be coupled to a component of the system-on-chip device 1022, such as an interface or a controller. Although the one or more motion sensors 104 are illustrated as included in the system-on-chip device 1022, in other embodiments the one or more motion sensors 104 may be external to the system-on-chip device 1022 and coupled to the system-on-chip device 1022.

[0067] FIG. 11 is a diagram of another particular illustrative embodiment of the mobile device 102, where the mobile device 102 is configured to compute 3D models according to a silhouette-based reconstruction technique (e.g., a visual hull technique) based on motion sensor data. Certain components and features of the mobile device 102 of FIG. 11 may be described generally with reference to the mobile device 102 of FIG. 10, such as the processor 1010, the one or more motion sensors 104, the camera controller 1090, the display controller 1026, the memory 1032, the instructions 1054, the CODEC 1034, the wireless controller 1040, the transceiver 1050, the system-in-package or system-on-chip device 1022, the wireless antenna 1042, the input device 1030, the display 320, the camera 108, the speaker 1036, the microphone 1038, and the power supply 1044. However, the 3D modeler 110 of FIG. 11 includes a silhouette generator 1192 to generate a silhouette 1194, for example in connection with a visual hull technique. The silhouette 1194 may be used by the 3D modeler 110 to generate the 3D model 116.

[0068] Although the present disclosure is described in connection with a mobile device (e.g., the mobile device 102) for ease of illustration, it should be appreciated that the functionalities described herein may be implemented using other devices. For example, the processor 1010 and the memory 1032 of FIGS. 10 and 11 may be integrated into a multimedia player, an entertainment unit, a navigation device, a personal digital assistant (PDA), a fixed location data unit, or a computer (e.g., a tablet computer, a laptop computer, a desktop computer, etc), a media device, another device configured to wirelessly communicate data, or a combination thereof. Further, although for convenience of illustration various techniques (e.g., reconstruction techniques, such as feature point-based and silhouette-based reconstruction techniques) have been described separately, one or more such techniques (or elements thereof) can be combined.

[0069] In conjunction with the described embodiments, an apparatus is disclosed that includes means for determining, based on data received from at least one means for sensing motion, a movement of a mobile device from a first position to a second position. For example, the means for determining may include the 3D modeler 110 of FIGS. 1-4, 10, and 11, the processor 1010 of FIGS. 10 and 11, one or more other devices configured to determine a movement of a mobile device, or any combination

thereof. The means for sensing motion may include the one or more motion sensors 104.

[0070] The apparatus further includes means for computing a three-dimensional (3D) model of an object based on a first image of the object corresponding to a first view of the object from the first position of the mobile device, based further on a second image of the object corresponding to a second view of the object from the second position of the mobile device, and based further on movement of the mobile device. For example, the means for computing the 3D model may include the 3D modeler 110 of FIGS. 1-4, 10, and 11, the processor 1010 of FIGS. 10 and 11, one or more other devices configured to compute a 3D model, or any combination thereof.

[0071] Those of skill would further appreciate that the various illustrative logical blocks, configurations, modules, circuits, and algorithm steps described in connection with the embodiments disclosed herein may be implemented as electronic hardware, processor-executable instructions stored at a computer-readable tangible medium, or combinations of both. Various illustrative components, blocks, configurations, modules, circuits, and steps have been described above generally in terms of their functionality. Whether such functionality is implemented as hardware or software depends upon the particular application and design constraints imposed on the overall system. Skilled artisans may implement the described functionality in varying ways for each particular application, but such implementation decisions should not be interpreted as causing a departure from the scope of the present disclosure.

[0072] The steps of a method or algorithm described in connection with the embodiments disclosed herein may be embodied directly in hardware, in a software module executed by a processor, or in a combination of the two. A software module may reside in random access memory (RAM), flash memory, read-only memory (ROM), programmable read-only memory (PROM), erasable programmable read-only memory (EPROM), electrically erasable programmable read-only memory (EEPROM), registers, hard disk, a removable disk, a compact disc read-only memory (CD-ROM), or any other form of non-transitory storage medium. An exemplary storage medium is coupled to the processor such that the processor can read information from, and write information to, the storage medium. In the alternative, the storage medium may be

integral to the processor. The processor and the storage medium may reside in an application-specific integrated circuit (ASIC). The ASIC may reside in a computing device or a user terminal (e.g., a mobile phone or a PDA). In the alternative, the processor and the storage medium may reside as discrete components in a computing device or user terminal.

[0073] The previous description of the disclosed embodiments is provided to enable a person skilled in the art to make or use the disclosed embodiments. Various modifications to these embodiments will be readily apparent to those skilled in the art, and the principles defined herein may be applied to other embodiments without departing from the scope of the disclosure. Thus, the present disclosure is not intended to be limited to the embodiments disclosed herein but is to be accorded the widest scope possible consistent with the principles and novel features as defined by the following claims.

**WHAT IS CLAIMED IS:**

1. A method comprising:  
capturing, at a camera of a mobile device, a first image of an object while the mobile device is at a first position;  
capturing a second image of the object while the mobile device is at a second position;  
determining, based on data output by at least one motion sensor of the mobile device, a movement of the mobile device from the first position to the second position; and  
computing a three-dimensional (3D) model of the object based on the first image, the second image, and the determined movement of the mobile device.
2. The method of claim 1, wherein the first image and the second image are video frames associated with a stored video file, and further comprising periodically accessing the stored video file to access the video frames.
3. The method of claim 1, wherein the first image and the second image are video frames associated with a video frame sequence, and further comprising accessing the video frames prior to storing the video frame sequence as a video file.
4. The method of claim 1, wherein the first image and the second image are snapshot images.
5. The method of claim 1, wherein computing the 3D model of the object is further based on one or more silhouettes of the object.
6. The method of claim 1, wherein the data output by the at least one motion sensor includes rotational motion associated with the mobile device, translational motion associated with the mobile device, or a combination thereof.

7. The method of claim 1, wherein the determined movement of the mobile device includes a rotation of the mobile device, a translation of the mobile device, or a combination thereof.

8. The method of claim 1, wherein computing the 3D model of the object further comprises determining feature points of the object based on a function of a first estimated reprojected image point, a second estimated reprojected image point, and the data output by the at least one motion sensor.

9. The method of claim 1, wherein computing the three-dimensional model of the object further comprises:

determining a first estimated reprojected image point of the object based on a product of a first estimate of the first position and one or more feature points of the object;

determining a second estimated reprojected image point of the object based on a product of a second estimate of the second position and the one or more feature points of the object;

determining a square of a first Euclidean distance between the first estimated reprojected image point and coordinates of the object in the first image;

determining a square of a second Euclidean distance between the second

estimated reprojected image point and coordinates of the object in the second image; and

modifying an estimate of the 3D model to minimize a sum of the square of the first Euclidean distance and the square of the second Euclidean distance, wherein the second estimate of the second position is generated based on the first estimate of the first position and the data output by the at least one motion sensor.

