



(22) Date de dépôt/Filing Date: 2011/08/29

(41) Mise à la disp. pub./Open to Public Insp.: 2011/11/02

(45) Date de délivrance/Issue Date: 2012/07/03

(51) Cl.Int./Int.Cl. *G02B 27/01* (2006.01),
A61B 3/10 (2006.01), *A61B 3/107* (2006.01),
A61B 3/11 (2006.01), *A61B 3/113* (2006.01),
G06F 19/00 (2011.01)

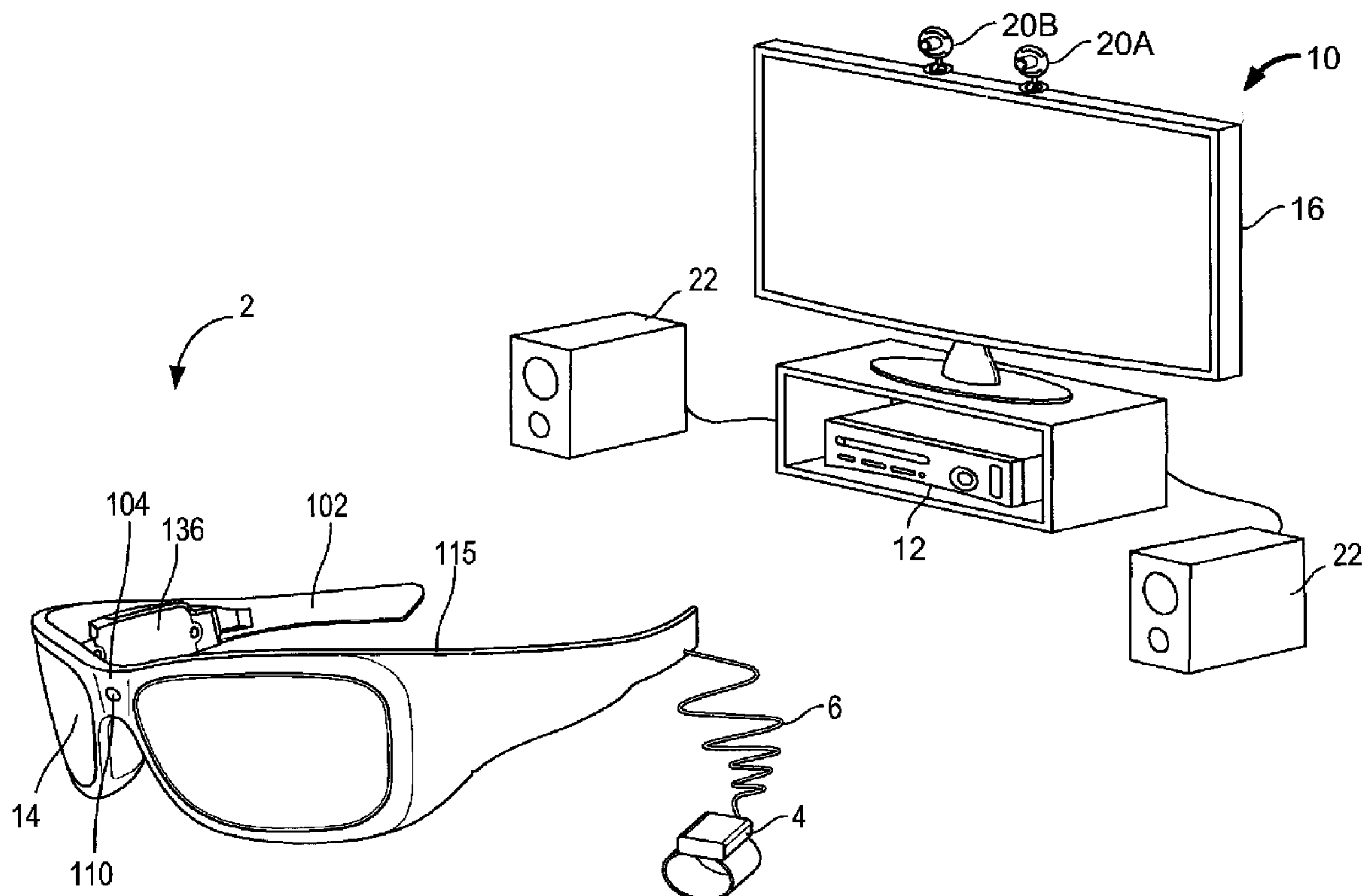
(72) Inventeurs/Inventors:
LEWIS, JOHN R., US;
WEI, YICHEN, US;
CROCCO, ROBERT L., US;
VAUGHT, BENJAMIN I., US;
KIPMAN, ALEX ABEN-ATHAR, US;
PEREZ, KATHRYN STONE, US

(73) Propriétaire/Owner:
MICROSOFT CORPORATION, US

(74) Agent: SMART & BIGGAR

(54) Titre : DETECTION DU REGARD DANS UN AFFICHAGE TRANSPARENT, PRES DE L'OEIL ET DE REALITE MIXTE

(54) Title: GAZE DETECTION IN A SEE-THROUGH, NEAR-EYE, MIXED REALITY DISPLAY



(57) Abrégé/Abstract:

The technology provides various embodiments for gaze determination within a see-through, near-eye, mixed reality display device. In some embodiments, the boundaries of a gaze detection coordinate system can be determined from a spatial relationship

(57) **Abrégé(suite)/Abstract(continued):**

between a user eye and gaze detection elements such as illuminators and at least one light sensor positioned on a support structure such as an eyeglasses frame. The gaze detection coordinate system allows for determination of a gaze vector from each eye based on data representing glints on the user eye, or a combination of image and glint data. A point of gaze may be determined in a three-dimensional user field of view including real and virtual objects. The spatial relationship between the gaze detection elements and the eye may be checked and may trigger a re-calibration of training data sets if the boundaries of the gaze detection coordinate system have changed.

52785-18(S)

ABSTRACT

The technology provides various embodiments for gaze determination within a see-through, near-eye, mixed reality display device. In some embodiments, the boundaries of a gaze detection coordinate system can be determined from a spatial relationship between a user eye and gaze detection elements such as illuminators and at least one light sensor positioned on a support structure such as an eyeglasses frame. The gaze detection coordinate system allows for determination of a gaze vector from each eye based on data representing glints on the user eye, or a combination of image and glint data. A point of gaze may be determined in a three-dimensional user field of view including real and virtual objects. The spatial relationship between the gaze detection elements and the eye may be checked and may trigger a re-calibration of training data sets if the boundaries of the gaze detection coordinate system have changed.

52785-18(S)

-1-

GAZE DETECTION IN A SEE-THROUGH, NEAR-EYE, MIXED REALITY DISPLAY

BACKGROUND

[0001] Augmented or mixed reality is a technology that allows virtual imagery to be mixed with a user's actual view of the real world. A near-eye display may be worn by a user to view the mixed imagery of virtual and real objects. A near-eye display displays virtual imagery in the user's field of view. However, the user's field of view is not stationary as a user moves his or her head. Furthermore, what the user is looking at in the field of view changes as the user shifts his or her eyes, even if his or her head does not move.

SUMMARY

[0002] The technology provides various embodiments for gaze determination within a see-through, near-eye mixed reality display device. Gaze is sometimes referred to as a line of sight from the user's eye to an object, real or virtual, at which the user is looking. Embodiments are provided for determining gaze based on glint data or a combination of both glint and image data. In some embodiments, a gaze determination coordinate system based on predetermined positioning of at least one light sensor and illuminators on the display device provides a three dimensional (3D) spatial relationship between the display device and each respective eye. A gaze vector for each eye may be determined based on the 3D spatial relationship. Based on gaze vectors for both eyes, a point of gaze may be determined which indicates one or more objects, real or virtual, at which a user is gazing, or more commonly stated as, at which the user is looking.

52785-18(S)

-2-

[0003] The technology provides an embodiment of a mixed reality display system with gaze determination. The system embodiment comprises a see-through, near-eye display device including a respective display optical system for each eye positioned to be seen through by the respective eye. For each eye, an image generation unit is attached to the see-through display device for generating at least one virtual image for display. Additionally, the embodiment includes a respective arrangement of gaze detection elements positioned on the display device for forming a spatial relationship between the gaze detection elements and each eye. The gaze detection elements include a set of illuminators for generating glints on the respective eye, and at least one respective sensor for capturing light reflected from the respective eye and generating data representing the captured reflected light. The system embodiment includes memory for storing software and the data. A processor is communicatively coupled to the at least one respective sensor for receiving and storing the data representing the captured reflected light in the accessible memory. The processor determines a gaze vector for each respective eye based on the data representing the captured reflected light and a point of gaze based on the gaze vectors in a three-dimensional (3D) user field of view.

[0004] The technology provides an embodiment of a method for determining gaze in a see-through, near-eye mixed reality display system. The method embodiment comprising determining boundaries of a gaze detection coordinate system based on positions of glints detected on a user eye, positions on the near-eye display system of illuminators for generating the glints; and a position of at least one sensor for detecting the glints. The method further comprises determining a gaze vector for each eye based on reflected eye data including the glints; determining a point of gaze based on the gaze vectors for the two eyes in a three-dimensional (3D) user field of view including real and virtual objects, and identifying any object at the point of gaze in the 3D user field of view.

52785-18(S)

-3-

[0005] Another system embodiment is also provided by the technology for a mixed reality display system with gaze determination based on glints. The system comprises a see-through, near-eye display device including a respective display optical system for each eye positioned to be seen through by the respective eye and a respective image generation unit for each eye attached to the see-through display device for generating at least one virtual image for display. The system further comprises a set of infra-red (IR) illuminators for producing glints for each eye and at least one respective sensor is positioned on the near-eye display device at a predetermined position to detect the glints and generate glint data to be stored in a memory accessible by a processor which determines a gaze vector for each eye based on the glint data. A point of gaze in a three-dimensional user field of view is determined by the software controlled processor based on the gaze vectors.

