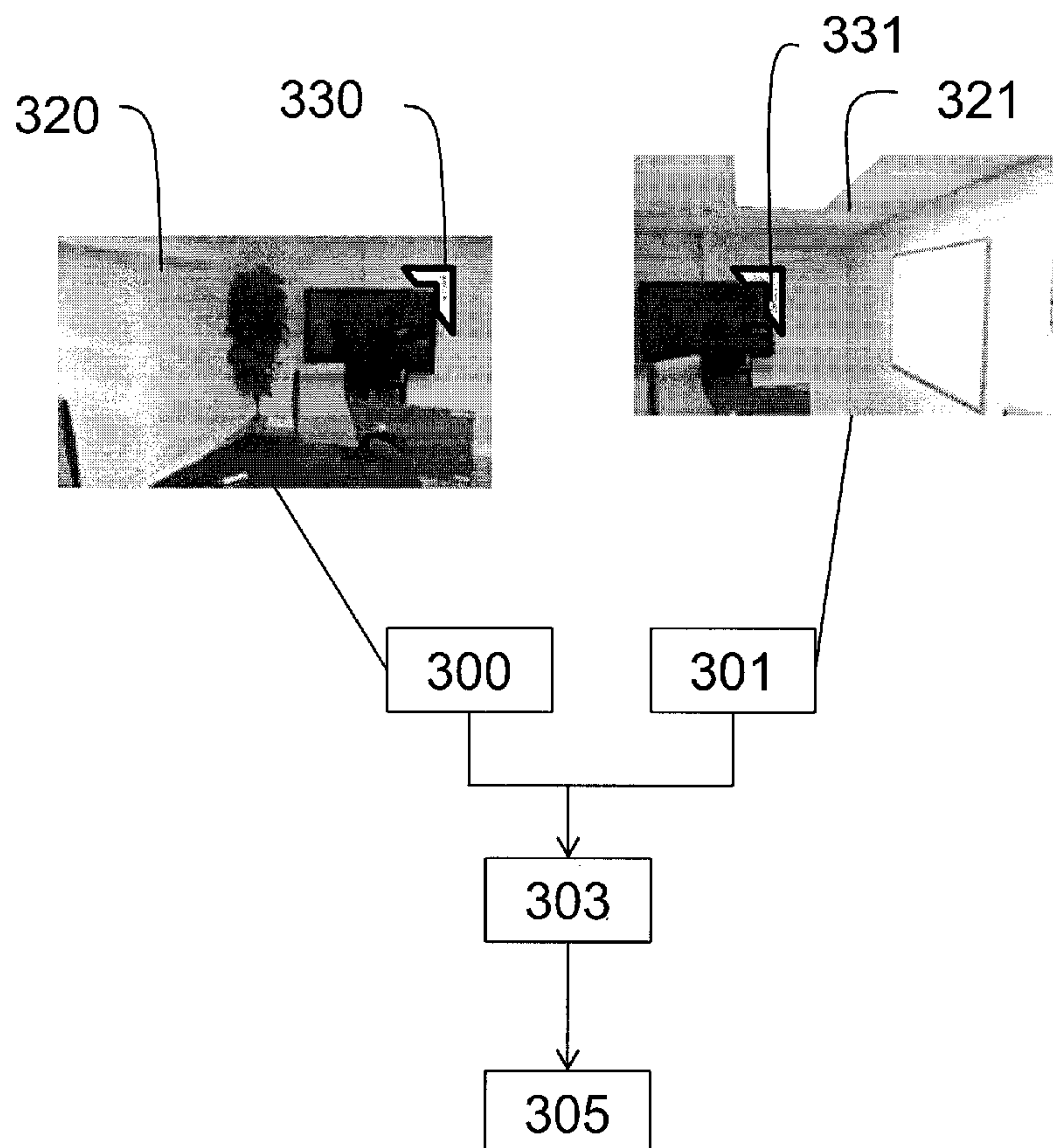




(86) **Date de dépôt PCT/PCT Filing Date:** 2014/10/03
(87) **Date publication PCT/PCT Publication Date:** 2015/04/09
(45) **Date de délivrance/Issue Date:** 2015/08/18
(85) **Entrée phase nationale/National Entry:** 2015/04/21
(86) **N° demande PCT/PCT Application No.:** CA 2014/050961
(87) **N° publication PCT/PCT Publication No.:** 2015/048906
(30) **Priorité/Priority:** 2013/10/03 (US61/886,437)

(51) **Cl.Int./Int.Cl.** *G01B 11/00* (2006.01),
G02B 27/01 (2006.01), *G06F 19/00* (2011.01),
H04N 13/00 (2006.01)
(72) **Inventeurs/Inventors:**
BALACHANDRESWARAN, DHANUSHAN, CA;
NJENGA, KIBAYA MUNGAI, CA;
ZHANG, JIAN, CA
(73) **Propriétaire/Owner:**
SULON TECHNOLOGIES INC., CA
(74) **Agent:** Bhole, Anil

(54) **Titre : SYSTEME A REALITE AUGMENTEE ET PROCEDE DE POSITIONNEMENT ET DE CARTOGRAPHIE**
(54) **Title: AUGMENTED REALITY SYSTEM AND METHOD FOR POSITIONING AND MAPPING**



(57) **Abrégé/Abstract:**

An augmented reality and virtual reality head mounted display is described. The head mounted display comprises a camera array in communication with a processor to map the physical environment for rendering an augmented reality of the physical environment.



ABSTRACT

An augmented reality and virtual reality head mounted display is described. The head mounted display comprises a camera array in communication with a processor to map the physical environment for rendering an augmented reality of the physical environment.

1 AUGMENTED REALITY SYSTEM AND METHOD FOR POSITIONING AND MAPPING

2 TECHNICAL FIELD

3 [0001] The following relates generally to systems and methods for augmented and virtual reality
4 environments, and more specifically to systems and methods for mapping a virtual or
5 augmented environment based on a physical environment, and displaying the virtual or
6 augmented environment on a head mounted device.

7 BACKGROUND

8 [0002] The range of applications for augmented reality (AR) and virtual reality (VR) visualization
9 has increased with the advent of wearable technologies and 3-dimensional (3D) rendering
10 techniques. AR and VR exist on a continuum of mixed reality visualization.

11 SUMMARY

12 [0003] In embodiments, a method is described for mapping a physical environment surrounding
13 a user wearing a wearable display for augmented reality. The method comprises: (i) capturing,
14 by at least one depth camera disposed upon the user, depth information for the physical
15 environment; (ii) by a processor, obtaining the depth information, determining the orientation of
16 the at least one depth camera relative to the wearable display, and assigning coordinates for the
17 depth information in a map of the physical environment based on the orientation of the at least
18 one depth camera.

19 [0004] In further embodiments, a system is described for mapping a physical environment
20 surrounding a user wearing a wearable display for augmented reality. The system comprises: (i)
21 at least one depth camera disposed upon the user, to capture depth information for the physical
22 environment; and (ii) at least one processor in communication with the at least one depth
23 camera, to obtain the depth information from the at least one depth camera, determine the
24 orientation of the at least one depth camera relative to the wearable display, and assign
25 coordinates for the depth information in a map of the physical information based on the
26 orientation of the at least one depth camera.

27 [0005] In still further embodiments, a system is described for displaying a rendered image
28 stream in combination with a physical image stream of a region of a physical environment
29 captured in the field of view of at least one image camera disposed upon a user wearing a
30 wearable display for augmented reality. The system comprises a processor configured to: (i)

1 obtain a map of the physical environment; (ii) determine the orientation and location of the
2 wearable display within the physical environment; (iii) determine, from the orientation and
3 location of the wearable display, the region of the physical environment captured in the field of
4 view of the at least one image camera; (iv) determine a region of the map corresponding to the
5 captured region of the physical environment; and (iv) generate rendered stream comprising
6 augmented reality for the corresponding region of the map.

7 [0006] In yet further embodiments, a method is described for displaying a rendered image
8 stream in combination with a physical image stream of a region of a physical environment
9 captured in the field of view of at least one image camera disposed upon a user wearing a
10 wearable display for augmented reality. The method comprises, by a processor: (i) obtaining a
11 map of the physical environment; (ii) determining the orientation and location of the wearable
12 display within the physical environment; (iii) determining, from the orientation and location of the
13 wearable display, the region of the physical environment captured in the field of view of the at
14 least one image camera; (iv) determining a region of the map corresponding to the captured
15 region of the physical environment; and (v) generating rendered stream comprising augmented
16 reality for the corresponding region of the map.

17 DESCRIPTION OF THE DRAWINGS

18 [0007] A greater understanding of the embodiments will be had with reference to the Figures, in
19 which:

20 [0008] Fig. 1 illustrates an embodiment of a head mounted display (HMD) device;

21 [0009] Fig. 2A illustrates an embodiment of an HMD having a single depth camera;

22 [0010] Fig. 2B illustrates an embodiment of an HMD having multiple depth cameras;

23 [0011] Fig. 3 is a flowchart illustrating a method for mapping a physical environment using a
24 depth camera;

25 [0012] Fig. 4 is a flowchart illustrating another method for mapping a physical environment
26 using a depth camera and an orientation detection system;

27 [0013] Fig. 5 is a flowchart illustrating a method for mapping a physical environment using
28 multiple depth cameras;

1 [0014] Fig. 6 is a flowchart illustrating a method for mapping a physical environment using at
2 least one depth camera and at least one imaging camera;

3 [0015] Fig. 7 is a flowchart illustrating a method for determining the location and orientation of
4 an HMD in a physical environment using at least one depth camera and/or at least one imaging
5 camera;

6 [0016] Fig. 8 is a flowchart illustrating a method for generating a rendered image stream of a
7 physical environment based on the position and orientation of an HMD within the physical
8 environment; and

9 [0017] Fig. 9 is a flowchart illustrating a method of displaying an augmented reality of a physical
10 environment by simultaneously displaying a physical image stream of the physical environment
11 and a rendered image stream.