10. The method of claim 9, wherein modifying the estimate of the 3D model is performed according to a bundle adjustment technique to determine a maximum likelihood solution of:

$$\min_{P^i, X_j} \sum_{i,j} d(P^i X_j, x_j^i),$$

wherein  $d$  is associated with a distance operation,  
wherein  $P^i$  is a projection matrix of the camera,  
wherein  $X_j$  is a j-th feature point of the object, and  
wherein  $x_j^i$  are coordinates of the j-th feature point in an i-th image.

11. An apparatus comprising:

a camera configured to capture a first image of an object while at a first position  
and to capture a second image of the object while at a second position;

at least one motion sensor;

a processor; and

a memory storing instructions executable by the processor to:

determine, based on data output by the at least one motion sensor, a  
movement of the camera from the first position to the second  
position; and

compute a three-dimensional (3D) model of the object based on the first  
image, the second image, and the determined movement of the  
camera.

12. The apparatus of claim 11, wherein the processor is integrated into a mobile device, the camera, a multimedia player, an entertainment unit, a navigation device, a personal digital assistant (PDA), a portable computer, or any combination thereof.

13. The apparatus of claim 11, wherein the data output by the at least one motion sensor includes rotational information associated with the camera, translational information associated with the camera, or a combination thereof.

14. The apparatus of claim 11, wherein the determined movement of the camera includes a rotation of the camera, a translation of the camera, or a combination thereof.

15. The apparatus of claim 11, wherein the instructions are further executable by the processor to determine feature points of the object based on a function of a first estimated reprojected image point, a second estimated reprojected image point, and the data output by the at least one motion sensor.

16. The apparatus of claim 11, wherein the instructions are further executable by the processor to:

determine a first estimated reprojected image point of the object based on a product of a first estimate of the first position and one or more feature points of the object;

determine a second estimated reprojected image point of the object based on a product of a second estimate of the second position and the one or more feature points of the object;

determine a square of a first Euclidean distance between the first estimated reprojected image point and coordinates of the object in the first image;

determine a square of a second Euclidean distance between the second estimated reprojected image point and coordinates of the object in the second image; and

modify an estimate of the 3D model to minimize a sum of the square of the first Euclidean distance and the square of the second Euclidean distance, wherein the second estimate of the second position is generated based on the first estimate of the first position and the data output by the at least one motion sensor.

17. The apparatus of claim 16, wherein modifying the estimate of the 3D model is performed according to a bundle adjustment technique to determine a maximum likelihood solution of:

$$\min_{P^i, X_j} \sum_{i,j} d(P^i X_j, x_j^i),$$

wherein  $d$  is associated with a distance operation,  
 wherein  $P^i$  is a projection matrix of the camera,  
 wherein  $X_j$  is a  $j$ -th feature point of the object, and  
 wherein  $x_j^i$  are coordinates of the  $j$ -th feature point in an  $i$ -th image.

18. A method comprising:  
determining, based on data received from at least one motion sensor, a  
movement of a mobile device from a first position to a second position;  
and  
computing a three-dimensional (3D) model of an object based on a first image of  
the object corresponding to a first view of the object from the first  
position of the mobile device, based further on a second image of the  
object corresponding to a second view of the object from the second  
position of the mobile device, and based further on the movement of the  
mobile device.

19. An apparatus comprising:  
means for determining, based on data received from at least one means for  
sensing motion, a movement of a mobile device from a first position to a  
second position; and  
means for computing a three-dimensional (3D) model of an object based on a  
first image of the object corresponding to a first view of the object from  
the first position of the mobile device, based further on a second image  
of the object corresponding to a second view of the object from the  
second position of the mobile device, and based further on the movement  
of the mobile device.

20. The apparatus of claim 19, further comprising means for capturing the first  
image and for capturing the second image.

21. The apparatus of claim 19, further comprising means for determining  
feature points of the object based on a function of a first estimated reprojected image  
point, a second estimated reprojected image point, and the data received from the at  
least one means for sensing motion.

22. A computer-readable non-transitory medium comprising instructions that, when executed by a processor, cause the processor to:

determine, based on data received from at least one motion sensor, a movement of a mobile device from a first position to a second position; and compute a three-dimensional (3D) model of an object based on a first image of the object corresponding to a first view of the object from the first position of the mobile device, based further on a second image of the object corresponding to a second view of the object from the second position of the mobile device, and based further on the movement of the mobile device.

23. The non-transitory computer-readable medium of claim 22, wherein the data includes rotational information associated with the mobile device, translational information associated with the mobile device, or a combination thereof.

24. The non-transitory computer-readable medium of claim 22, further comprising instructions executable by the processor to determine feature points of the object based on a function of a first estimated reprojected image point, a second estimated reprojected image point, and the data.

25. The non-transitory computer-readable medium of claim 22, further comprising instructions executable by the processor to generate one or more silhouettes associated with the object, wherein the 3D model of the object is based further on the one or more silhouettes.