According to one aspect of the present invention, there is provided a mixed reality display system with gaze determination comprising: a see-through, near-eye, mixed reality display device including a respective display optical system for each eye positioned to be seen through by the respective eye and at least one image generation unit for generating at least one virtual image for display in the display optical systems; a respective arrangement of gaze detection elements positioned on the display device for forming a spatial relationship between the gaze detection elements and each eye, the gaze detection elements including a set of illuminators for generating glints on the respective eye, each illuminator positioned on the respective see-through display device at a respective predetermined position and generating illumination about a predetermined wavelength, and at least one respective sensor for capturing reflections from the respective eye and generating data representing the captured reflections, the at least one respective sensor positioned at a predetermined position in relation to predetermined positions of the

52785-18(S)

-3a-

set of illuminators on the respective see-through display device; a memory for storing software and the data; and a processor communicatively coupled to the at least one respective sensor to receive the data representing the captured reflections and having access to the memory for storing the data, the processor determining a gaze vector for each respective eye based on the data representing the captured reflections and a point of gaze based on the gaze vectors in a three-dimensional (3D) user field of view; wherein data representing the captured reflections includes glint intensity data and wherein for each eye the processor determining the gaze vector comprises: at least some of the time using a first technique that comprises determining the center of the cornea and the center of the pupil based on image data of the respective eye; and at least some of the time using a second different, less computationally intensive technique based on the glint intensity data independent of the image data of the respective eye.

According to another aspect of the present invention, there is provided a method for determining gaze in a see-through, near-eye mixed reality display system comprising: determining boundaries of a three dimensional gaze detection coordinate system based on positions of glints detected on a user eye, positions on the near-eye display system of illuminators for generating the glints, and a position of at least one sensor for detecting the glints; determining a gaze vector for each eye based on reflected eye data including the glints; determining a point of gaze based on the gaze vectors for the two eyes in a three-dimensional (3D) user field of view including real and virtual objects; and identifying any object at the point of gaze in the 3D user field of view.

According to still another aspect of the present invention, there is provided a mixed reality display system with gaze determination based on glints comprising: a see-through, near-eye display device including a respective display

52785-18(S)

-3b-

optical system for each eye positioned to be seen through by the respective eye and at least one image generation unit for generating at least one virtual image for display; a set of infra-red illuminators for producing glints for each eye, each illuminator positioned on the respective see-through display device at a respective predetermined position and generating infra-red radiation about a predetermined wavelength; at least one respective sensor for detecting the glints reflected from each eye and generating glint data including intensity data for the glints, the at least one respective sensor positioned in relation to one or more respective illuminator positions on the respective see-through display device at at least one predetermined position to detect infra-red reflected radiation about the predetermined wavelength; a memory for storing software and data including glint data; and a processor communicatively coupled to the at least one respective sensor to receive the glint data and having access to the memory for storing the glint data, the processor determining one or more glint positions relative to an eye part based on the glint intensity data generated by the at least one respective sensor, the respective predetermined position of each illuminator in the set of infra-red illuminators, and the at least one predetermined position of the at least one respective sensor for detecting the glints reflected from each eye; the processor determining a gaze vector for each respective eye based on the one or more glint positions relative to an eye part; and the processor determining a point of gaze based on the gaze vectors in a three dimensional user field of view.

[0006] This Summary is provided to introduce a selection of concepts in a simplified form that are further described below in the Detailed Description. This Summary is not intended to identify key features or essential features of the claimed subject matter, nor is it intended to be used as an aid in determining the scope of the claimed subject matter.

52785-18(S)

-3c-

BRIEF DESCRIPTION OF THE DRAWINGS

[0007] Figure 1A is a block diagram depicting example components of one embodiment of a see-through, mixed reality display device with gaze determination.

[0008] Figure 1B is a block diagram depicting example components of another embodiment of a see-through, mixed reality display device with gaze determination.

[0009] Figure 1C illustrates an exemplary arrangement of positions of respective sets of gaze detection elements in a gaze detection system for each eye

52785-18(S)

4

positioned facing each respective eye on a mixed reality display device embodied in a set of eyeglasses.

[0010] Figure 1D illustrates another exemplary arrangement of positions of respective sets of gaze detection elements in a gaze detection system for each eye positioned facing each respective eye on a mixed reality display device embodied in a set of eyeglasses.

[0011] Figure 1E illustrates yet another exemplary arrangement of positions of respective sets of gaze detection elements in a gaze detection system for each eye positioned facing each respective eye by the set of eyeglasses.

[0012] Figure 2 is a side view of an eyeglass temple in an embodiment of a mixed reality display device providing support for hardware and software components.

[0013] Figure 3A is a top view of an embodiment of a display optical system of a see-through, near-eye, mixed reality device including an arrangement of gaze detection elements in a gaze detection system.

[0014] Figure 3B is a top view of another embodiment of a display optical system of a see-through, near-eye, mixed reality device including an arrangement of gaze detection elements in a gaze detection system.

[0015] Figure 3C is a top view of a third embodiment of a display optical system of a see-through, near-eye, mixed reality device including an arrangement of gaze detection elements in a gaze detection system.

[0016] Figure 3D is a top view of a fourth embodiment of a display optical system of a see-through, near-eye, mixed reality device including an arrangement of gaze detection elements in a gaze detection system.

52785-18(S)

5

[0017] Figure 4A is a block diagram of one embodiment of hardware and software components of a see-through, near-eye, mixed reality display unit as may be used for the embodiment of Figure 2.

[0018] Figure 4B is a block diagram of one embodiment of the hardware and software components of a processing unit associated with a see-through, near-eye, mixed reality display unit.

[0019] Figure 5 is a top view illustrating examples of gaze vectors intersecting at a point of gaze where a user's eyes are focused.

[0020] Figure 6A is a flowchart of a method embodiment for determining a three-dimensional user field of view.

[0021] Figure 6B is a flowchart of a method embodiment for identifying one or more real objects in a user field of view.

[0022] Figure 6C is a flowchart of a method embodiment for generating a three-dimensional model of a user space.

[0023] Figure 6D is a flowchart of a method embodiment for identifying one or more objects in a user field of view.

[0024] Figure 6E is a flowchart of a method embodiment for identifying one or more objects in a user field of view.

[0025] Figure 6F is a block diagram of a system embodiment for determining positions of objects within a user field of view of a see-through, near-eye display device.

[0026] Figure 7 is a flowchart of a method embodiment for determining gaze in

52785-18(S)

6

a see-through, near-eye mixed reality display system.

[0027] Figure 8 is a flowchart of a method embodiment for identifying glints in image data.

[0028] Figure 9 is a flowchart of a method embodiment which may be used to determine boundaries for a gaze detection coordinate system.

[0029] Figure 10 is a flowchart illustrating a method embodiment for determining a position of a center of a cornea in the coordinate system with optical gaze detection elements of the see-through, near-eye, mixed reality display.

[0030] Figure 11 provides an illustrative example of defining a plane using the geometry provided by the arrangement of optical elements to form the gaze detection coordinate system which may be used by the embodiment of Figure 10 to find the cornea center.

[0031] Figure 12 is a flowchart illustrating a method embodiment for determining a pupil center from image data generated by a sensor.

[0032] Figure 13 is a flowchart illustrating a method embodiment for determining a gaze vector based on the determined centers for the pupil, the cornea and a center of rotation of an eyeball.

[0033] Figure 14 is a flowchart illustrating a method embodiment for determining gaze based on glint data.

[0034] Figure 15a is a flowchart illustrating a method embodiment for generating a set of training data sets for a comparison based determination of gaze.

[0035] Figure 15b is a flowchart illustrating a method embodiment for

52785-18(S)

7

determining gaze based on the training data sets.

[0036] Figure 15c is a flowchart of an interpolation method embodiment which may be used with the comparison step of Figure 14.

[0037] Figure 16 is a flowchart illustrating a method embodiment for checking whether re-calibration of a training gaze data sets is to be done.

[0038] Figure 17 is a block diagram of an exemplary mobile device which may operate in embodiments of the technology.

[0039] Figure 18 is a block diagram of one embodiment of a computing system that can be used to implement a hub computing system.

DETAILED DESCRIPTION

[0040] The technology provides various embodiments for gaze determination within a see-through, near-eye, mixed reality display device. Gaze is sometimes referred to as a line of sight or a visual axis of an eye. The visual axis extends from the fovea, sometimes referred to as the fovealcentralis, of the retina through the center of the pupil. Extending the visual axis from the fovea through the pupil and a see-through lens for each eye, one can determine a point of gaze in a user's field of view which may include images of virtual objects, and an actual direct view of the real world.

[0041] The use of the term "actual direct view" refers to the ability to see real world objects directly with the human eye, rather than seeing created image representations of the objects. For example, looking through glass at a room allows

52785-18(S)

8

a user to have an actual direct view of the room, while viewing a video of a room on a television is not an actual direct view of the room. Based on the context of executing software, for example, a gaming application, the system can project images of virtual objects, sometimes referred to as virtual images, on the display that are viewable by the person wearing the see-through display device while that person is also viewing real world objects through the display.

[0042] Geometry of one or more gaze detection elements with respect to the visible portion of a human eye forms a basis for various embodiments of gaze determination. In some embodiments, data from only glints may be used to track changing intensities due to different reflectivities on parts of the eye like the sclera, sometimes referred to as the white section of the eye, the pupil and the iris. A glint is a very small and often very bright reflection of light from a light source off of a surface of the cornea of an eye. The glint is an image of the light source, typically a narrow beam source focused on the eye. In some embodiments, a training or calibration gaze data set for the glints may be used in comparisons for detecting a current gaze.