12 DETAILED DESCRIPTION

13 [0018] It will be appreciated that for simplicity and clarity of illustration, where considered
14 appropriate, reference numerals may be repeated among the figures to indicate corresponding
15 or analogous elements. In addition, numerous specific details are set forth in order to provide a
16 thorough understanding of the embodiments described herein. However, it will be understood by
17 those of ordinary skill in the art that the embodiments described herein may be practiced without
18 these specific details. In other instances, well-known methods, procedures and components
19 have not been described in detail so as not to obscure the embodiments described herein. Also,
20 the description is not to be considered as limiting the scope of the embodiments described
21 herein.

22 [0019] It will also be appreciated that any module, unit, component, server, computer, terminal
23 or device exemplified herein that executes instructions may include or otherwise have access to
24 computer readable media such as storage media, computer storage media, or data storage
25 devices (removable and/or non-removable) such as, for example, magnetic disks, optical disks,
26 or tape. Computer storage media may include volatile and non-volatile, removable and non-
27 removable media implemented in any method or technology for storage of information, such as
28 computer readable instructions, data structures, program modules, or other data. Examples of
29 computer storage media include RAM, ROM, EEPROM, flash memory or other memory
30 technology, CD-ROM, digital versatile disks (DVD) or other optical storage, magnetic cassettes,
31 magnetic tape, magnetic disk storage or other magnetic storage devices, or any other medium

1 which can be used to store the desired information and which can be accessed by an
2 application, module, or both. Any such computer storage media may be part of the device or
3 accessible or connectable thereto. Any application or module herein described may be
4 implemented using computer readable/executable instructions that may be stored or otherwise
5 held by such computer readable media and executed by the one or more processors.

6 [0020] The present disclosure is directed to systems and methods for augmented reality (AR).
7 However, the term "AR" as used herein may encompass several meanings. In the present
8 disclosure, AR includes: the interaction by a user with real physical objects and structures
9 along with virtual objects and structures overlaid thereon; and the interaction by a user with a
10 fully virtual set of objects and structures that are generated to include renderings of physical
11 objects and structures and that may comply with scaled versions of physical environments to
12 which virtual objects and structures are applied, which may alternatively be referred to as an
13 "enhanced virtual reality". Further, the virtual objects and structures could be dispensed with
14 altogether, and the AR system may display to the user a version of the physical environment
15 which solely comprises an image stream of the physical environment. Finally, a skilled reader
16 will also appreciate that by discarding aspects of the physical environment, the systems and
17 methods presented herein are also applicable to virtual reality (VR) applications, which may be
18 understood as "pure" VR. For the reader's convenience, the following refers to "AR" but is
19 understood to include all of the foregoing and other variations recognized by the skilled reader.

20 [0021] A head mounted display (HMD) or other wearable display worn by a user situated in a
21 physical environment may comprise a display system and communicate with: at least one depth
22 camera disposed upon or within the HMD, or worn by (i.e., disposed upon) the user, to generate
23 depth information for the physical environment; and at least one processor disposed upon, or
24 within, the HMD, or located remotely from the HMD (such as, for example, a processor of a
25 central console, or a server) to generate a map of the physical environment from the depth
26 information. The processor may generate the map as, for example, a point cloud, in which the
27 points correspond to the obtained depth information for the physical environment.

28 [0022] Mapping a physical environment from a scanning system tied to the user may be
29 referred to as inside-out mapping or first-person-view mapping. In contrast, outside-in mapping
30 involves mapping a physical environment from one or more scanning systems situated in the
31 physical environment and directed to scan towards one or more users. It has been found that
32 user engagement with an AR may be enhanced by allowing a user to move throughout a

1 physical environment in an unconstrained manner. Inside-out mapping may provide greater
2 portability because mapping of a physical environment is performed by equipment tied to a user
3 rather than the physical environment.

4 [0023] The processor further generates an AR (also referred to as "rendered") image stream
5 comprising computer-generated imagery (CGI) for the map, and provides the AR image stream
6 to the display system for display to the user. The processor may continuously adapt the
7 rendered image stream to correspond to the user's actual position and orientation within the
8 physical environment. The processor may therefore obtain real-time depth information from the
9 depth camera to determine the user's real-time orientation and location with the physical
10 environment, as described herein in greater detail. The processor provides the rendered image
11 stream to the display system for display to the user.

12 [0024] The display system of the HMD may display an image stream of the physical
13 environment, referred to herein as a "physical image stream", to the user. The display system
14 obtains the image stream from at least one image camera disposed upon the HMD or the user,
15 either directly, or by way of the processor. The at least one image camera may be any suitable
16 image capture device operable to capture visual images of the physical environment in digital
17 format, such as, for example, a colour camera or video camera. In operation, the at least one
18 image camera dynamically captures the physical image stream for transmission to the display
19 system.

20 [0025] The display system may further simultaneously display the physical image stream
21 provided by the at least one image camera, and the rendered image stream obtained from the
22 processor. Further systems and methods are described herein.

23 [0026] Referring now to Fig. 1, an exemplary HMD 12 configured as a helmet is shown;
24 however, other configurations are contemplated. The HMD 12 may comprise: a processor 130
25 in communication with one or more of the following components: (i) at least one depth camera
26 127 (e.g., a time-of-flight camera) to capture depth information for a physical environment, and
27 at least one image camera 123 to capture at least one physical image stream of the physical
28 environment; (ii) at least one display system 121 for displaying to a user of the HMD 12 an AR
29 and/or VR and/or the image stream of the physical environment; (iii) at least one power
30 management system 113 for distributing power to the components; (iv) at least one sensory
31 feedback system comprising, for example, haptic feedback devices 120, for providing sensory

1 feedback to the user; and (v) an audio system 124 with audio input and output to provide audio
2 interaction. The processor 130 may further comprise a wireless communication system 126
3 having, for example, antennae, to communicate with other components in an AR and/or VR
4 system, such as, for example, other HMDs, a gaming console, a router, or at least one
5 peripheral 13 to enhance user engagement with the AR and/or VR. The power management
6 system may comprise a battery to generate power for the HMD, or it may obtain power from a
7 power source located remotely from the HMD, such as, for example, from a battery pack
8 disposed upon the user or located within the physical environment, through a wired connection
9 to the HMD.

10 [0027] In certain applications, the user views an AR comprising a completely rendered version
11 of the physical environment (i.e., "enhanced VR"). In such applications, the user may determine
12 the locations for obstacles or boundaries in the physical environment based solely on the
13 rendering displayed to the user in the display system 121 of the user's HMD 12.

14 [0028] As shown in Figs. 2A and 2B, an HMD 212 may comprise a display system 221 and at
15 least one depth camera 227, which are both in communication with a processor 230 configured
16 to: obtain depth information from the at least one depth camera 227, map the physical
17 environment from the depth information, and determine substantially real-time position
18 information for the HMD 212 within the physical environment; and generate a rendered image
19 stream for the map based on the real-time position information. As shown in Fig. 2A, the HMD
20 212 comprises a single depth camera 227 or, as shown in Fig. 2B, multiple depth cameras 227.
21 If the HMD 212 is equipped with multiple depth cameras 227, the multiple depth camera 227
22 may be disposed at angles to one another, or in other orientation with respect to each other
23 permitting the image cameras to capture, in combination, a wider field of view than a single
24 depth camera 227 might capture. For example, as shown in Fig. 2B, the four depth cameras
25 227 shown in Fig. 2B, are directed substantially orthogonally with respect to each other and
26 outwardly toward the physical environment with respect to the HMD 212. As configured, the four
27 depth cameras capture a 360 degree view of the regions of the physical environment outside
28 the intersection points of the fields of view of the depth cameras 227. Each of the four depth
29 cameras 227 has a field of view that is sufficiently wide to intersect with the field of view of each
30 of its neighbouring depth cameras 227. It will be appreciated that the field of view of each of the
31 depth cameras 227 is illustrated in Figs. 2A and 2B by the broken lines extending from each
32 depth camera 227 outwardly from the HMD 212 in the direction of each arrow.

1 [0029] If the at least one depth camera 227 of the HMD 212 has a combined field of view that is
2 less than 360 degrees about the HMD 212, as shown in Fig. 2A, a 360 degree view of the
3 physical space may be obtained if a user wearing the HMD 212 makes a rotation and, possibly
4 a translation, in the physical environment while the at least one depth camera 227 continuously
5 captures depth information for the physical space. However, during the user's rotation, the
6 user's head may tilt front to back, back to front, and/or shoulder to shoulder, such that the
7 continuously captured depth information is captured at different angles over the course of the
8 rotation. Therefore, the processor may invoke a stitching method, as hereinafter described, to
9 align the depth information along the rotation.