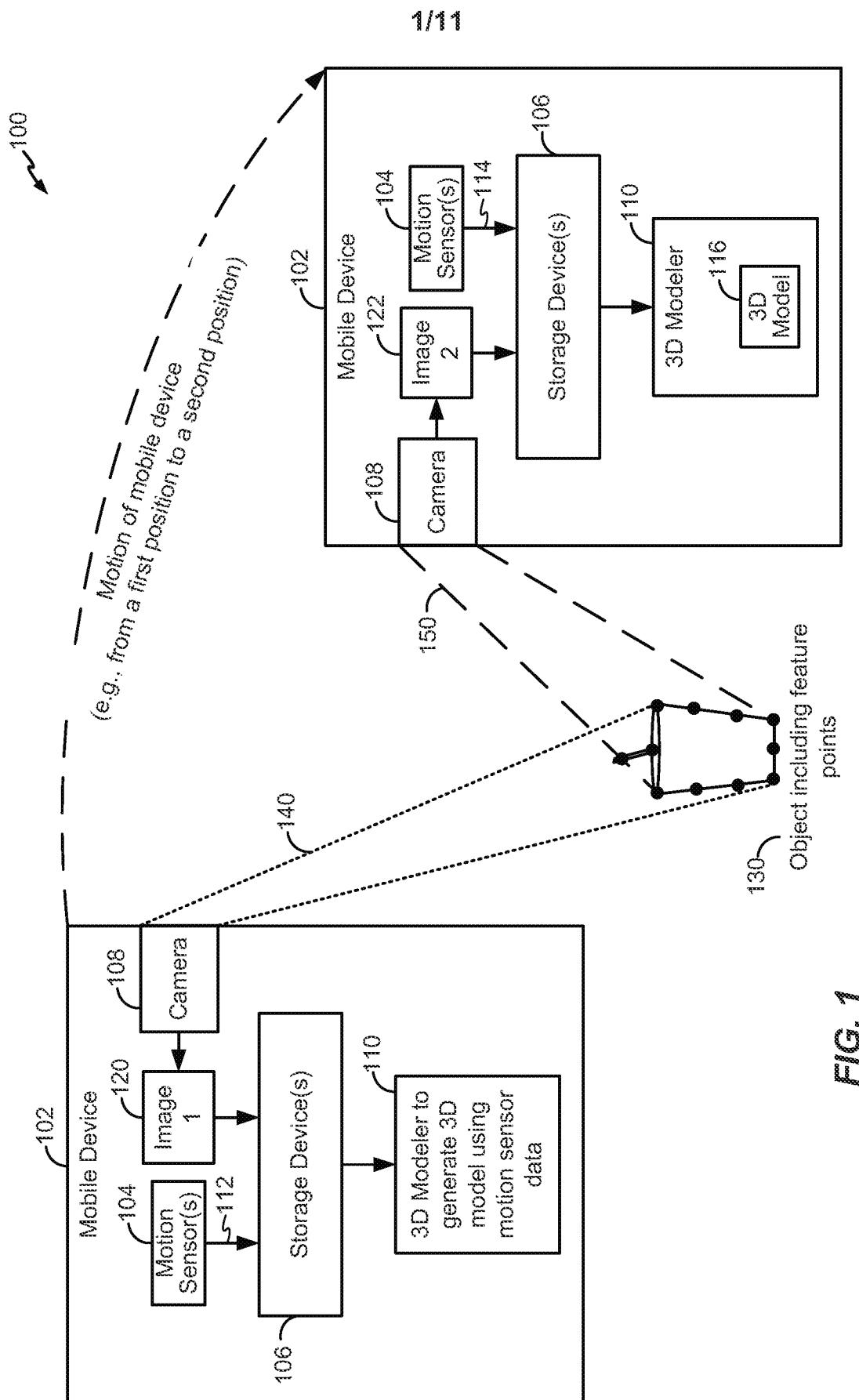


FIG. 1

2/11

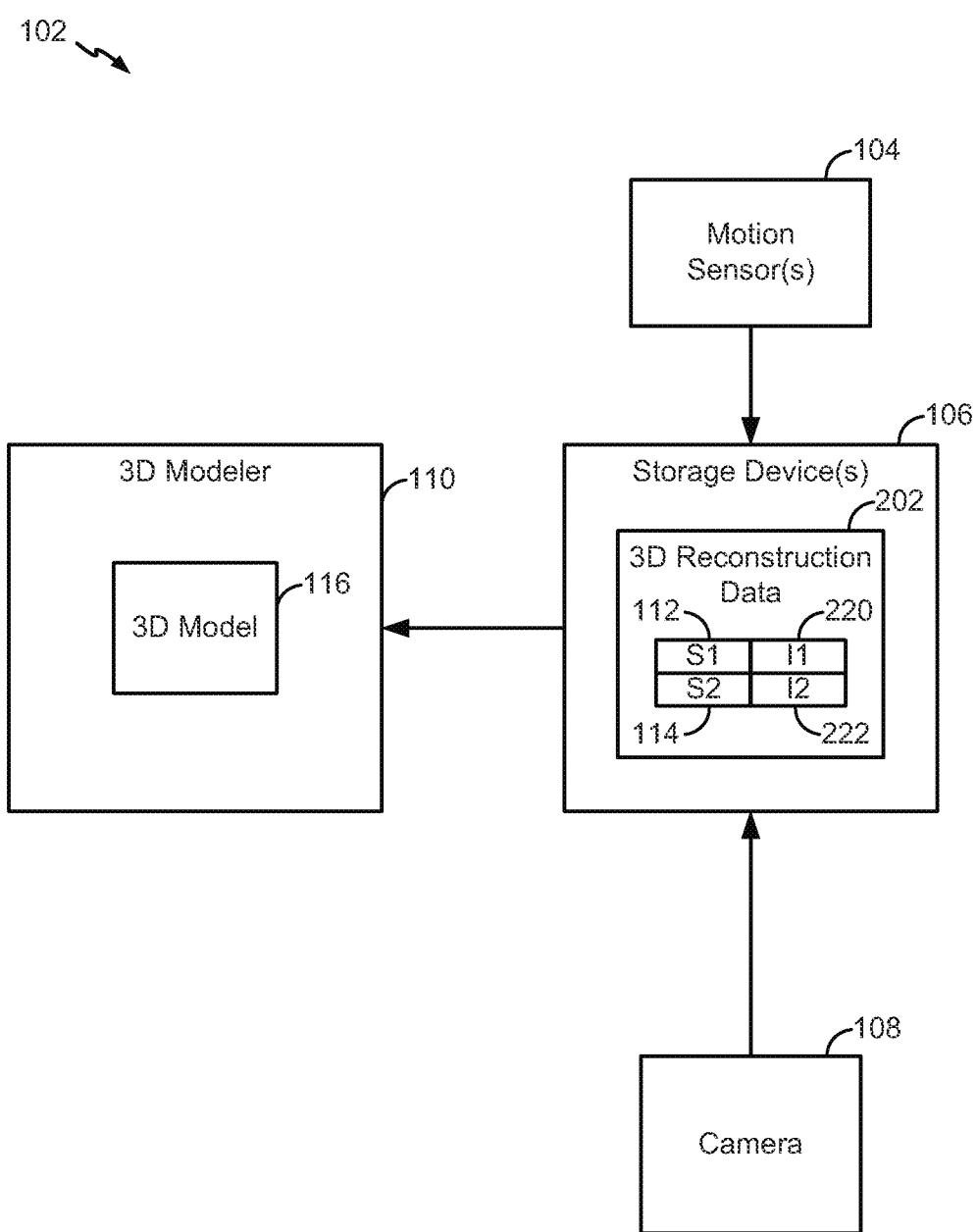