[0043] Other embodiments use both image data of the eye and data representing glints in the context of a geometry of the illuminators and at least one image sensor to determine boundaries of a three-dimensional (3D) spatial relationship between positions of parts of the eye and a respective system of gaze detection elements. Examples of such parts of the eye are a center of a cornea determined based on glint data, a center of a pupil determined from image data of an eye, and a center of rotation of the eye a position of which is estimated based on the position of the cornea center. For accuracy considerations in gaze determination purposes, the center of rotation of the eyeball may be considered fixed. A gaze vector for the respective eye is determined based on the cornea

52785-18(S)

9

center, pupil center, and center of rotation which form an optical axis for the respective eye. An angle offset may be applied to the optical axis in order to obtain a visual axis for the eye which may be selected as the gaze vector.

[0044] Different gaze detection techniques may be used within the same system. For example, due to obstructions of the eye or update processing time, less computationally intensive techniques, like a version of the approach based on correlating glint intensity values with pupil position, may be used more frequently in combination with more computationally intensive techniques run with longer time intervals in between like a version of determining the gaze vector based on the 3D spatial relationship between the cornea center, pupil center, center of rotation and a gaze detection system of optical elements. Changes in the spatial relationship including depth changes between the eye and the gaze detection elements can be determined also as an indicator triggering recalibration of the system, for example in embodiments using a training gaze data set.

[0045] In the embodiments discussed below, the see-through display device is in a set of eyeglasses but other head mounted display (HMD) formats and near-eye display holders suitable for consumer, everyday use can be used as well.

[0046] Figure 1A is a block diagram depicting example components of one embodiment of a mixed reality display system with gaze determination. System 10 includes a see-through display device as a near-eye, head mounted display device 2 in communication with processing unit 4 via wire 6. In other embodiments, head mounted display device 2 communicates with processing unit 4 via wireless communication. Processing unit 4 may take various embodiments. In some embodiments, processing unit 4 is a separate unit which may be worn on the user's body, e.g. the wrist in the illustrated example or in a pocket, and includes much of the computing power used to operate near-eye display device 2. Processing unit 4

52785-18(S)

10

may communicate wirelessly (e.g., WiFi, Bluetooth™, infra-red, or other wireless communication means) to one or more hub computing systems 12. In other embodiments, the functionality of the processing unit 4 may be integrated in software and hardware components of the display device 2.

[0047] Head mounted display device 2, which in one embodiment is in the shape of eyeglasses in a frame 115, is worn on the head of a user so that the user can see through a display, embodied in this example as a display optical system 14 for each eye, and thereby have an actual direct view of the space in front of the user. Frame 115 provides a support for holding elements of the system in place as well as a conduit for electrical connections. In this embodiment, frame 115 provides a convenient eyeglass frame as support for the elements of the system discussed further below. In other embodiments, other support structures can be used. An example of such a structure is a visor or goggles. The frame 115 includes a temple or side arm for resting on each of a user's ears. Temple 102 is representative of an embodiment of the right temple. Nose bridge 104 of the frame includes a microphone 110 for recording sounds and transmitting audio data to processing unit 4.

[0048] Hub computing system 12 may be a computer, a gaming system or console, or the like. According to an example embodiment, the hub computing system 12 may include hardware components and/or software components such that hub computing system 12 may be used to execute applications such as gaming applications, non-gaming applications, or the like. In one embodiment, hub computing system 12 may include a processor such as a standardized processor, a specialized processor, a microprocessor, or the like that may execute instructions stored on a processor readable storage device for performing the processes described herein.

52785-18(S)

11

[0049] Hub computing system 12 further includes one or more capture devices, such as capture devices 20A and 20B. In other embodiments, more or less than two capture devices can be used to capture the room or other physical environment of the user. Capture devices 20A and 20B may be, for example, cameras that visually monitor one or more users and the surrounding space such that gestures and/or movements performed by the one or more users, as well as the structure of the surrounding space, may be captured, analyzed, and tracked to perform one or more controls or actions within an application and/or animate an avatar or on-screen character. An application may be executing on hub computing system 12, the display device 2, as discussed below on a mobile device 5 or a combination of these.

[0050] Hub computing system 12 may be connected to an audiovisual device 16 such as a television, a monitor, a high-definition television (HDTV), or the like that may provide game or application visuals. In some instances, the audiovisual device 16 may be a three-dimensional display device. For example, hub computing system 12 may include a video adapter such as a graphics card and/or an audio adapter such as a sound card that may provide audiovisual signals associated with the game application, non-game application, etc. The audiovisual device 16 may receive the audiovisual signals from hub computing system 12 and may then output the game or application visuals and/or audio associated with the audiovisual signals. According to one embodiment, the audiovisual device 16 may be connected to hub computing system 12 via, for example, an S-Video cable, a coaxial cable, an HDMI cable, a DVI cable, a VGA cable, component video cable, RCA cables, etc. In one example, audiovisual device 16 includes internal speakers. In other embodiments, audiovisual device 16, a separate stereo or hub computing system 12 is connected to external speakers 22.

52785-18(S)

12

[0051] Figure 1B is a block diagram depicting example components of another embodiment of a mixed reality display system with gaze determination. In this embodiment, the near-eye display device 2 communicates with a mobile computing device 5 as an example embodiment of the processing unit 4. In the illustrated example, the mobile device 5 communicates via wire 6, but communication may also be wireless in other examples.

[0052] Furthermore, as in the hub computing system 12, gaming and non-gaming applications may execute on a processor of the mobile device 5 which user actions control or which user actions animate an avatar as may be displayed on a display 7 of the device 5. The mobile device 5 also provides a network interface for communicating with other computing devices like hub computing system 12 over the Internet or another communication network via a wired or wireless communication medium using a wired or wireless communication protocol. A remote network accessible computer system like hub computing system 12 may be leveraged for processing power and remote data access by a processing unit 4 like mobile device 5. Examples of hardware and software components of a mobile device 5 such as may be embodied in a smartphone or tablet computing device are described in Figure 16, and these components can embody the hardware and software components of a processing unit 4 such as those discussed in the embodiment of Figure 4A. Some other examples of mobile devices 5 are a laptop or notebook computer and a netbook computer.

[0053] As noted above, in some embodiments, gaze detection of each of a user's eyes is based on a three dimensional coordinate system of gaze detection elements on a near-eye, mixed reality display device like the eyeglasses 2 in relation to one or more human eye elements such as a cornea center, a center of eyeball rotation and a pupil center. Examples of gaze detection elements which

52785-18(S)

13

may be part of the coordinate system including glint generating illuminators and at least one sensor for capturing data representing the generated glints. As discussed in the embodiment of Figure 7, a center of the cornea can be determined based on two glints using planar geometry. The center of the cornea links the pupil center and the center of rotation of the eyeball, which may be treated as a fixed location for determining an optical axis of the user's eye at a certain gaze or viewing angle.

[0054] Figure 1C illustrates an exemplary arrangement of positions of respective sets of gaze detection elements in a gaze detection system for each eye positioned facing each respective eye by a see-through, near-eye, mixed reality display system embodied in a set of eyeglasses 2. What appears as a lens for each eye represents a display optical system 14 for each eye, e.g. 14r and 14l. A display optical system includes a see-through lens, e.g. 118 and 116 in Figures 3A-3D, as in an ordinary pair of glasses, but also contains optical elements (e.g. mirrors, filters) for seamlessly fusing virtual content with the actual and direct real world view seen through the lens 118, 116. A display optical system 14 has an optical axis which is generally in the center of the see-through lens 118, 116 in which light is generally collimated to provide a distortionless view. For example, when an eye care professional fits an ordinary pair of eyeglasses to a user's face, a goal is that the glasses sit on the user's nose at a position where each pupil is aligned with the center or optical axis of the respective lens resulting in generally collimated light reaching the user's eye for a clear or distortionless view.

[0055] In the example of Figure 1C, a detection area 139r, 139l of at least one sensor is aligned with the optical axis of its respective display optical system 14r, 14l so that the center of the detection area 139r, 139l is capturing light along the optical axis. If the display optical system 14 is aligned with the user's pupil, each detection area 139 of the respective sensor 134 is aligned with the user's pupil.

52785-18(S)

14

Reflected light of the detection area 139 is transferred via one or more optical elements to the actual image sensor 134 of the camera, in this example illustrated by dashed line as being inside the frame 115.

[0056] In one example, a visible light camera also commonly referred to as an RGB camera may be the sensor, and an example of an optical element or light directing element is a visible light reflecting mirror which is partially transmissive and partially reflective. The visible light camera provides image data of the pupil of the user's eye, while IR photodetectors 152 capture glints which are reflections in the IR portion of the spectrum. If a visible light camera is used, reflections of virtual images may appear in the eye data captured by the camera. An image filtering technique may be used to remove the virtual image reflections if desired. An IR camera is not sensitive to the virtual image reflections on the eye.

[0057] In other examples, the at least one sensor 134 is an IR camera or a position sensitive detector (PSD) to which IR radiation may be directed. For example, a hot reflecting surface may transmit visible light but reflect IR radiation. The IR radiation reflected from the eye may be from incident radiation of the illuminators 153, other IR illuminators (not shown) or from ambient IR radiation reflected off the eye. In some examples, sensor 134 may be a combination of an RGB and an IR camera, and the optical light directing elements may include a visible light reflecting or diverting element and an IR radiation reflecting or diverting element. In some examples, a camera may be small, e.g. 2 millimeters (mm) by 2mm. An example of such a camera sensor is the Omnivision OV7727. In other examples, the camera may be small enough, e.g. the Omnivision OV7727, e.g. that the image sensor or camera 134 may be centered on the optical axis or other location of the display optical system 14. For example, the camera 134 may be embedded within a lens of the system 14. Additionally, an image filtering

52785-18(S)

15

technique may be applied to blend the camera into a user field of view to lessen any distraction to the user.