10 [0030] As shown in Fig. 3, at block 300, a depth camera on an HMD captures depth information
11 for a physical space at time $t=0$. At block 301, the depth camera captures depth information for
12 the physical space at time $t=1$. Continuously captured depth information may be understood as
13 a series of frames representing the captured depth information for a discrete unit of time.

14 [0031] At block 303, a processor receives the depth information obtained at blocks 300 and 301
15 during the user's rotation and "stitches" the depth information received during the user's
16 rotation. Stitching comprises aligning subsequent frames in the continuously captured depth
17 information to create a substantially seamless map, as outlined herein with reference to blocks
18 303 and 305.

19 [0032] The region of the physical space captured within the depth camera's field of view at time
20 $t=0$ is illustrated by the image 320; similarly, the region of the physical space captured within the
21 depth camera's field of view at time $t=1$ is illustrated by the image 321. It will be appreciated that
22 the user capturing the sequence shown in Fig. 3 must have rotated her head upwardly between
23 time $t=0$ and $t=1$. Still at block 303, the processor uses the depth information obtained at block
24 300 as a reference for the depth information obtained at block 301. For example, the television
25 shown in the image 320 has an upper right-hand corner represented by a marker 330. Similarly,
26 the same television shown in the image 321 has an upper right-hand corner defined by a marker
27 331. Further, in both images, the region underneath the markers is defined by a wall having a
28 depth profile. The processor identifies the shared topography corresponding to the markers 330
29 and 331. At block 305, the processor generates a map of the physical environment by using the
30 depth information captured at block 300 as a reference for the depth information captured at
31 block 301 based on the shared topographical feature or features identified at block 303. For
32 example, if the processor assigns coordinates x_{tr} , y_{tr} , z_{tr} to the top right-hand corner of the

1 television based on the depth information captured at block 300, the processor will then assign
2 the same coordinate to that same corner for the depth information obtained at block 301. The
3 processor thereby establishes a reference point from which to map the remaining depth
4 information obtained at block 301. The processor repeats the processes performed at blocks
5 303 and 305 for further instances of depth capture at time $t > 1$, until the depth camera has
6 obtained depth information for all 360 degrees.

7 [0033] It will be appreciated that accuracy may be enhanced if, instead of identifying a single
8 topographical feature common to subsequent depth information captures, the processor
9 identifies more than one common topographical feature between frames. Further, capture
10 frequency may be increased to enhance accuracy.

11 [0034] Alternatively, or in addition, to identifying common features between subsequent frames
12 captured by the at least one depth camera, the processor may obtain real-time orientation
13 information from an orientation detecting system for the HMD, as shown in Fig. 4. The HMD
14 may comprise an inertial measurement unit, such as a gyroscope or accelerometer, a 3D
15 magnetic positioning system, or other suitable orientation detecting system to provide
16 orientation information for the HMD to the processor, at block 311. For example, if the
17 orientation detecting system is embodied as an accelerometer, the processor may obtain real-
18 time acceleration vectors from the accelerometer to calculate the orientation of the HMD at a
19 point in time. At block 303A, the processor associates the real-time orientation of the HMD to
20 the corresponding real-time depth information. At block 305, as previously described, the
21 processor uses the depth information obtained at block 300 as a reference for depth
22 coordinates captured at block 302. However, instead of, or in addition to, the identifying of at
23 least one topographical common element between the first captured depth information and the
24 subsequently captured depth information, the processor uses the change in orientation of the
25 HMD at the time of capture of the subsequent information (as associated at block 303A) to
26 assign coordinates to that depth information relative to the first captured depth information. The
27 processor repeats steps 303A and 305 for further subsequently captured depth information until
28 the depth camera has obtained depth information for all 360 degrees about the HMD, thereby
29 generating a 360 degree map of the physical environment.

30 [0035] As shown in Fig. 2B, the HMD 212 may comprise an array of depth cameras 227, such
31 as, for example, four depth cameras 227, configured to obtain depth information for the physical
32 space for all 360 degrees about the HMD 212, even though the HMD 212 may remain stationary

1 during depth capture for mapping. As shown in Fig. 5, a first, second, third and fourth depth
2 camera each captures depth information for the physical environment, at blocks 501, 502, 503
3 and 504, respectively. All depth cameras may capture the depth information substantially
4 simultaneously. The processor obtains the depth information from the depth cameras, at block
5 505. The processor identifies each camera and its respective depth information by a unique ID.
6 At block 509, the processor obtains from a memory the orientation of each depth camera
7 relative to the HMD, which is associated in the memory to the unique ID of the depth camera, at
8 block 507. At block 511, the processor generates a map for the physical environment based on
9 the depth information received from, and the orientation of, each of the depth cameras. The
10 processor assigns a coordinate in the map for each point in the depth information; however,
11 since each of the depth cameras in the array is directed toward a different direction from the
12 other depth cameras, the processor rotates the depth information from each depth camera by
13 the rotation of each depth camera from the reference coordinate system according to which the
14 processor maps the physical environment. For example, with reference to Fig. 2B, the
15 processor may render the point P_1 on the map as the base point from which all other points in
16 the map are determined by assigning map coordinates $x, y, z = 0, 0, 0$. It will be appreciated,
17 then, that the forward-facing depth camera 227 which generates the depth information for point
18 P_1 , may return depth information that is already aligned with the map coordinates. However, at
19 block 511, the processor adjusts the depth information from the remaining depth cameras by
20 their respective relative orientations with respect to the forward-facing depth camera. The
21 processor may thereby render a map of the physical environment, at block 513.

22 [0036] The HMD may further comprise at least one imaging camera to capture a physical image
23 stream of the physical environment. The processor may enhance the map of the physical
24 environment generated using depth information from the at least one depth camera by adding
25 information from the physical image stream of the physical environment. During mapping
26 according to the previously described mapping methods, the processor may further obtain a
27 physical image stream of the physical environment from the at least one imaging camera, as
28 shown in Fig. 6. The at least one imaging camera captures a physical image stream of the
29 physical environment, at block 601. Substantially simultaneously, the at least one depth camera
30 captures depth information for the physical environment, at block 603. At block 609, the
31 processor obtains the depth information from the at least one depth camera and the physical
32 image stream of the physical environment from the at least one imaging camera. Each imaging
33 camera may have a predetermined relationship, in terms of location, orientation and field of

1 view, with respect to the at least one depth camera, defined in a memory at block 605. The
2 processor obtains the definition at block 607. At block 609, the processor assigns depth data to
3 each pixel in the physical image stream based on the depth information and the predetermined
4 relationship for the time of capture of the relevant pixel. At block 611, the processor stitches the
5 physical images captured in the physical image stream using stitching methods analogous to
6 those described above, with suitable modification for images, as opposed to depth data. For
7 example, the processor may identify common graphic elements or regions within subsequent
8 frames. At block 613, the processor generates an image and depth map of the physical
9 environment.

10 [0037] Once the processor has mapped the physical environment, the processor may track
11 changes in the user's orientation and position due to the user's movements. As shown in Fig. 7,
12 at block 701, the at least one image camera continues to capture a physical image stream of the
13 physical environment. Further, or alternatively, the at least one depth camera continues to
14 capture depth information for the physical environment at block 703. At block 705, the processor
15 continues to obtain data from each or either of the real-time image stream and depth
16 information. At block 711, the processor may compare the real-time image stream to the image
17 map generated according to, for example, the method described above with reference to Fig. 6,
18 in order to identify a graphic feature common to a mapped region and the image stream. Once
19 the processor has identified a common region, it determines the user's location and orientation
20 with respect to the map at a point in time corresponding to the compared portion (i.e., frame) of
21 the image stream. By determining the transformation required to scale and align the graphic
22 feature in the physical image stream with the same graphic feature in the image map, the
23 processor may determine the user's position and orientation with reference to the map. Further,
24 or alternatively, the processor may perform an analogous method for depth information obtained
25 in real-time from the at least one depth camera. At block 713, the processor identifies a
26 topographical feature for a given point in time in the real-time depth information, and also
27 identifies the same topographical feature in the depth map of the physical environment. At block
28 723, the processor determines the transformation required to scale and align the topographical
29 feature between the real-time depth information and the depth map in order to determine the
30 user's position and orientation at the given point in time. The processor may verify the position
31 and orientation determined at blocks 721 and 723 with reference to each other to resolve any
32 ambiguities, or the common regions identified at blocks 711 and 713. For example, if the image
33 map for the physical environment comprises two or more regions which are graphically identical,

1 a graphical comparison alone would return a corresponding number of locations and
2 orientations for the HMD; however, as shown by the dashed lines in Fig. 7, the processor may
3 use the depth comparison to resolve erroneous image matching, and vice versa.

4 [0038] Alternatively, the HMD may comprise a local positioning system and/or an orientation
5 detecting system, such as, for example, a 3D magnetic positioning system, laser positioning
6 system and/or inertial measurement unit to determine the real-time position and orientation of
7 the user.