FIG. 2

3/11

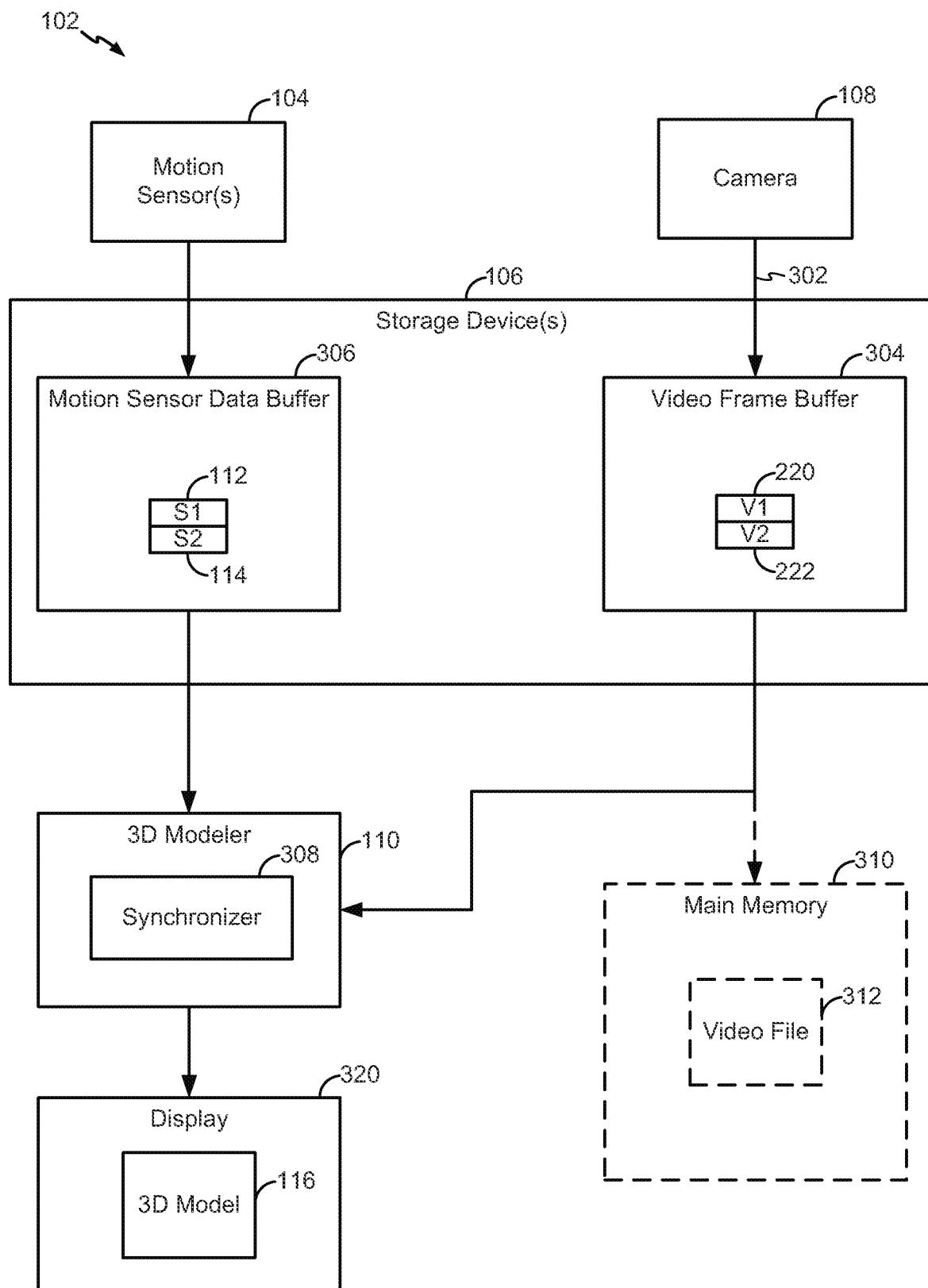


FIG. 3

4/11

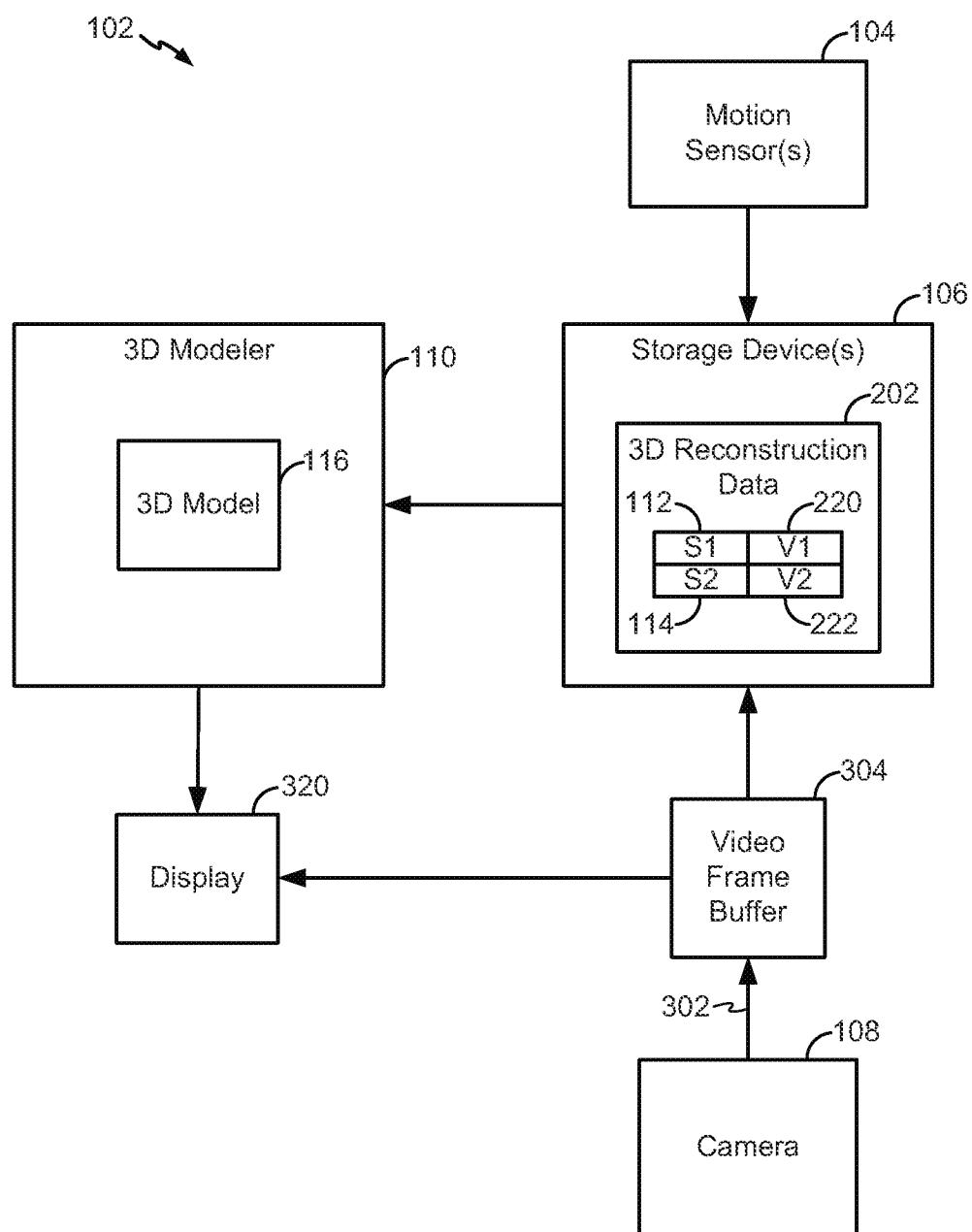