[0058] In the example of Figure 1C, there are four sets of an illuminator 153 paired with a photodetector 152 and separated by a barrier 154 to avoid interference between the incident light generated by the illuminator 153 and the reflected light received at the photodetector 152. To avoid unnecessary clutter in the drawings, drawing numerals are shown with respect to a representative pair. Each illuminator may be an infra-red (IR) illuminator which generates a narrow beam of light at about a predetermined wavelength. Each of the photodetectors may be selected to capture light at about the predetermined wavelength. Infra-red may also include near-infrared. As there can be wavelength drift of an illuminator or photodetector or a small range about a wavelength may be acceptable, the illuminator and photodetector may have a tolerance range about a wavelength for generation and detection. In embodiments where the sensor is an IR camera or IR position sensitive detector (PSD), the photodetectors may be additional data capture devices and may also be used to monitor the operation of the illuminators, e.g. wavelength drift, beam width changes, etc. The photodetectors may also provide glint data with a visible light camera as the sensor 134.

[0059] As mentioned above, in some embodiments which calculate a cornea center as part of determining a gaze vector, two glints, and therefore two illuminators will suffice. However, other embodiments may use additional glints in determining a pupil position and hence a gaze vector. As eye data representing the glints is repeatedly captured, for example at 30 frames a second or greater, data for one glint may be blocked by an eyelid or even an eyelash, but data may be gathered by a glint generated by another illuminator.

[0060] Figure 1D illustrates another exemplary arrangement of positions of

52785-18(S)

16

respective sets of gaze detection elements in a gaze detection system for each eye positioned facing each respective eye on a mixed reality display device embodied in a set of eyeglasses. In this embodiment, two sets of illuminator 153 and photodetector 152 pairs are positioned near the top of each frame portion 115 surrounding a display optical system 14, and another two sets of illuminator and photodetector pairs are positioned near the bottom of each frame portion 115 for illustrating another example of a geometrical relationship between illuminators and hence the glints they generate. This arrangement of glints may provide more information on a pupil position in the vertical direction.

[0061] Figure 1E illustrates yet another exemplary arrangement of positions of respective sets of gaze detection elements in a gaze detection system for each eye positioned facing each respective eye by the set of eyeglasses. In this example, the sensor 134r, 134l is in line or aligned with the optical axis of its respective display optical system 14r, 14l but located on the frame 115 below the system 14. Additionally, in some embodiments, the camera 134 may be a depth camera or include a depth sensor. In this example, there are two sets of illuminators 153 and photodetectors 152.

[0062] Figure 2 is a side view of an eyeglass temple 102 of the frame 115 in an embodiment of the see-through, mixed reality display device embodied as eyeglasses providing support for hardware and software components. At the front of frame 115 is physical environment facing video camera 113 that can capture video and still images. Particularly in some embodiments where the display device 2 is not operating in conjunction with depth cameras like capture devices 20a and 20b of the hub system 12, the physical environment facing camera 113 may be a depth camera as well as a visible light sensitive camera. For example, the depth camera may include an IR illuminator transmitter and a hot reflecting surface like a

52785-18(S)

17

hot mirror in front of the visible image sensor which lets the visible light pass and directs reflected IR radiation within a wavelength range or about a predetermined wavelength transmitted by the illuminator to a CCD or other type of depth sensor. The data from the sensors may be sent to a processor 210 of the control circuitry 13, or the processing unit 4,5 or both which may process them but which the unit 4,5 may also send to hub computing system 12 in some embodiments like Figure 1A or over a network to one or more computer systems (e.g. like hub computing system 12) for processing. The processing identifies and maps the user's real world field of view. Additionally, the physical environment facing camera 113 may also include a light meter for measuring ambient light. A change of a certain amount may trigger a message for recalibration of training gaze data sets in some embodiments as discussed further below.

[0063] Control circuits 136 provide various electronics that support the other components of head mounted display device 2. More details of control circuits 136 are provided below with respect to Figure 4A. Inside, or mounted to temple 102, are ear phones 130, inertial sensors 132, GPS transceiver 144 and temperature sensor 138. In one embodiment inertial sensors 132 include a three axis magnetometer 132A, three axis gyro 132B and three axis accelerometer 132C (See Figure 4A). The inertial sensors are for sensing position, orientation, and sudden accelerations of head mounted display device 2. From these movements, head position may also be determined.

[0064] The display device 2 provides an image generation unit which can create one or more images including one or more virtual objects. In some embodiments, a microdisplay may be used as the image generation unit. A microdisplay assembly 173 comprises light processing elements and a variable focus adjuster 135. An example of a light processing element is a microdisplay

52785-18(S)

18

unit 120. Other examples include one or more optical elements such as one or more lenses of a lens system 122 and one or more reflecting elements such as surfaces 124a and 124b in Figures 3A and 3B or 124 in Figures 3C and 3D. Lens system 122 may comprise a single lens or a plurality of lenses.

[0065] Mounted to or inside temple 102, the microdisplay unit 120 includes an image source and generates an image of a virtual object. The microdisplay unit 120 is optically aligned with the lens system 122 and the reflecting surface 124 or reflecting surfaces 124a and 124b as illustrated in the following figures. The optical alignment may be along an optical axis 133 or an optical path 133 including one or more optical axes. The microdisplay unit 120 projects the image of the virtual object through lens system 122, which may direct the image light, onto reflecting element 124 which directs the light into lightguide optical element 112 as in Figures 3C and 3D or onto reflecting surface 124a (e.g. a mirror or other surface) which directs the light of the virtual image to a partially reflecting element 124b which combines the virtual image view along path 133 with the natural or actual direct view along the optical axis 142 as in Figures 3A – 3D. The combination of views are directed into a user's eye.

[0066] The variable focus adjuster 135 changes the displacement between one or more light processing elements in the optical path of the microdisplay assembly or an optical power of an element in the microdisplay assembly. The optical power of a lens is defined as the reciprocal of its focal length, e.g. $1/\text{focal length}$, so a change in one effects the other. The change in focal length results in a change in the region of the field of view, e.g. a region at a certain distance, which is in focus for an image generated by the microdisplay assembly 173.

[0067] In one example of the microdisplay assembly 173 making displacement changes, the displacement changes are guided within an armature 137 supporting at

52785-18(S)

19

least one light processing element such as the lens system 122 and the microdisplay 120 in this example. The armature 137 helps stabilize the alignment along the optical path 133 during physical movement of the elements to achieve a selected displacement or optical power. In some examples, the adjuster 135 may move one or more optical elements such as a lens in lens system 122 within the armature 137. In other examples, the armature may have grooves or space in the area around a light processing element so it slides over the element, for example, microdisplay 120, without moving the light processing element. Another element in the armature such as the lens system 122 is attached so that the system 122 or a lens within slides or moves with the moving armature 137. The displacement range is typically on the order of a few millimeters (mm). In one example, the range is 1-2 mm. In other examples, the armature 137 may provide support to the lens system 122 for focal adjustment techniques involving adjustment of other physical parameters than displacement. An example of such a parameter is polarization.

[0068] For more information on adjusting a focal distance of a microdisplay assembly, see U.S. patent no. 12/941,825 entitled “Automatic Variable Virtual Focus for Augmented Reality Displays,” filed November 8, 2010, having inventors Avi Bar-Zeev and John Lewis.

[0069] In one example, the adjuster 135 may be an actuator such as a piezoelectric motor. Other technologies for the actuator may also be used and some examples of such technologies are a voice coil formed of a coil and a permanent magnet, a magnetostriction element, and an electrostriction element.

[0070] There are different image generation technologies that can be used to implement microdisplay 120. For example, microdisplay 120 can be implemented using a transmissive projection technology where the light source is modulated by optically active material, backlit with white light. These technologies are usually

52785-18(S)

20

implemented using LCD type displays with powerful backlights and high optical energy densities. Microdisplay 120 can also be implemented using a reflective technology for which external light is reflected and modulated by an optically active material. The illumination is forward lit by either a white source or RGB source, depending on the technology. Digital light processing (DLP), liquid crystal on silicon (LCOS) and Mirasol® display technology from Qualcomm, Inc. are all examples of reflective technologies which are efficient as most energy is reflected away from the modulated structure and may be used in the system described herein. Additionally, microdisplay 120 can be implemented using an emissive technology where light is generated by the display. For example, a PicoP™ engine from Microvision, Inc. emits a laser signal with a micro mirror steering either onto a tiny screen that acts as a transmissive element or beamed directly into the eye (e.g., laser).

[0071] As mentioned above, the configuration of the light processing elements of the microdisplay assembly 173 create a focal distance or focal region in which a virtual object appears in an image. Changing the configuration changes the focal region for the virtual object image. The focal region determined by the light processing elements can be determined and changed based on the equation $1/S1 + 1/S2 = 1/f$.

[0072] The symbol f represents the focal length of a lens such as lens system 122 in the microdisplay assembly 173. The lens system 122 has a front nodal point and a rear nodal point. If light rays are directed toward either nodal point at a given angle relative to the optical axis, the light rays will emerge from the other nodal point at an equivalent angle relative to the optical axis. In one example, the rear nodal point of lens system 122 would be between itself and the microdisplay 120. The distance from the rear nodal point to the microdisplay 120 may be denoted as

52785-18(S)

21

S2. The front nodal point is typically within a few mm of lens system 122. The target location is the location of the virtual image to be generated by the microdisplay 120 in a three-dimensional physical space. The distance from the front nodal point to the target location of the virtual image may be denoted as S1. Since the image is to be a virtual image appearing on the same side of the lens as the microdisplay 120, sign conventions give that S1 has a negative value.

[0073] If the focal length of the lens is fixed, S1 and S2 are varied to focus virtual objects at different depths. For example, an initial position may have S1 set to infinity, and S2 equal to the focal length of lens system 122. Assuming lens system 122 has a focal length of 10mm, consider an example in which the virtual object is to be placed about 1 foot or 300 mm into the user's field of view. S1 is now about -300mm, f is 10mm and S2 is set currently at the initial position of the focal length, 10mm, meaning the rear nodal point of lens system 122 is 10 mm from the microdisplay 120. The new distance or new displacement between the lens 122 and microdisplay 120 is determined based on $1/(-300) + 1/S2 = 1/10$ with all in units of mm. The result is about 9.67 mm for S2.