8 [0039] Augmented reality involves combining CGI (also understood as renderings generated
9 by a processor) with a physical image stream of the physical environment. An HMD for AR and
10 VR applications is shown in Fig. 1, as previously described. The display system 121 may be
11 operable to receive either a combined image stream (i.e., a physical image stream and a
12 rendered image stream) from the processor, or to simultaneously receive a physical image
13 stream from the at least one imaging camera and the rendered image stream from the
14 processor, thereby displaying an AR to the user of the HMD 12. The processor generates a
15 rendered image stream according to any suitable rendering techniques for display on the
16 display system of the HMD. The rendered image stream may comprise, for example, CGI within
17 the map of the physical environment.

18 [0040] The display system of an HMD may display the rendered image stream alone (enhanced
19 VR) or overlaid over the physical image stream to combine the visual and typographic aspects
20 of the physical environment (AR).

21 [0041] In an enhanced VR application, the processor may enhance a user's interaction with the
22 physical environment by accounting for the user's real-time location and orientation within the
23 physical environment when generating the rendered image stream. As the user moves about
24 the physical environment, the VR of that physical environment displayed to the user will reflect
25 changes in the user's position and/or orientation. As shown in Fig. 8, at block 801, the processor
26 determines the orientation and location of the user's HMD according to any suitable method,
27 including the orientation and positioning methods described above. In an enhanced VR
28 application, parameters corresponding to a notional or virtual camera may be defined, at block
29 803, in a memory accessible by the processor. For example, the notional camera may have a
30 defined notional field of view and relative location on the HMD. At block 805, the processor
31 determines which region of the map lies within the field of view of the notional camera based on

1 the orientation and location information obtained at block 801 in conjunction with the camera
2 parameters defined at block 803. At block 807, the processor generates a rendered image
3 stream of the region of the map lying within the notional field of view, including any CGI within
4 that region. At block 809, the display system of the HMD may display the rendered image
5 stream in substantially real-time, where processing time for generating the image stream is
6 responsible for any difference between the actual and orientation and location, and the
7 displayed notional orientation and location, of the user's HMD.

8 [0042] In an AR application, the display system of an HMD may display the rendered image
9 stream overlaid over, or combined with, the physical image stream. When the at least one
10 image camera captures an image stream of the physical environment, the captured physical
11 image stream at any given moment will comprise elements of the physical environment lying
12 within the field of view of the camera at that time.

13 [0043] The physical image stream obtained by the camera is either transmitted to the processor
14 for processing and/or transmission to the display system, or directly to the display system for
15 display to the user.

16 [0044] Referring now to Fig. 9, a method of overlapping the physical image stream with the
17 rendered image stream is shown. At block 901, the at least one image camera captures the
18 physical image stream of the physical environment. As the at least one image camera captures
19 the physical image stream, the processor determines, at block 903, the real-time orientation and
20 location of the HMD in the physical environment. Parameters corresponding to the field of view
21 of the at least one image camera, and the position and orientation of the at least one image
22 camera relative to the HMD are defined in a memory, at block 905. At block 907, the processor
23 determines the region of the physical environment lying within the field of view of the at least
24 one image camera in real-time using the real-time orientation and location of the HMD, as well
25 as the defined parameters for the at least one image camera. At block 909, the processor
26 generates a rendered image stream comprising rendered CGI within a region of the map of the
27 physical environment corresponding to the region of the physical environment lying within the
28 field of view of the at least one image camera. The region in the rendered image stream may be
29 understood as a region within the field of view of a notional camera having the same orientation,
30 location and field of view in the map as the at least one image camera has in the physical
31 environment, since the map is generated with reference to the physical environment. At block
32 911, the display system of the HMD obtains the rendered and physical image streams and

simultaneously displays both. The physical image stream may be provided directly to the display system, or it may first pass to the processor for combined transmission to the display system along with the rendered image stream.

[0045] If the fields of view of the notional and physical cameras are substantially aligned and identical, simultaneous and combined display of both image streams provides a combined stream that is substantially matched.

[0046] In embodiments, the processor may increase or decrease the signal strength of one or the other of the physical and rendered image streams to vary the effective transparency.

[0047] In embodiments, the processor only causes the display system to display the physical image stream upon the user selecting display of the physical image stream. In further embodiments, the processor causes the display system to display the physical image stream in response to detecting proximity to an obstacle in the physical environment. In still further embodiments, the processor increases the transparency of the rendered image stream in response to detecting proximity to an obstacle in the physical environment. Conversely, the processor may reduce the transparency of the rendered image stream as the HMD moves away from obstacles in the physical environment.

[0048] In still further embodiments, the display system displays the physical and rendered image streams according to at least two of the techniques described herein.

[0049] The scope of the claims should not be limited by the preferred embodiments set forth in the examples, but should be given the broadest interpretation consistent with the description as a whole.

CLAIMS

What is claimed is:

1. A method for mapping a physical environment in which a user wearing a wearable display for augmented reality is situated, the method comprising:
 - (a) capturing, by at least one depth camera disposed upon the user, depth information for the physical environment;
 - (b) by a processor, obtaining the depth information, determining the orientation of the at least one depth camera relative to the wearable display, and assigning coordinates for the depth information in a map of the physical environment based on the orientation of the at least one depth camera.

wherein:

- i) the capturing comprises continuously capturing a sequence of frames of depth information for the physical environment during rotation and translation of the at least one depth camera in the physical environment; and
 - ii) the obtaining further comprises continuously determining the translation and the rotation of the at least one depth camera between each of the frames, and;
 - iii) the assigning comprises assigning first coordinates to the depth information from a first frame and assigning subsequent coordinates to the depth information from each of the subsequent frames according to the rotation and translation of the at least one depth camera between each of the frames.
2. The method of claim 1, further comprising:
 - (a) identifying topography shared between first and second ones of subsequent frames;

- (b) assigning shared coordinates to the shared topography for each of the first and second ones of the subsequent frames; and
 - (c) assigning coordinates for the second one of the subsequent frames with reference to the coordinates for the shared topography.
3. The method of claim 1, further comprising:
- (a) capturing, by at least one image camera disposed upon the user, a physical image stream of the physical environment;
 - (b) obtaining the physical image stream, determining the orientation of the at least one image camera relative to the wearable display, and assigning coordinates to a plurality of pixels in the physical image stream in the map of the physical environment based on the orientation of the at least one image camera.
4. A system for mapping a physical environment surrounding a user wearing a wearable display for augmented reality, the system comprising:
- (a) at least one depth camera disposed upon the user, to capture depth information for the physical environment;
 - (b) at least one processor in communication with the at least one depth camera, to obtain the depth information from the at least one depth camera, determine the orientation of the at least one depth camera relative to the wearable display, and assign coordinates for the depth information in a map of the physical information based on the orientation of the at least one depth camera.
- wherein the at least one depth camera is configured to continuously capture a sequence of frames of depth information for the physical environment during rotation and translation of the at least one depth camera in the physical environment, and the processor is configured to:
- (c) continuously determine the rotation and the translation of the at least one depth camera between each of the frames;
 - (d) assign coordinates for the depth information by assigning first coordinates to the depth information from a first frame and assigning subsequent coordinates to the

depth information from each of subsequent frames based on the rotation and translation of at least one physical between each of the frames.

5. The system of claim 4, wherein the processor is further configured to:
 - (a) identify topography shared between each frame and a subsequent frame;
 - (b) assign the same coordinates to the shared topography for each frame and the subsequent frame; and
 - (c) assign coordinates for the subsequent frame with reference to the coordinates for the shared topography.
6. The system of claim 4, further comprising at least one image camera disposed upon the user, operable to capture a physical image stream of the physical environment, and wherein the processor is configured to:
 - (a) obtain the physical image stream, determine the orientation of the at least one image camera relative to the wearable display; and
 - (b) assign coordinates to a plurality of pixels in the physical image stream in the map of the physical environment based on the orientation of the at least one image camera.