FIG. 4

5/11

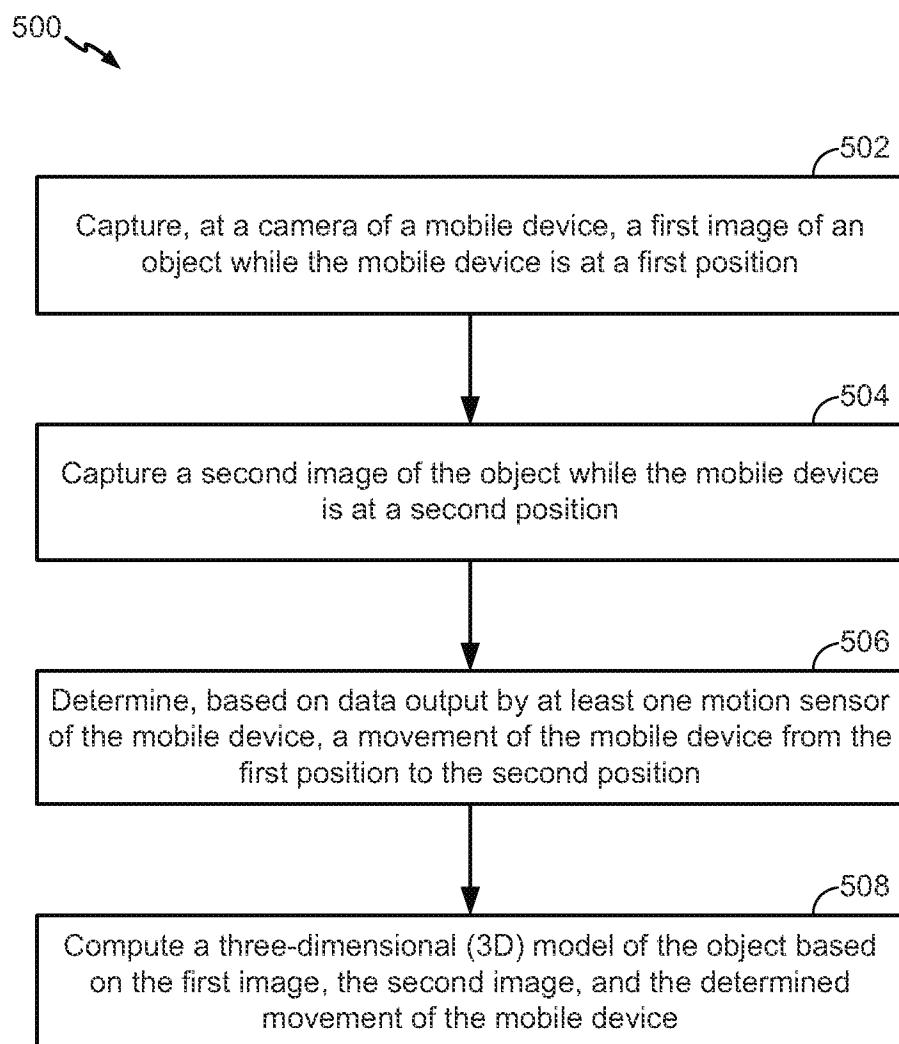
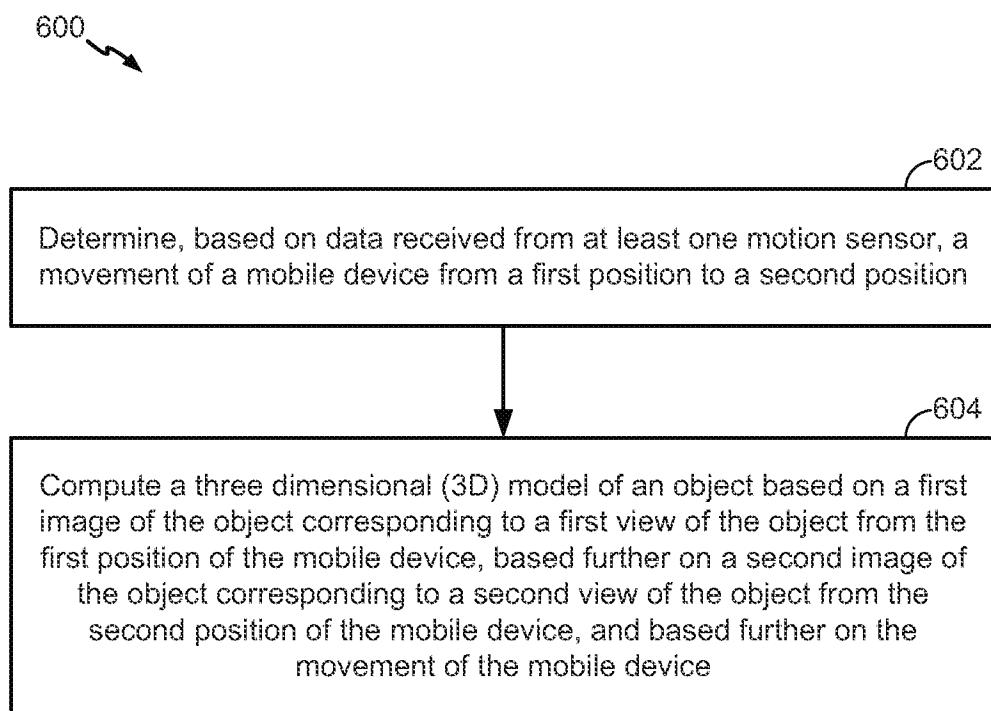