[0074] In one example, one or more processors such as in the control circuitry, the processing unit 4, 5 or both can calculate the displacement values for S1 and S2, leaving the focal length f fixed and cause the control circuitry 136 to cause a variable adjuster driver 237 (see Figure 4A) to send drive signals to have the variable virtual focus adjuster 135 move the lens system 122 along the optical path 133 for example. In other embodiments, the microdisplay unit 120 may be moved instead or in addition to moving the lens system 122. In other embodiments, the focal length of at least one lens in the lens system 122 may be changed instead or with changes in the displacement along the optical path 133 as well.

[0075] Figure 3A is a top view of an embodiment of a display optical system

52785-18(S)

22

14 of a see-through, near-eye, mixed reality device including an arrangement of gaze detection elements in a gaze detection system. A portion of the frame 115 of the near-eye display device 2 will surround a display optical system including providing support for one or more lenses as illustrated. In order to show the components of the display system 14, in this case 14r for the right eye system, a top portion of the frame 115 surrounding the display optical system is not depicted.

[0076] The display optical system 14 in this embodiment has an optical axis 142 and includes a see-through lens 118 allowing the user an actual direct view of the real world. In this example, the see-through lens 118 is a standard lens used in eye glasses and can be made to any prescription (including no prescription). In another embodiment, see-through lens 118 can be replaced by a variable prescription lens. In some embodiments, see-through, near-eye display device 2 will include additional lenses.

[0077] The display optical system 14 further comprises reflecting surfaces 124a and 124b. In this embodiment, light from the microdisplay 120 is directed along optical path 133 via a reflecting element 124a to a partially reflective element 124b embedded in lens 118 which combines the virtual object image view traveling along optical path 133 with the natural or actual direct view along the optical axis 142 so that the combined views are directed into a user's eye, right one in this example, at the optical axis, the position with the most collimated light for a clearest view.

[0078] A detection area 139r of a light sensor is also part of the display optical system 14r. An optical element 125 embodies the detection area 139r by capturing reflected light from the user's eye received along the optical axis 142 and directs the captured light to the sensor 134r, in this example positioned in the bridge 104.

52785-18(S)

23

As shown, the arrangement allows the detection area 139 of the sensor 134r to have its center aligned with the center of the display optical system 14. For example, if sensor 134r is an image sensor, sensor 134r captures the detection area 139, so an image captured at the image sensor is centered on the optical axis because the detection area 139 is. In one example, sensor 134 r is a visible light camera or a combination of RGB/IR camera, and the optical element 125 includes an optical element which reflects visible light reflected from the user's eye, for example a partially reflective mirror.

[0079] In other embodiments, the sensor 134r is an IR sensitive device such as an IR camera, and the element 125 includes a hot reflecting surface which lets visible light pass through it and reflects IR radiation to the sensor 134r. An IR camera may capture not only glints, but also an infra-red or near-infra-red image of the user's eye including the pupil.

[0080] In other embodiments, the IR sensor device 134r is a position sensitive device (PSD), sometimes referred to as an optical position sensor. The position of detected light on the surface of the sensor is identified. A PSD can be selected which is sensitive to a wavelength range or about a predetermined wavelength of IR illuminators for the glints. When light within the wavelength range or about the predetermined wavelength of the position sensitive device is detected on the sensor or light sensitive portion of the device, an electrical signal is generated which identifies the location on the surface of the detector. In some embodiments, the surface of a PSD is divided into discrete sensors like pixels from which the location of the light can be determined. In other examples, a PSD isotropic sensor may be used in which a change in local resistance on the surface can be used to identify the location of the light spot on the PSD. Other embodiments of PSDs may also be used. By operating the illuminators 153 in a predetermined sequence, the location

52785-18(S)

24

of the reflection of glints on the PSD can be identified and hence related back to their location on a cornea surface.

[0081] The depiction of the light directing elements, in this case reflecting elements, 125, 124, 124a and 124b in Figures 3A-3D are representative of their functions. The elements may take any number of forms and be implemented with one or more optical components in one or more arrangements for directing light to its intended destination such as a camera sensor or a user's eye.

[0082] The display optical system 14 includes other gaze detection elements in this embodiment. In this embodiment, attached to frame 115 and on the sides of lens 118, are at least two (2) but may be more, infra-red (IR) illuminating devices 153 which direct narrow infra-red light beams within a particular wavelength range or about a predetermined wavelength at the user's eye to each generate a respective glint on a surface of the respective cornea. In other embodiments, the illuminators and any photodiodes may be on the lenses, for example at the corners or edges. In this embodiment, in addition to the at least 2 infra-red (IR) illuminating device 153 are IR photodetectors 152. Each photodetector 152 is sensitive to IR radiation within the particular wavelength range of its corresponding IR illuminator 153 across the lens 118 and is positioned to detect a respective glint. As shown in Figures 1C-1E, the illuminator and photodetector are separated by a barrier 154 so that incident IR light from the illuminator 153 does not interfere with reflected IR light being received at the photodetector 152. In the case where the sensor 134 is an IR sensor, the photodetectors 152 may not be needed or may be an additional glint data capture source. With a visible light camera, the photodetectors 152 capture light from glints and generate glint intensity values.

[0083] Figure 3B is a top view of another embodiment of a display optical system 14 of a see-through, near-eye, mixed reality device including an

52785-18(S)

25

arrangement of gaze detection elements in a gaze detection system. In this embodiment, in addition to the at least 2 infra-red (IR) illuminating devices 153 are IR photodetectors 152. In this embodiment, the hot reflecting surface 125 has been removed to show operation without a position sensitive detector.

[0084] In the embodiment of Figure 3B, light detector 134r may be embodied as a visible light camera, sometimes referred to as an RGB camera, or it may be embodied as an IR camera or a camera capable of processing light in both the visible and IR ranges e.g. a depth camera. In this example, the image sensor 134r is the detection area 139r, and the image sensor 134 of the camera is located vertically on the optical axis 142 of the display optical system. In some examples, the camera may be located on frame 115 either above or below see-through lens 118 or embedded in the lens 118. In some embodiments, the illuminators 153 provide light for the camera, and in other embodiments the camera captures images with ambient lighting or light from its own light source.

[0085] In one embodiment, glint reflections can estimate gaze based on a few data points of the intensity values detected for the glints, rather than processing much, much larger sets of image data of eyes. The position of the illuminators 153 on the eyeglass frame 115 or other support structure of a near-eye display device may be fixed so that the position of glints detected by one or more sensors is fixed in the sensor detection area. The cornea and hence the iris and the pupil rotate with the eyeball about a fixed center. The iris, pupil, and the sclera which is sometimes referred to as the white portion of the eyeball, move underneath the glint as the user's gaze changes. So a glint detected at a same sensor location may result in different intensity values due to different reflectivities associated with the different eye parts. As the pupil is a hole with tissue that absorbs most incoming light, the intensity value for it would be very low or near zero, while that for the iris would

52785-18(S)

26

be a higher intensity value due to its higher reflectivity. An intensity value for the sclera may be highest as the sclera has the highest reflectivity. In some examples, an illuminator may be positioned as in Figures 3A through 3D on either side of the display optical system 14 and hence on either side of the pupil of the user's eye. In other embodiments, additional illuminators may be positioned on the frame 115 or lens 118, for example, four illuminators may be positioned to generate a surrounding geometric shape, e.g. a box, of glints on the eyeball which would be approximately centered on the pupil when a user is looking straight ahead. The microdisplay assembly 173 can display a virtual image or send a message, e.g. a visual virtual image or an audio instruction to a user to cause the user to look straight ahead for initializing the glints on or near the pupil. In other embodiments, gaze detection based on glints is based on intensity values generated from illuminators with the glint positioning being independent of being centered on the pupil.

[0086] Figure 3C is a top view of a third embodiment of a display optical system 14 of a see-through, near-eye, mixed reality device including an arrangement of gaze detection elements in a gaze detection system. The display includes a light guide optical element 112 between an additional see-through lens 116 and see-through lens 118. Lightguide optical element 112 channels artificial light to the eye.

[0087] Lightguide optical element 112 transmits light from microdisplay 120 to the eye of the user wearing head mounted display device 2. Lightguide optical element 112 also allows light from in front of the head mounted display device 2 to be transmitted through lightguide optical element 112 to the user's eye thereby allowing the user to have an actual direct view of the space in front of head mounted display device 2 in addition to receiving a virtual image from

52785-18(S)

27

microdisplay 120. Thus, the walls of lightguide optical element 112 are see-through. Lightguide optical element 112 includes a first reflecting surface 124 (e.g., a mirror or other surface). Light from microdisplay 120 passes through lens 122 and becomes incident on reflecting surface 124. The reflecting surface 124 reflects the incident light from the microdisplay 120 such that light is trapped inside a planar, substrate comprising lightguide optical element 112 by internal reflection.

[0088] After several reflections off the surfaces of the substrate, the trapped light waves reach an array of selectively reflecting surfaces 126. Note that only one of the five surfaces is labeled 126 to prevent over-crowding of the drawing. Reflecting surfaces 126 couple the light waves incident upon those reflecting surfaces out of the substrate into the eye of the user. More details of a lightguide optical element can be found in United States Patent Application Publication 2008/0285140, Serial No. 12/214,366, published on November 20, 2008, "Substrate-Guided Optical Devices".

[0089] In this embodiment, as in Figure 1E and one of the examples for Figure 3B, the display optical system 14 is similarly arranged with IR illuminators 153 and photodetectors 152, and a visible light or IR camera 134r located on the frame 115 or lens 118 below or above optical axis 142, typically at a center of lenses 116 and 118 supporting the lightguide optical element 112.

[0090] Figure 3D is a top view of a fourth embodiment of a display optical system 14 of a see-through, near-eye, mixed reality device including an arrangement of gaze detection elements in a gaze detection system. This embodiment is similar to Figure 3C's embodiment including a light guide optical element 112. However, the only light detectors are the IR photodetectors 152, so this embodiment relies on glint detection only for gaze detection as discussed in the examples below.