1 / 9

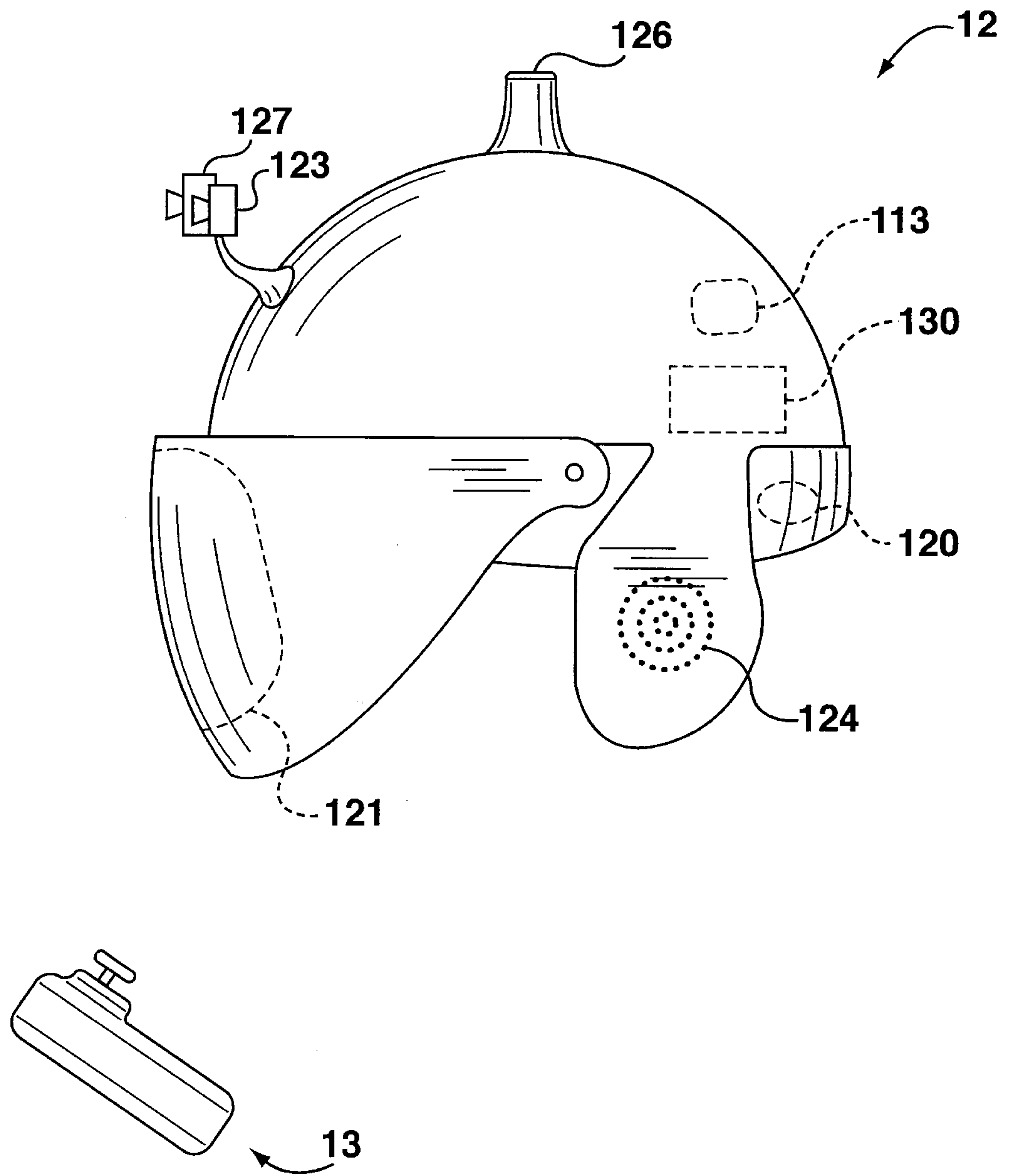


FIG. 1

2 / 9

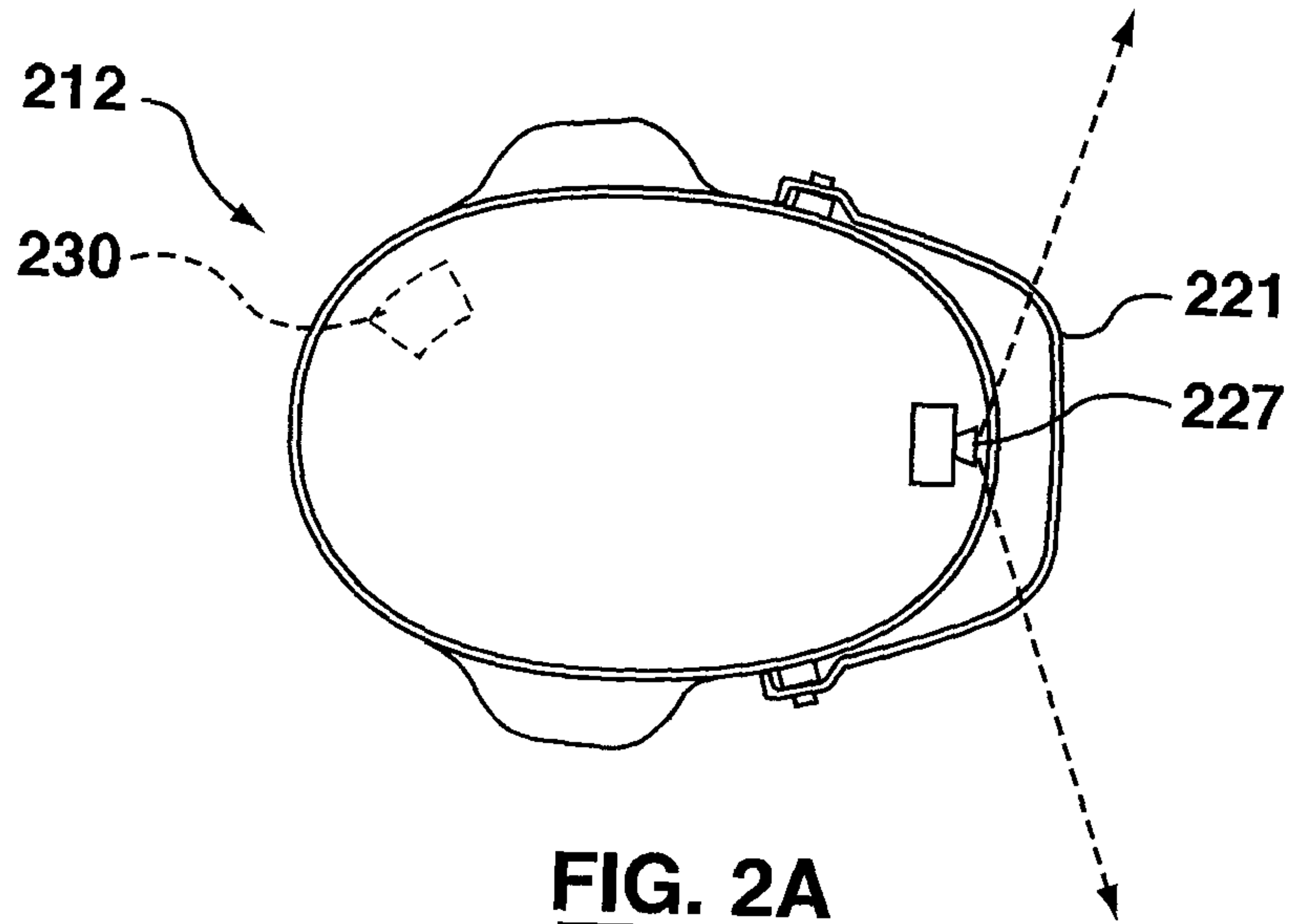


FIG. 2A

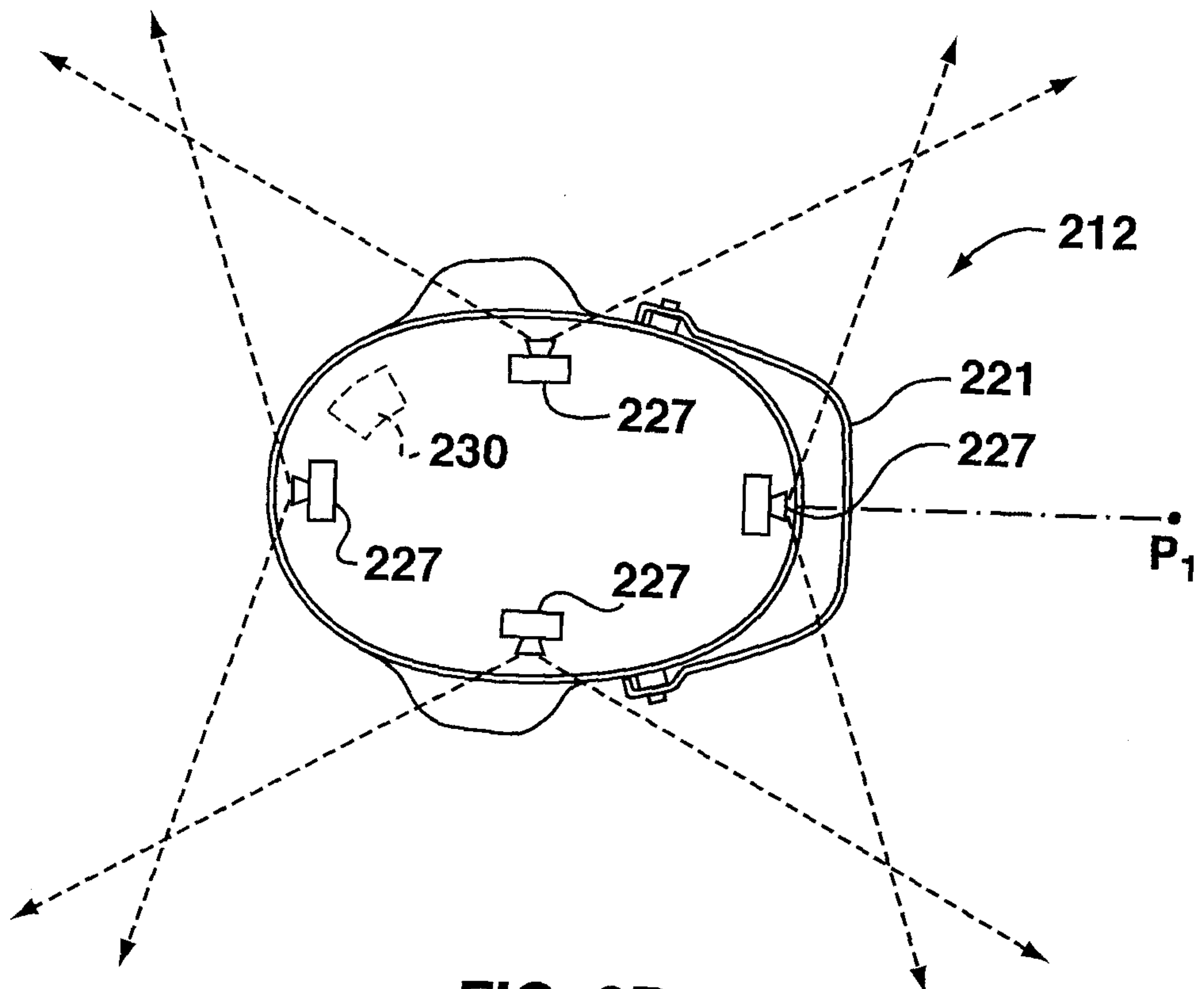