FIG. 5

6/11

**FIG. 6**

7/11

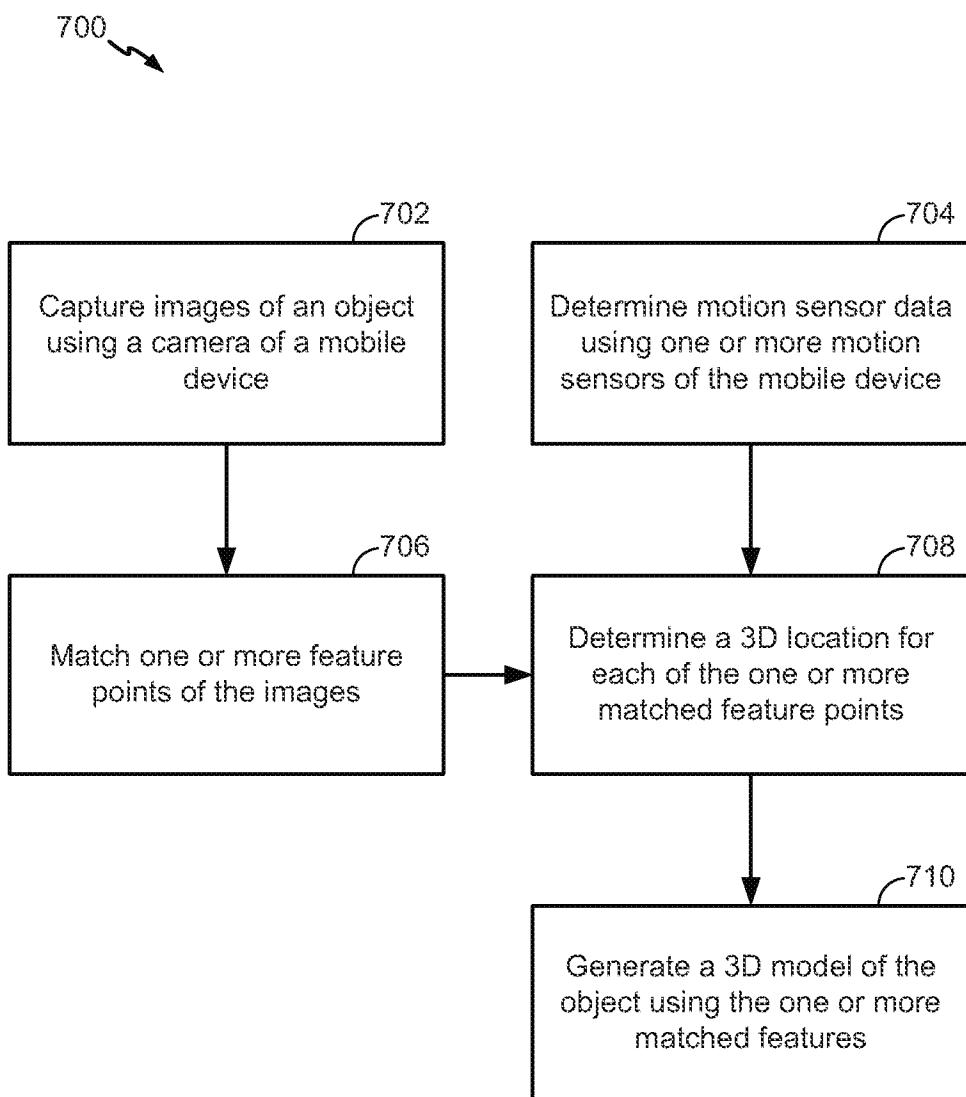


FIG. 7

8/11

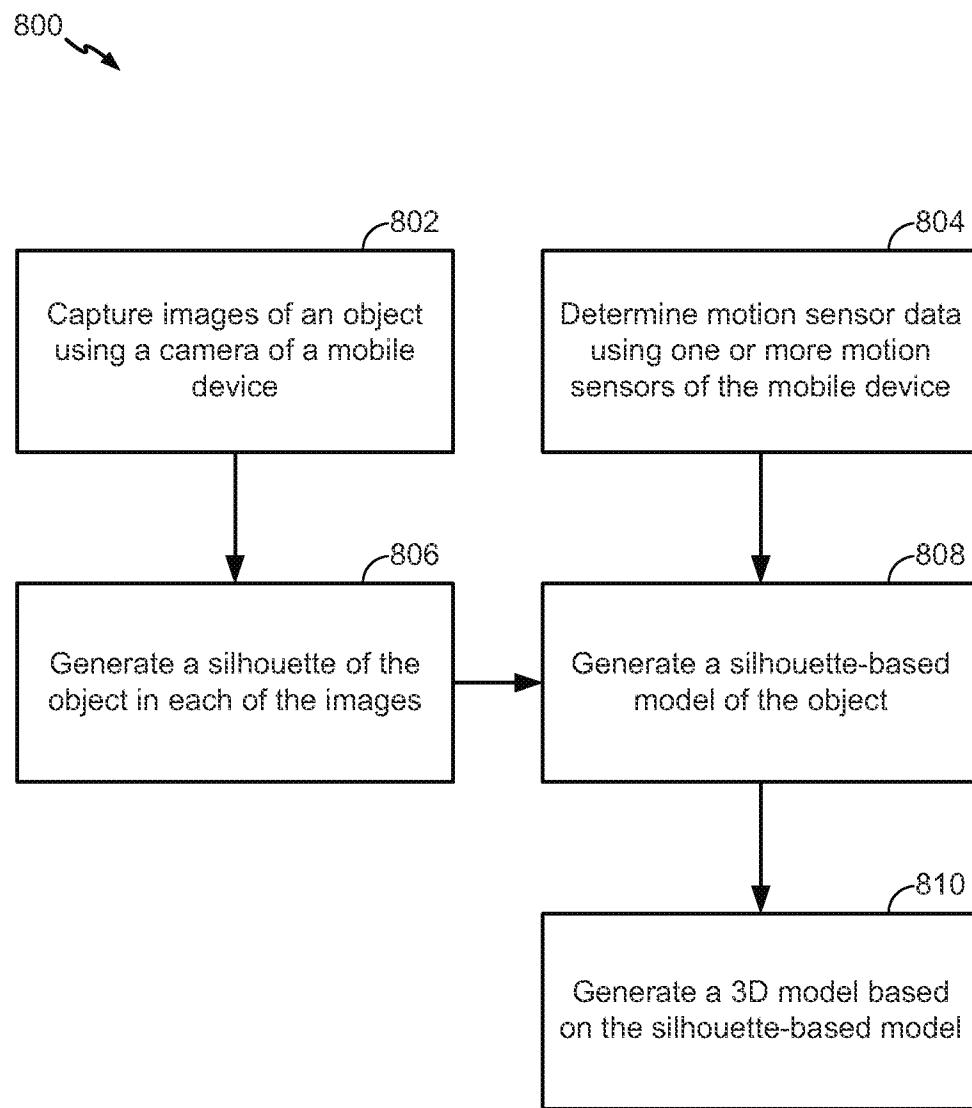


FIG. 8

9/11

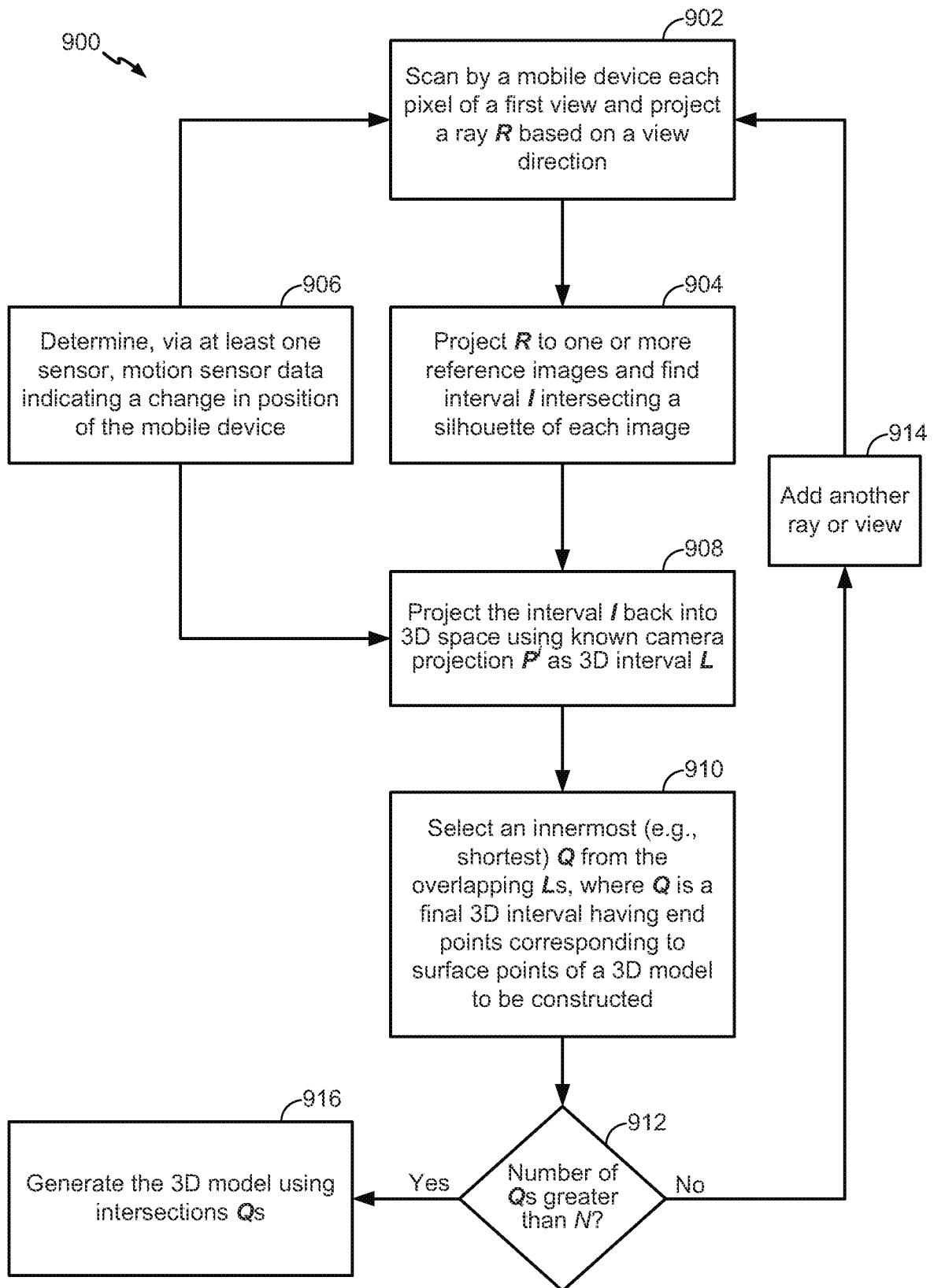


FIG. 9

10/11

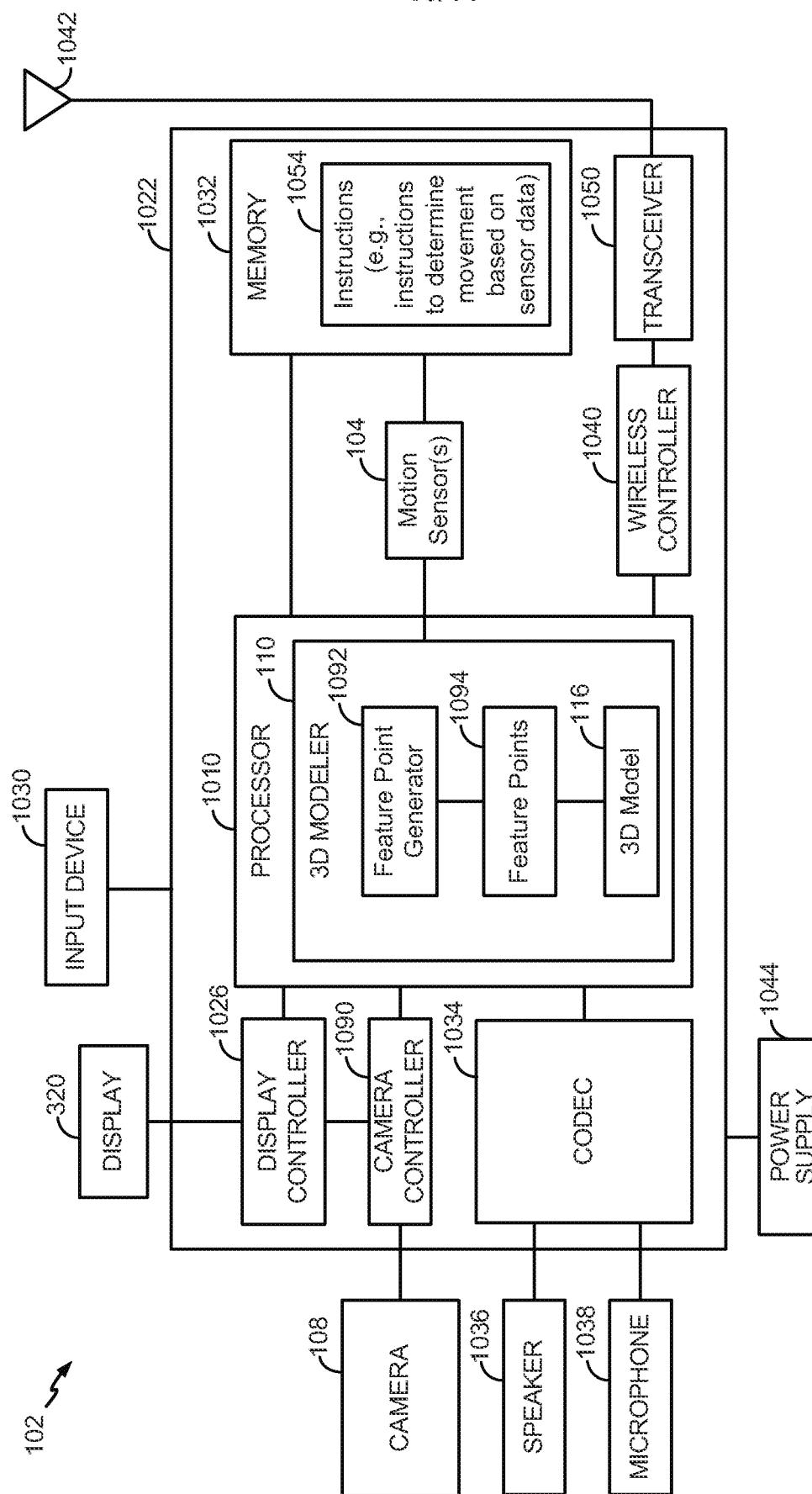


FIG. 10

11/11

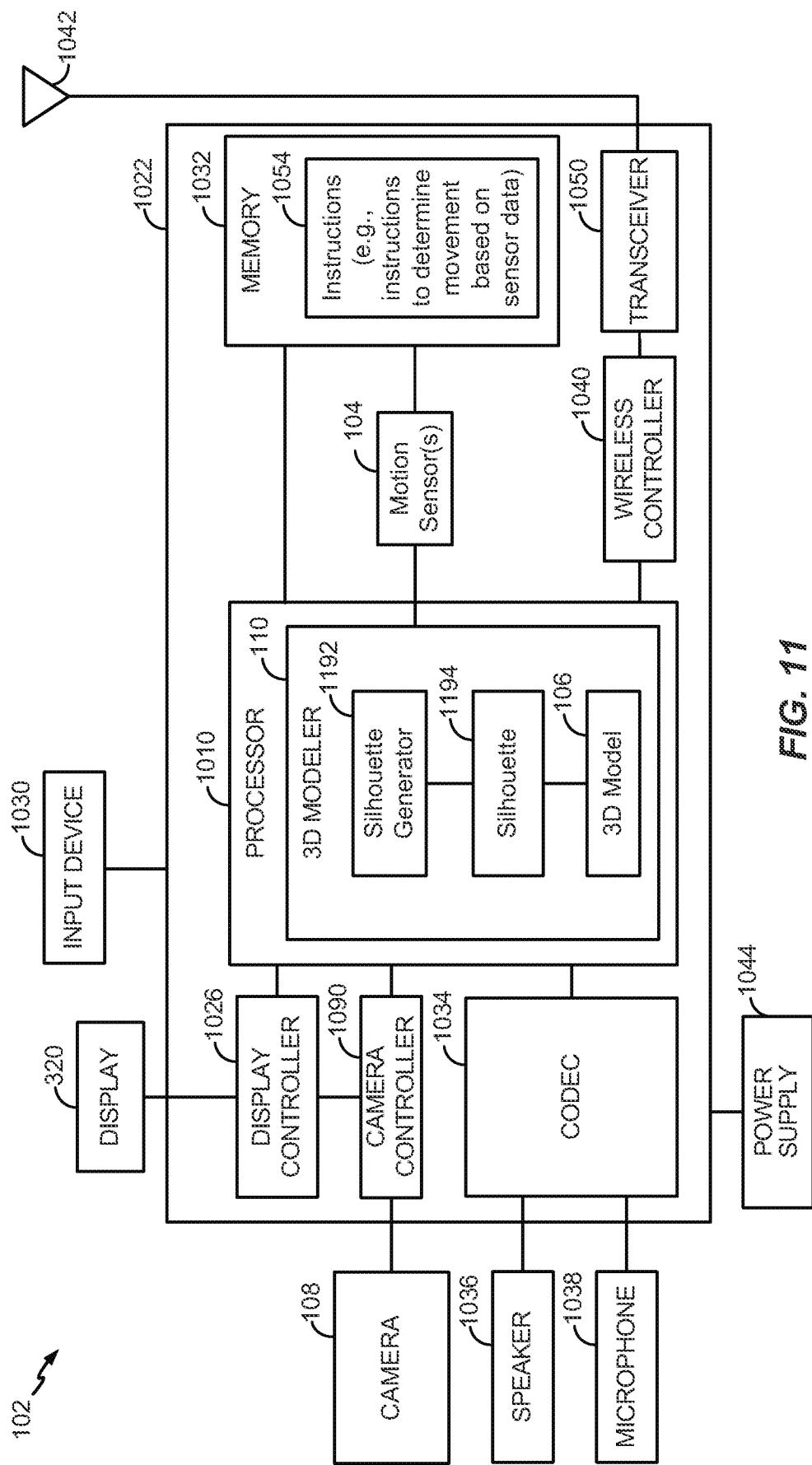


FIG. 11

# INTERNATIONAL SEARCH REPORT

International application No  
PCT/US2012/069088

**A. CLASSIFICATION OF SUBJECT MATTER**  
INV. G06T7/00  
ADD.

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

**B. FIELDS SEARCHED**

Minimum documentation searched (classification system followed by classification symbols)  
G06T G06K

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

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

EPO-Internal, WPI Data, INSPEC

**C. DOCUMENTS CONSIDERED TO BE RELEVANT**

Category*	Citation of document, with indication, where appropriate, of the relevant passages	Relevant to claim No.
X	US 2010/316282 A1 (HOPE CLINTON B [US]) 16 December 2010 (2010-12-16)	1-7, 11-14, 18-20, 22,23,25
Y	abstract figures 3,5 paragraphs [0015], [0019], [0020], [0025], [0026], [0031] ----- -/-	8-10, 15-17, 21,24

Further documents are listed in the continuation of Box C.

See patent family annex.

\* Special categories of cited documents :

- "A" document defining the general state of the art which is not considered to be of particular relevance