52785-18(S)

28

[0091] In the embodiments of Figures 3A-3D, the positions of the gaze detection elements, e.g. the detection area 139 and the illuminators 153 and photodetectors 152 are fixed with respect to each other. In these examples, they are also fixed in relation to the optical axis of the display optical system 14.

[0092] In the embodiments above, the specific number of lenses shown are just examples. Other numbers and configurations of lenses operating on the same principles may be used. Additionally, in the examples above, only the right side of the see-through, near-eye display 2 are shown. A full near-eye, mixed reality display device would include as examples another set of lenses 116 and/or 118, another lightguide optical element 112 for the embodiments of Figures 3C and 3D, another micro display 120, another lens system 122, likely another environment facing camera 113, another eye tracking camera 134 for the embodiments of Figures 3A to 3C, earphones 130, and a temperature sensor 138.

[0093] Figure 4A is a block diagram of one embodiment of hardware and software components of a see-through, near-eye, mixed reality display unit as may be used with the embodiments described in this disclosure. Figure 4B is a block diagram describing the various components of processing unit 4. In this embodiment, near-eye display device 2, receive instructions about a virtual image from processing unit 4 and provides the sensor information back to processing unit 4. Processing unit 4, the components of which are depicted in Figure 4B, will receive the sensory information from the display device 2 and may also receive sensory information from hub computing device 12 (See Figure 1). Based on that information, processing unit 4 will determine where and when to provide a virtual image to the user and send instructions accordingly to the control circuitry 136 of the display device 2.

[0094] Note that some of the components of Figure 4A (e.g., physical

52785-18(S)

29

environment facing camera 113, eye camera 134, variable virtual focus adjuster 135, photodetector interface 139, micro display 120, illumination device 153 or illuminators, earphones 130, and temperature sensor 138) are shown in shadow to indicate that there are two of each of those devices, one for the left side and one for the right side of head mounted display device 2. Figure 4A shows the control circuit 200 in communication with the power management circuit 202. Control circuit 200 includes processor 210, memory controller 212 in communication with memory 214 (e.g., D-RAM), camera interface 216, camera buffer 218, display driver 220, display formatter 222, timing generator 226, display out interface 228, and display in interface 230. In one embodiment, all of components of control circuit 220 are in communication with each other via dedicated lines of one or more buses. In another embodiment, each of the components of control circuit 200 are in communication with processor 210.

[0095] Camera interface 216 provides an interface to the two physical environment facing cameras 113 and each eye camera 134 and stores respective images received from the cameras 113, 134 in camera buffer 218. Display driver 220 will drive microdisplay 120. Display formatter 222 may provide information, about the virtual image being displayed on microdisplay 120 to one or more processors of one or more computer systems, e.g. 4, 12, 210 performing processing for the augmented reality system. Timing generator 226 is used to provide timing data for the system. Display out 228 is a buffer for providing images from physical environment facing cameras 113 and the eye cameras 134 to the processing unit 4. Display in 230 is a buffer for receiving images such as a virtual image to be displayed on microdisplay 120. Display out 228 and display in 230 communicate with band interface 232 which is an interface to processing unit 4.

[0096] Power management circuit 202 includes voltage regulator 234, eye

52785-18(S)

30

tracking illumination driver 236, variable adjuster driver 237, photodetector interface 239, audio DAC and amplifier 238, microphone preamplifier and audio ADC 240, temperature sensor interface 242 and clock generator 244. Voltage regulator 234 receives power from processing unit 4 via band interface 232 and provides that power to the other components of head mounted display device 2. Illumination driver 236 controls, for example via a drive current or voltage, the illumination devices 153 to operate about a predetermined wavelength or within a wavelength range. Audio DAC and amplifier 238 receives the audio information from earphones 130. Microphone preamplifier and audio ADC 240 provides an interface for microphone 110. Temperature sensor interface 242 is an interface for temperature sensor 138. Power management unit 202 also provides power and receives data back from three axis magnetometer 132A, three axis gyro 132B and three axis accelerometer 132C. Power management unit 202 also provides power and receives data back from and sends data to GPS transceiver 144.

[0097] The variable adjuster driver 237 provides a control signal, for example a drive current or a drive voltage, to the adjuster 135 to move one or more elements of the microdisplay assembly 173 to achieve a displacement for a focal region calculated by software executing in the a processor 210 of the control circuitry 13, or the processing unit 4,5 or the hub computer 12 or both. In embodiments of sweeping through a range of displacements and, hence, a range of focal regions, the variable adjuster driver 237 receives timing signals from the timing generator 226, or alternatively, the clock generator 244 to operate at a programmed rate or frequency.

[0098] The photodetector interface 239 performs any analog to digital conversion needed for voltage or current readings from each photodetector, stores the readings in a processor readable format in memory via the memory controller

52785-18(S)

31

212, and monitors the operation parameters of the photodetectors 152 such as temperature and wavelength accuracy.

[0099] Figure 4B is a block diagram of one embodiment of the hardware and software components of a processing unit 4, 5 associated with a see-through, near-eye display unit. The mobile device 5 may include this embodiment of hardware and software components as well or similar components which perform similar functions. Figure 4B shows controls circuit 304 in communication with power management circuit 306. Control circuit 304 includes a central processing unit (CPU) 320, graphics processing unit (GPU) 322, cache 324, RAM 326, memory control 328 in communication with memory 330 (e.g., D-RAM), flash memory controller 332 in communication with flash memory 334 (or other type of non-volatile storage), display out buffer 336 in communication with see-through, near-eye display device 2 via band interface 302 and band interface 232, display in buffer 338 in communication with near-eye display device 2 via band interface 302 and band interface 232, microphone interface 340 in communication with an external microphone connector 342 for connecting to a microphone, PCI express interface for connecting to a wireless communication device 346, and USB port(s) 348.

[00100] In one embodiment, wireless communication component 346 can include a Wi-Fi enabled communication device, Bluetooth communication device, infrared communication device, etc. The USB port can be used to dock the processing unit 4, 5 to hub computing device 12 in order to load data or software onto processing unit 4, 5, as well as charge processing unit 4, 5. In one embodiment, CPU 320 and GPU 322 are the main workhorses for determining where, when and how to insert virtual images into the view of the user.

[00101] Power management circuit 306 includes clock generator 360, analog to

52785-18(S)

32

digital converter 362, battery charger 364, voltage regulator 366, see-through, near-eye display power source 376, and temperature sensor interface 372 in communication with temperature sensor 374 (located on the wrist band of processing unit 4). An alternating current to direct current converter 362 is connected to a charging jack 370 for receiving an AC supply and creating a DC supply for the system. Voltage regulator 366 is in communication with battery 368 for supplying power to the system. Battery charger 364 is used to charge battery 368 (via voltage regulator 366) upon receiving power from charging jack 370. Device power interface 376 provides power to the display device 2.

[00102] Figure 5 is a top view illustrating examples of gaze vectors intersecting at a point of gaze where a user's eyes are focused. A model of the eye 160l, 160r is illustrated for each eye based on the Gullstrand schematic eye model. For each eye, an eyeball 160 is modeled as a sphere with a center of rotation 166 and includes a cornea 168 modeled as a sphere too and having a center 164. The cornea rotates with the eyeball, and the center 166 of rotation of the eyeball may be treated as a fixed point. The cornea covers an iris 170 with a pupil 162 at its center. In this example, on the surface 172 of the respective cornea are glints 174 and 176.

[00103] The axis 178 formed from the center of rotation 166 through the cornea center 164 to the pupil 162 is the optical axis of the eye. A gaze vector 180 is sometimes referred to as the line of sight or visual axis which extends from the fovea through the center of the pupil 162. The fovea is a small area of about 1.2 degrees located in the retina. The angular offset between the optical axis computed in the embodiment of Figure 9 and the visual axes has horizontal and vertical components. The horizontal component is up to 5 degrees from the optical axis, and the vertical component is between 2 and 3 degrees. In many embodiments, the optical axis is determined and a small correction determined through user

52785-18(S)

33

calibration is applied to obtain the visual axis which is selected as the gaze vector. For each user, a small virtual object may be displayed by the display device at each of a number of predetermined positions at different horizontal and vertical positions. An optical axis may be computed for during display of the object at each position, and a ray modeled as extending from the position into the user eye. An offset angle with horizontal and vertical components may be determined based on how the optical axis must be moved to align with the modeled ray. From the different positions, an average offset angle with horizontal or vertical components can be selected as the small correction to be applied to each computed optical axis. In some embodiments, only a horizontal component is used for the offset angle correction.

[00104] In the illustrated embodiment of Figure 5, a sensor detection area 139 is aligned with the optical axis of each display optical system 14 within an eyeglass frame 115. The respective image sensor in this example is a camera capable of capturing image data representing glints 174l and 176l generated respectively by illuminators 153a and 153b on the left side of the frame 115 and data representing glints 174r and 176r generated respectively by illuminators 153c and 153d.

[00105] Through the display optical systems, 14l and 14r in the eyeglass frame 115, the user's field of view includes both real objects 190, 192 and 194 and virtual objects 182, 184, and 186. In this example, the cornea 168l of the left eye is rotated to the right or towards the user's nose, and the cornea 168r of the right eye is rotated to the left or towards the user's nose. Both pupils are gazing at a virtual object 186. Gaze vectors 180l and 180r from each eye enter the Panum's fusional region 195 in which virtual object 186 is located. The Panum's fusional region is the area of single vision in a binocular viewing system like that of human vision. The intersection of the gaze vectors 180l and 180r indicates that the user is looking

52785-18(S)

at virtual object 186.

[00106] For a see-through mixed reality display device, the gaze vectors are determined to identify a point of gaze in a three-dimensional (3D) user field of view which includes both real objects, typically not under computer control, and virtual objects generated by an application. The gaze vectors may intersect at an object 10 feet away or at a distance effectively at infinity. The following figures briefly discuss embodiments for determining a 3D user field of view.