FIG. 2B

3 / 9

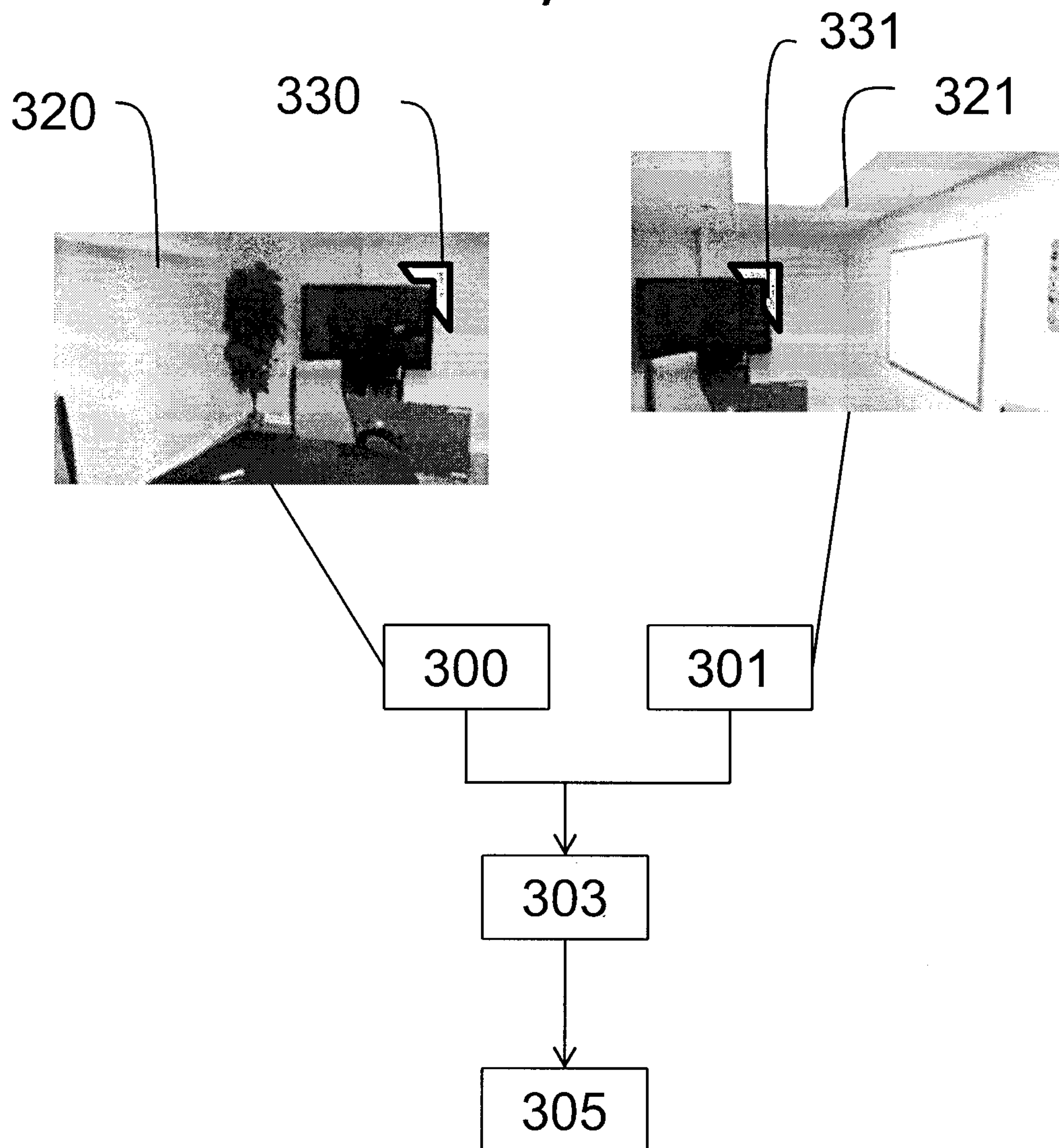


FIG. 3

4 / 9

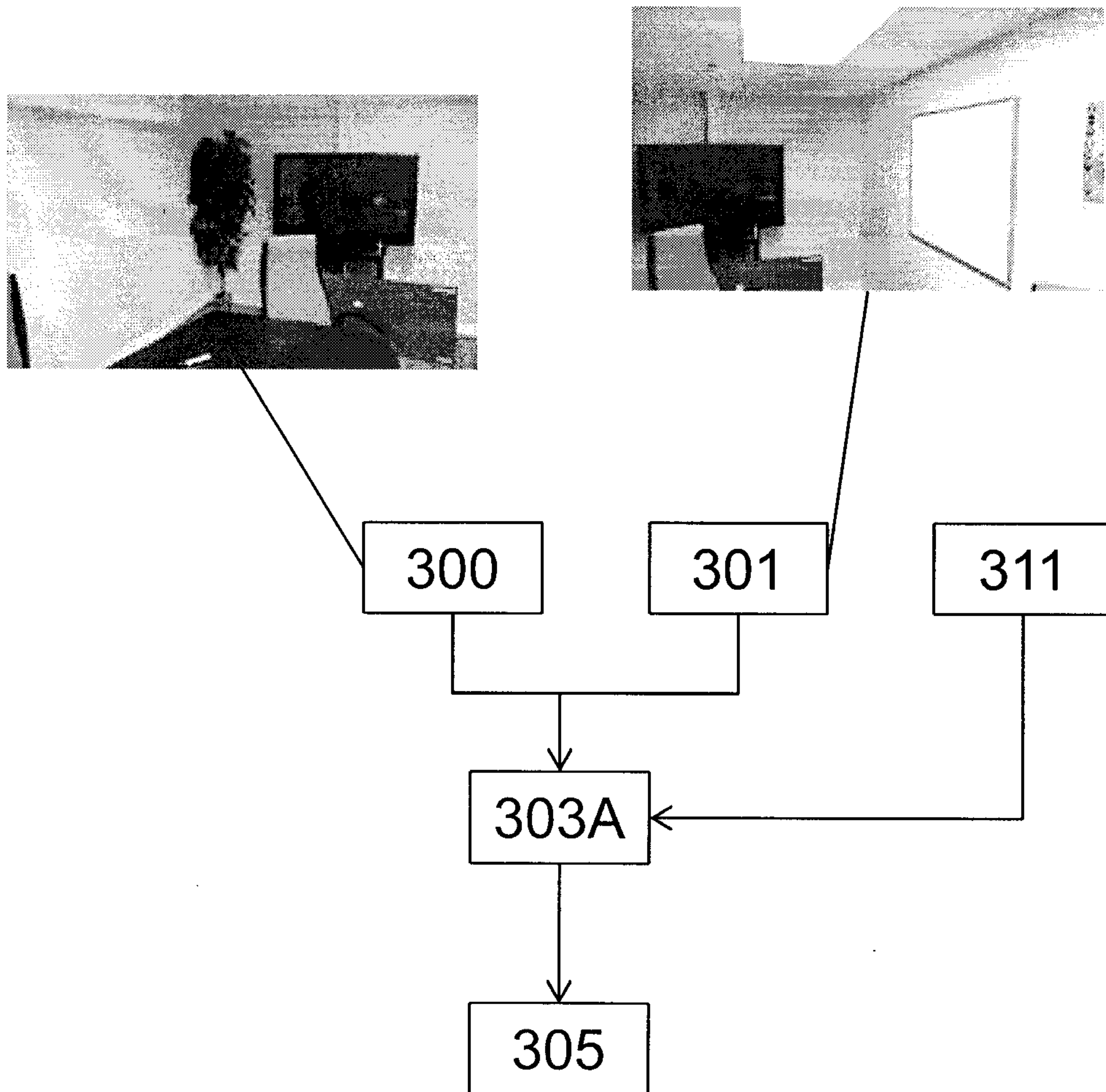


FIG. 4