- "E" earlier application or patent but published on or after the international filing date
- "L" document which may throw doubts on priority claim(s) or which is cited to establish the publication date of another citation or other special reason (as specified)
- "O" document referring to an oral disclosure, use, exhibition or other means
- "P" document published prior to the international filing date but later than the priority date claimed

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

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

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

"&" document member of the same patent family

Date of the actual completion of the international search	Date of mailing of the international search report
8 March 2013	19/03/2013
Name and mailing address of the ISA/ European Patent Office, P.B. 5818 Patentlaan 2 NL - 2280 HV Rijswijk Tel. (+31-70) 340-2040, Fax: (+31-70) 340-3016	Authorized officer  Scholz, Volker

## INTERNATIONAL SEARCH REPORT

International application No
PCT/US2012/069088

C(Continuation). DOCUMENTS CONSIDERED TO BE RELEVANT

Category*	Citation of document, with indication, where appropriate, of the relevant passages	Relevant to claim No.
A	<p>BASTIAN J ET AL: "Interactive modelling for AR applications", MIXED AND AUGMENTED REALITY (ISMAR), 2010 9TH IEEE INTERNATIONAL SYMPOSIUM ON, IEEE, PISCATAWAY, NJ, USA, 13 October 2010 (2010-10-13), pages 199-205, XP032291046, DOI: 10.1109/ISMAR.2010.5643570 ISBN: 978-1-4244-9343-2</p> <p>abstract figure 3 section 2</p> <p>-----</p>	5,25
X	<p>R. ROBERT CLARK ET AL: "&lt;title&gt;3D environment capture from monocular video and inertial data&lt;/title&gt;", PROCEEDINGS OF SPIE, vol. 6056, 2 February 2006 (2006-02-02), page 60560G, XP055054107, ISSN: 0277-786X, DOI: 10.1117/12.650805</p> <p>abstract sections 1-3</p> <p>-----</p>	1,18,22
Y	<p>&amp; Manolis I.A. Lourakis ET AL: "The design and implementation of a generic sparse bundle adjustment software package based on the levenberg-marquardt algorithm", Technical Report FORTH-ICS TR-340, 1 August 2004 (2004-08-01), pages 1-23, XP055054108, Retrieved from the Internet: URL:<a href="http://www.ics.forth.gr/~lourakis/publ/2004_tr340.pdf">http://www.ics.forth.gr/~lourakis/publ/2004_tr340.pdf</a> [retrieved on 2013-02-21]</p> <p>section 3</p> <p>-----</p>	8-10, 15-17, 21,24

## INTERNATIONAL SEARCH REPORT

#### Information on patent family members

International application No  
PCT/US2012/069088

Patent document cited in search report	Publication date	Patent family member(s)			Publication date
US 2010316282	A1 16-12-2010	CN 102012625 A	JP 2011027718 A	KR 20100135196 A	13-04-2011 10-02-2011 24-12-2010
		TW 201101812 A	US 2010316282 A1		01-01-2011 16-12-2010