[00107] References to front facing image data are referring to image data from one or more front facing camera like camera 113 in Figures 1A and 1B. In these embodiments, the field of view of the front facing cameras 113 approximates the user field of view as the camera is located at a relatively small offset from the optical axis 142 of each display optical system 14. The offset may be taken into account in the image data.

[00108] Figure 6A is a flowchart of a method embodiment for determining a three-dimensional user field of view. In step 510, one or more processors of the control circuitry 136, the processing unit 4,5, the hub computing system 12 or a combination of these receive image data from one or more front facing cameras, and in step 512 identify one or more real objects in front facing image data. Data from the orientation sensor 132, e.g. the three axis accelerometer 132C and the three axis magnetometer 132A, can also be used with the front facing camera 113 image data for mapping what is around the user, the position of the user's face and head in order to determine which objects, real or virtual, he or she is likely focusing on at the time. Based on an executing application, the one or more processors in step 514 identify virtual object positions in a user field of view which may be determined to be the field of view captured in the front facing image data. In step 516, a three-dimensional position is determined for each object in the user field of

52785-18(S)

35

view. In other words, where each object is located with respect to the display device 2, for example with respect to the optical axis 142 of each display optical system 14.

[00109] Figure 6B is a flowchart of a method embodiment for identifying one or more real objects in a user field of view. This embodiment may be used to implement step 512. Each of the implementing examples in Figures 6B, 6D and 6E may be used separately or in conjunction with one another to identify the location of objects in the user field of view. In step 520, a location of user wearing the display device 2 is identified. For example, GPS data via a GPS unit 965 in the mobile device 5 or GPS transceiver 144 on the display device 2 may identify the location of the user. In step 522, one or more processors, retrieve one or more images of the location from a database (e.g. 470), and uses pattern recognition in step 524 to select one or more images matching image data from the one or more front facing cameras. In some embodiments, steps 522 and 524 may be performed remotely by a more powerful computer, e.g. hub 12, having access to image databases. Based on GPS data, in step 526 the one or more processors determines a relative position of one or more objects in front facing image data to one or more GPS tracked objects 528 in the location, and determines in step 529 a position of user from the one or more real objects based on the one or more relative positions.

[00110] In some embodiments such as in Figure 1A, a user wearing a see-through, near-eye display may be in a location in which a computer system or one or more computers provides a three-dimensional mapping of objects within a space, e.g. a store. Figure 6C is a flowchart of a method embodiment for generating a three-dimensional model of a user space. In step 530, a computer system with access to depth cameras like hub system 12 with capture devices 20A and 20B creates a three-dimensional model of a space based on depth images. The depth

52785-18(S)

36

images may be from multiple perspectives and may be combined based on a common coordinate space, e.g. the store space, and creates a volumetric or three dimensional description of the space. In step 532, objects are detected in the space. For example, edge detection may be performed on the depth images to distinguish objects, including people, from each other. In step 534, the computer system 12 identifies one or more detected objects including their positions in the space. The objects may also be identified based on comparisons of shape and pattern recognition techniques including facial recognition techniques with reference images of things and people from image databases.

[00111] Figure 6D is a flowchart of a method embodiment for identifying one or more objects in a user field of view based on depth data transmitted to the see-through, mixed reality display device 2. The processing unit 4,5 in step 540 sends front facing image data to a three-dimensional modeling system such as may be implemented by a depth image processing application executing on a computer system like hub computing system 12 communicatively coupled to depth cameras 20A and 20B. Data from the orientation sensor 132 may also be sent for identifying face or head position. For example, when a user enters a store, a computer system at the store provides a 3D mapping of the store and what and who is in it. In step 542, the display device 2 receives data identifying one or more objects in a field of view for the user and their positions in a 3D model of a space. The image data from the one or more front facing cameras 113 approximates the user field of view, so the hub system 12 identifies the object in the front facing image data, for example through image recognition or pattern recognition software. Orientation data may also be used with the front facing image data to refine the user field of view and identify objects tracked by the computer system 12 falling within the user field of view. (The hub system 12 also aligns the front facing image data when received from two or more cameras 113 for identifying the user field of

52785-18(S)

37

view.) The processing unit 4,5 in step 544 receives a position of the user in the 3D model of the space, and in step 546 the processing unit 4,5, or the processor 210 of the control circuitry 136 or both determines a position of one or more objects in the user field of view based on the positions of the user and the one or more objects in the 3D model of the space. In another example, the processing unit 4,5 receives the position of the user and the one or more objects as determined by the computer system 12.

[00112] Figure 6E is a flowchart of a method embodiment for identifying one or more objects in a user field of view when the front facing camera 113 is a depth camera providing depth image data or has a depth sensor for providing depth data which can be combined with image data to provide depth image data. In step 550, the one or more processors of the display device 2, e.g. processor 210 of the control circuitry or the processing unit 4,5, or both identifies one or more real objects in a user field of view including their three-dimensional positions based on depth image data from one or more front facing cameras. The one or more processors may also map the user field of view based on orientation data from an orientation sensor 132 in addition to the image data. The one or more processors perform step 514 of identifying virtual object positions in the user field of view based on an executing application and step 516 of determining a three-dimensional position of each object in the user field of view. Additionally, a remote computer system 12 may also providing additional processing power to the other processors for performing the steps of Figure 6E.

[00113] Each of the method embodiments of Figures 6A through 6E are typically performed repeatedly as the user and objects within the user's environment move around.

[00114] Figure 6F is a block diagram of a system embodiment for determining

52785-18(S)

38

positions of objects within a user field of view of a see-through, near-eye display device. This embodiment illustrates how the various devices may leverage networked computers to map a three-dimensional model of a user field of view and the real and virtual objects within the model. An application 456 executing in a processing unit 4,5 communicatively coupled to a display device 2 can communicate over one or more communication networks 50 with a computing system 12 for processing of image data to determine and track a user field of view in three dimensions. The computing system 12 may be executing an application 452 remotely for the processing unit 4,5 for providing images of one or more virtual objects. Either or both of the applications 456 and 452 working together may map a 3D model of space around the user. A depth image processing application 450 detects objects, identifies objects and their locations in the model. The application 450 may perform its processing based on depth image data from depth camera like 20A and 20B, two-dimensional or depth image data from one or more front facing cameras 113, and GPS metadata associated with objects in the image data obtained from a GPS image tracking application 454.

[00115] The GPS image tracking application 454 identifies images of the user's location in one or more image database(s) 470 based on GPS data received from the processing unit 4,5 or other GPS units identified as being within a vicinity of the user, or both. Additionally, the image database(s) may provide accessible images of a location with metadata like GPS data and identifying data uploaded by users who wish to share their images. The GPS image tracking application provides distances between objects in an image based on GPS data to the depth image processing application 450. Additionally, the application 456 may perform processing for mapping and locating objects in a 3D user space locally and may interact with the GPS image tracking application for receiving distances between objects. Many combinations of shared processing are possible between the

52785-18(S)

39

applications by leveraging network connectivity.

[00116] Figure 7 is a flowchart of a method embodiment for determining gaze in a see-through, near-eye mixed reality display system and provides an overall view of how a near-eye display device can leverage its geometry of optical components to determine gaze. One or more processors such as that in processing unit 4, the mobile device 5, the control circuitry 136, or the hub computing system 12 alone or in combination 12 determine in step 602 boundaries for a gaze detection coordinate system. In step 604, a gaze vector for each eye is determined based on reflected eye data including glints, and in step 606 a point of gaze, e.g. what the user is looking at, is determined for the two eyes in a three-dimensional (3D) user field of view. As the positions and identity of objects in the user field of view are tracked, for example, by embodiments like in Figures 6A-6F, in step 608, any object at the point of gaze in the 3D user field of view is identified. In many embodiments, the three-dimensional user field of view includes displayed virtual objects and an actual direct view of real objects. The term object includes a person.

[00117] The method embodiment in Figure 7 and other method embodiments discussed below which use glint data for other ways of detecting gaze, may identify such glints from image data of the eye. When IR illuminators are used, typically an IR image sensor is used as well. The following method may also work with a discrete surface position sensitive detector (PSD), e.g. one with pixels. Figure 8 is a flowchart of a method embodiment for identifying glints in image data. As noted above, a glint is a very small and a very bright reflection of light from a light source off of a specularly reflective surface such as the cornea of an eye. In the method embodiment below, each of the steps is performed for a data sample set. In some examples, that may include data from one image or image frame, and in others, the data sample set may be for a number of images or image frames. In step

52785-18(S)

40

605, the processor identifies each connected set of pixels having their intensity values within a predetermined intensity range, for example, the range of intensity values may begin at 220 and end at the brightest pixel value 255. In step 607, the candidate glints are pruned by identifying as a candidate glint each connected set of pixels which satisfies glint geometry criteria. An example of glint geometry criteria is size and shape for the glints. Some may be too large, too small, or have too irregular a shape. Furthermore, the illuminators are positioned for the resulting glints to have a spatial or geometric relationship to each other. For example, the illuminators 153 are arranged for the glints to form a rectangle. In the embodiment discussed in Figure 9 in which a pupil center is determined from image data as well, a spatial relationship to the pupil may also be a criteria, e.g. a distance too far from the pupil may indicate a connected set is not a candidate glint.

[00118] In step 609, the one or more processors determine whether there are less candidate glints than a predetermined number. For example, for four illuminators, four glints are expected but the predetermined number may be two. In the example of the rectangle as the geometric relationship, two glints which form a horizontal line or a diagonal line of a predetermined length may have been selected as candidates. There may be an eyelid or eyelash obstruction for the other glints. If there are less than the predetermined number of glints, the data sample set is dropped for further processing, and processing returns in step 611 to step 605 of a next data sample set. If there are not less candidates than a predetermined number, then step 613 determines whether there are more candidate glints than a predetermined number. If there are more candidates, in step 615, the one or more processors select as glints the predetermined number of candidates which most closely fit the predetermined geometrical relationship between the glints. For example, for the rectangle, which candidates most closely form the rectangle of the predetermined size and shape. If there are not more candidates than the number,

52785-18(S)

41

the number of candidates matches the predetermined number of glints, and the candidates are selected as the glints in step 617.