5 / 9

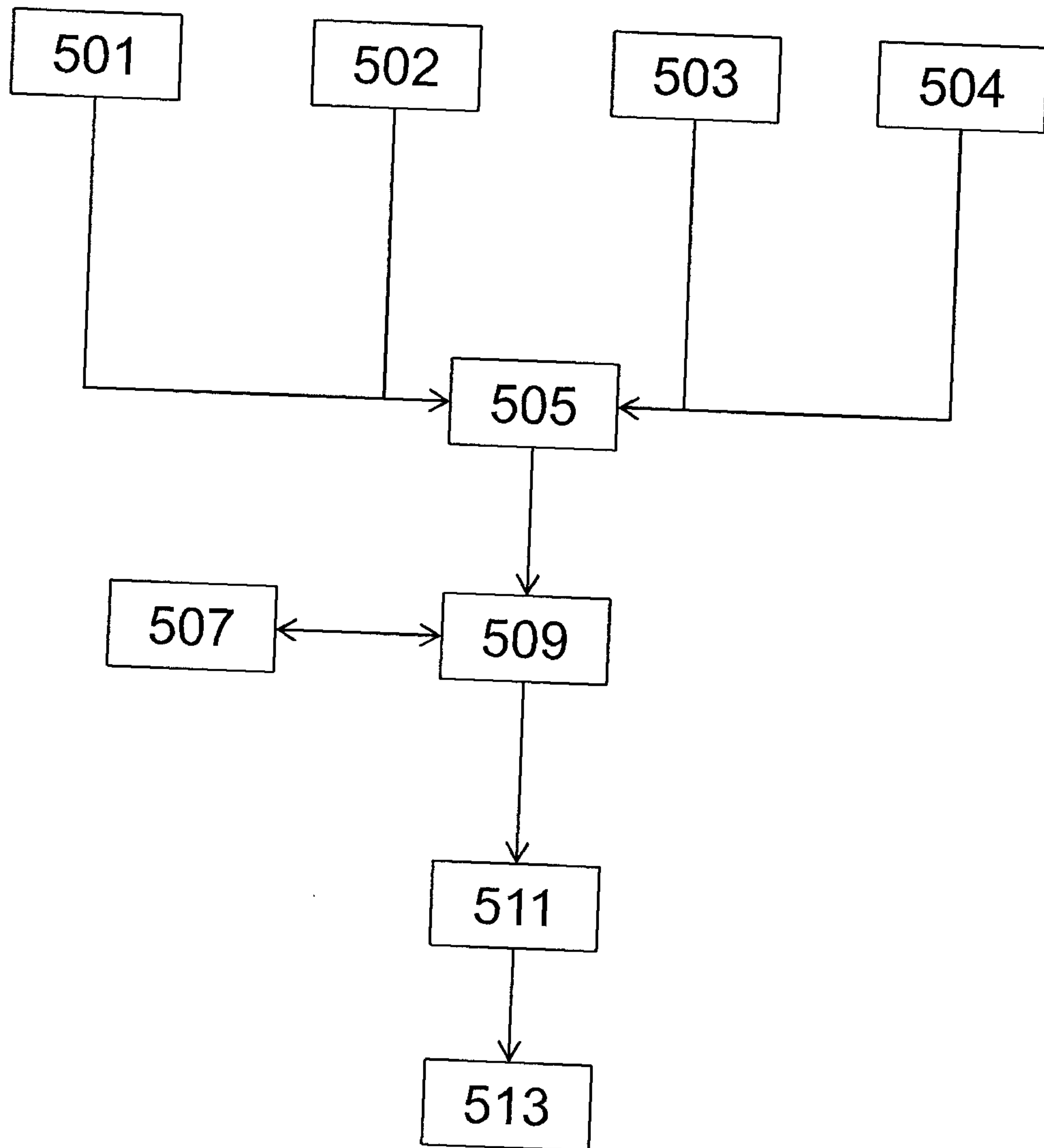


FIG. 5

6 / 9

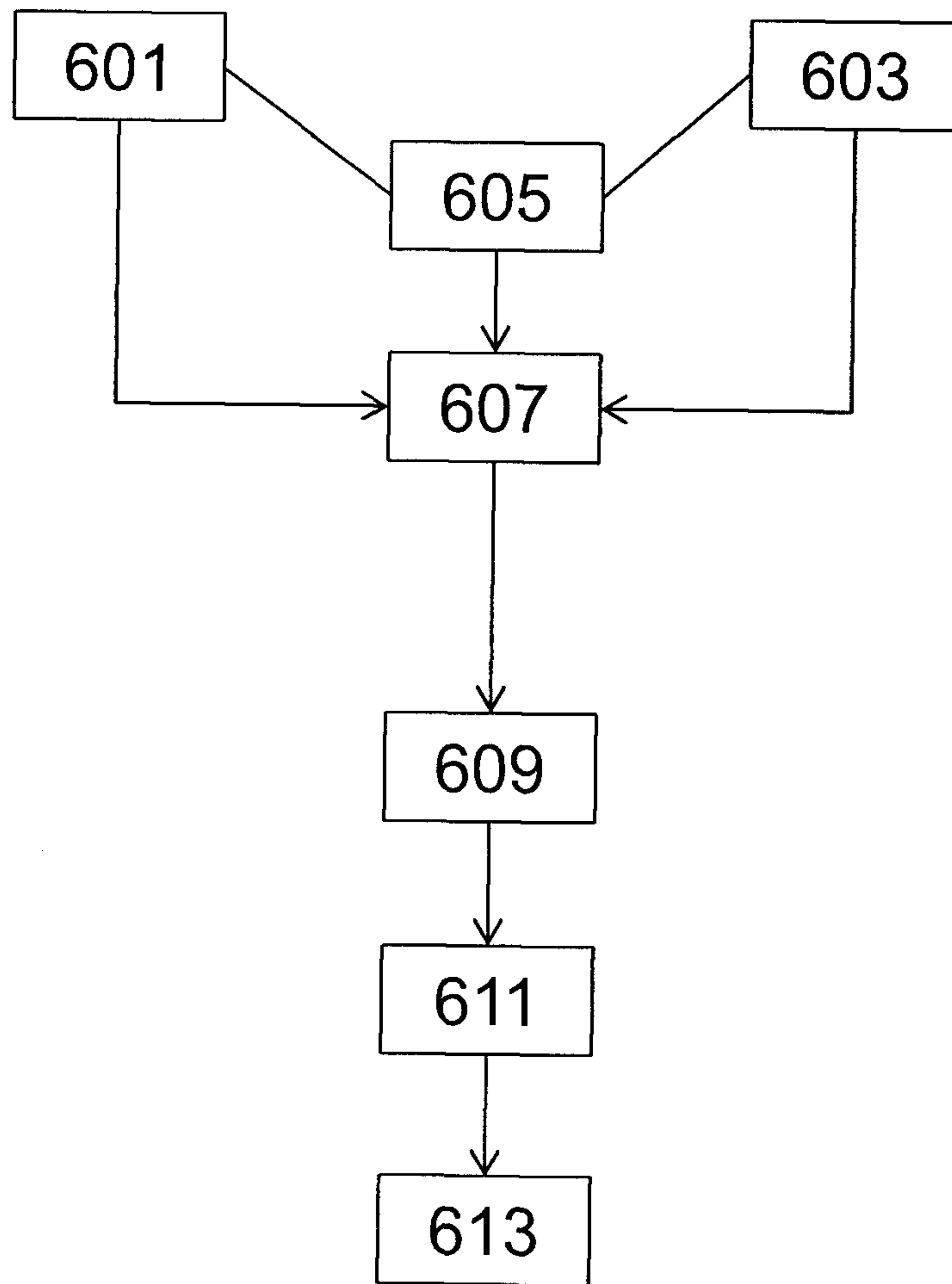


FIG. 6

7 / 9

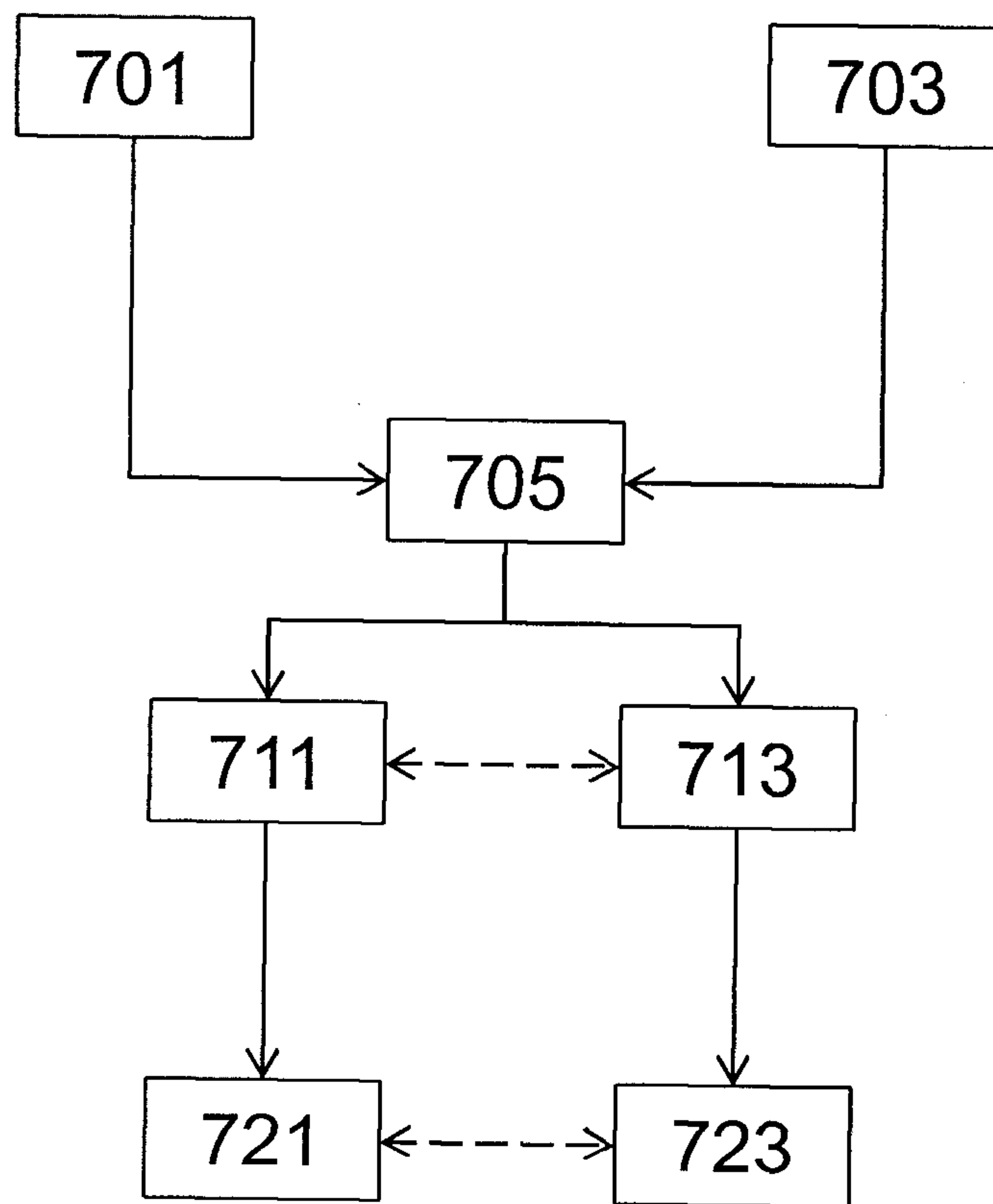


FIG. 7

8 / 9

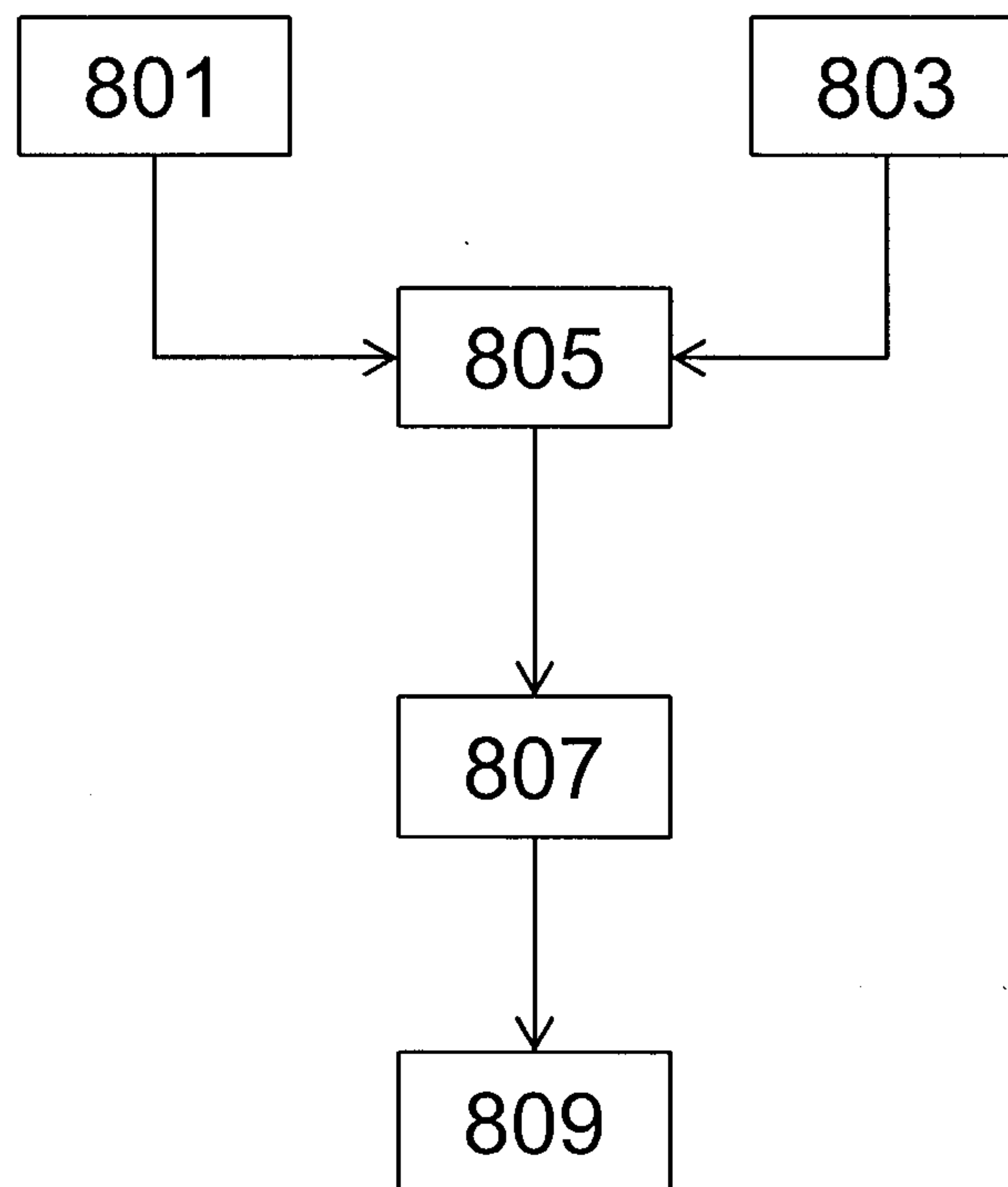


FIG. 8

9 / 9

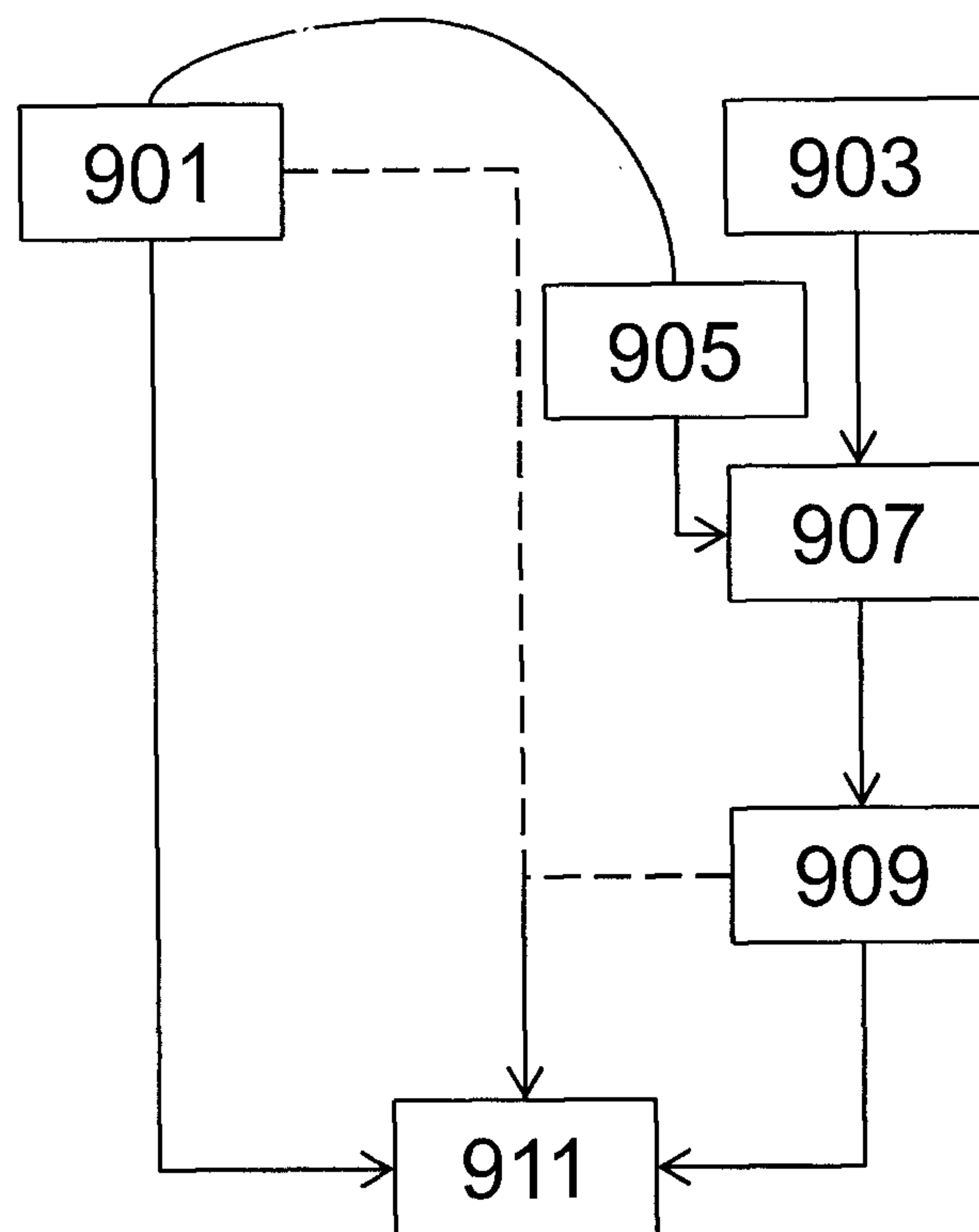


FIG. 9