[00119] Due to the geometry of the placement of illuminators for generating the glints as discussed above, the glints appear in the same locations, barring movement of the frame 115 with respect to the eye. Furthermore, as the positioning of the illuminators with respect to each other on the support structure of the frame 115 or lens 118 is fixed, the spatial relationship of the glints to each other in the image is fixed as well. As for size, as the glints are very small, the number of pixels making up the glint area on the sensor and in the sensed image would be correspondingly small. For example, if the image sensor of the camera has a 1000 pixels, each glint may take up less than ten pixels. Glints may be monitored in each image frame taken for example at 30 or 60 frames a second and an area may be identified as a glint from a number of frame samples. There may not be glint data in every frame. Sampling accommodates or smoothes out obstructions of glint, and pupil data, in different image frames such as due to factors like an eyelid or eyelash covering the glint and/or pupil. An image frame is an example of an image format.

[00120] Figure 9 is a flowchart of a method embodiment which may be used to implement step 602 of determining boundaries for a gaze detection coordinate system. One or more processors determines a position of a center 164 of a cornea of each eye with respect to the illuminators 153 and at least one light sensor, e.g. 134 or 152, based on glints in step 612. Based on image data provided by the at least one sensor, in step 614, the one or more processors determine a pupil center of each eye. In step 616, the position of the center of eyeball rotation, which may be treated as fixed, is determined relative to the cornea and pupil centers. For example, based on the pupil center, a ray can be extended back through the

52785-18(S)

42

determined cornea center 164 to the fixed center 166 of eyeball rotation. Additionally, distance or length approximations are used for approximating the length on the optical axis between the pupil and the cornea, for example about 3 mm, and the length on the optical axis between the center of curvature of cornea and the center of eyeball rotation, about 6 mm. These values have been determined from population studies of human eye parameters such as those compiled by Gullstrand. (See *Hennessey*, p. 88).

[00121] Optionally, the one or more processors in step 618 determines a position of the fixed center of eyeball rotation with respect to the illuminators and the at least one sensor for the respective eye. This position determined in step 618 provides a depth distance between a fixed point, or one that can be approximated as fixed for accuracy considerations of gaze detection, and the display optical system. In effect, a depth axis has been defined for the gaze detection coordinate system. Changes detected along the depth axis may be used to indicate that the near-eye display system has moved and trigger determination of boundaries of the coordinate system again or re-calibration of training gaze data sets as discussed below.

[00122] Figure 10 illustrates a method embodiment for determining a position of the center of the cornea in the coordinate system with optical elements of the see-through, near-eye, mixed reality display. The one or more processors generate in step 622 a first plane including points including positions of a first illuminator for generating a first glint, a pupil center of the at least one image sensor, e.g. camera entrance pupil center, and the first glint. As in the embodiment of Figure 3A, the pupil center of the camera may be positioned in relation to the detection area 139 which acts as an image plane and which directs the light it receives to an image sensor in another location. In other examples, like in Figures 3B and 3C, the

52785-18(S)

43

detection area 139 may be the image sensor itself which is the image plane. This first plane will also include a position of the cornea center. Similarly, the one or more processors generate in step 624 a second plane including points including positions of a second illuminator for generating a second glint, the same pupil center of at least one sensor and the second glint. The two planes share the same camera pupil center as an origin and a distance vector to each illuminator is fixed with respect to the camera pupil center as the image sensor and illuminators are positioned on the near-eye display device at predetermined locations. These predetermined locations allow the various points in the planes to be related to each other in a third coordinate system including the two illuminators, the position of the camera pupil center, and the cornea center of curvature. The processor determines in step 626 the position of the cornea center of curvature based on the intersection of the first and second planes.

[00123] Figure 11 provides an illustrative example of the geometry of a gaze detection coordinate system 500 which may be used by the embodiment of Figure 10 to find the cornea center. In this embodiment, the at least one sensor is a camera modeled as a pin-hole camera. The geometry depicted is a slightly modified version of Figure 3 on page 89 of *Hennessey et al.* "A Single Camera Eye-Gaze Tracking System with Free Head Motion," ETRA 2006, San Diego, CA, ACM p. 88, pp. 87-94 (hereafter *Hennessey*). A list of variables is provided as follows:

[00124] $\hat{\mathbf{q}}_i$ is a position of an illuminator_{*i*}, the light of which produces glint $\hat{\mathbf{g}}_i$ (e.g. 174)

[00125] $\hat{\mathbf{g}}_i$ is the glint produced by illuminator_{*i*} (153) on a cornea surface,

[00126] $\hat{\mathbf{o}}$ is a camera pupil center of the pin-hole camera model,

[00127] $\hat{\mathbf{i}}_i$ is the image of glint $\hat{\mathbf{g}}_i$ on the image plane which is the detection area

52785-18(S)

44

139 of the camera sensor,

[00128] length_i is the scalar distance or length from point $\hat{o}$ to $\hat{q}_i$,

[00129] $\hat{\mathbf{i}}_i$ is the vector from the camera pupil center $\hat{o}$ to the image $\hat{\mathbf{i}}_i$ on the image sensor of the glint $\hat{\mathbf{g}}_i$,

[00130] $\hat{\mathbf{Q}}_i$ is the vector from the camera pupil center $\hat{o}$ to the position $\hat{q}_i$ of illuminator_{*i*},

[00131] the $\hat{\mathbf{X}}_i$ axis is defined along $\hat{\mathbf{Q}}_i$, in this example

[00132] and the $\hat{\mathbf{Z}}_i$ axis of the coordinate system is such so that $\hat{\mathbf{i}}_i$ which connects the image $\hat{\mathbf{i}}_i$ of the glint $\hat{\mathbf{g}}_i$ on image plane 139 (detection area) lies in a plane formed by the $\hat{\mathbf{X}}_i$ and $\hat{\mathbf{Z}}_i$ axes.

[00133] $\hat{\beta}$ is an angle formed in the $\hat{\mathbf{X}}_i\hat{\mathbf{Z}}_i$ plane between a line 502 representing the incident ray of light from the illuminator (153) position $\hat{q}_i$ to the glint $\hat{\mathbf{g}}_i$ (174) on a cornea surface.

[00134] $\hat{\alpha}$ is the angle formed in the $\hat{\mathbf{X}}_i\hat{\mathbf{Z}}_i$ plane between a line 504 representing the reflected ray from the glint $\hat{\mathbf{g}}_i$ to the camera pupil center of the camera, $\hat{o}$, which is also the origin of the coordinate system.

[00135] $\hat{\mathbf{c}}$ is the position of the cornea center which also lies in the $\hat{\mathbf{X}}_i\hat{\mathbf{Z}}_i$ plane.

As the cornea is modeled as a sphere, r is the radius of the corneal sphere, and each glint $\hat{\mathbf{g}}_i$ is a point on the first or external surface of the sphere, so each glint is separated from the cornea center by the radius r . In the above example, the glint $\hat{\mathbf{g}}_i$ is modeled as a point on the exterior surface or first surface of the cornea. In such a

52785-18(S)

45

model, the light of the illuminator is bouncing off the cornea in the same medium, air, of the same index of refraction as the reflected light of the glint directed back to the camera sensor.

As shown in Figure 11, a line or ray 506 normal to the glint $\hat{\mathbf{g}}_i$ on the surface of the cornea can be extended from the glint in the direction of the cornea and also extended to intersect with the $\hat{\mathbf{x}}_i$ axis of the $\hat{\mathbf{x}}_i\hat{\mathbf{z}}_i$ plane of the coordinate system. Also as shown in Figure 11, the incident ray 502 and the reflected ray 504 make a right triangle with the line length l_i between the position of the illuminator $\hat{\mathbf{q}}_i$ and the camera pupil center $\hat{\mathbf{o}}$. Thus angle A and angle D is each represented by

$$\frac{\pi - \tilde{\alpha}_i - \tilde{\beta}_i}{2} \quad \text{wherein} \quad \tilde{\alpha}_i = \cos^{-1} \left(\frac{-\tilde{l}_i \cdot \tilde{Q}_i}{\|\tilde{l}_i\| \cdot \|\tilde{Q}_i\|} \right) \quad \text{and} \quad \tilde{\beta}_i =$$

$$\tan^{-1} \left(\frac{\hat{\mathbf{g}}_{ix} \cdot \tan(\tilde{\alpha}_i)}{\tilde{l}_i - \hat{\mathbf{g}}_{ix}} \right).$$

[00136] According to *Hennessey*, the center of the cornea $\hat{\mathbf{c}}_i$ can be defined in the coordinate system 500 in terms of the unknown parameter $\hat{\mathbf{g}}_{ix}$ resulting in 3 equations for 4 unknowns $(\hat{\mathbf{c}}_{ix}, \hat{\mathbf{c}}_{iy}, \hat{\mathbf{c}}_{iz}, \hat{\mathbf{g}}_{ix})$ as follows:

52785-18(S)

46

$$[00137] \quad \begin{bmatrix} \hat{c}_{ix} \\ \hat{c}_{iy} \\ \hat{c}_{iz} \end{bmatrix} = \begin{bmatrix} \hat{g}_{ix} - r \cdot \sin\left(\frac{\hat{\alpha}_i - \hat{\beta}_i}{2}\right) \\ 0 \\ \hat{g}_{ix} \cdot \tan(\hat{\alpha}_i) + r \cdot \cos\left(\frac{\hat{\alpha}_i - \hat{\beta}_i}{2}\right) \end{bmatrix}$$

[00138] Another two-dimensional plane including the cornea center, $\hat{c}$, another glint $\hat{g}_i$, the camera pupil center $\hat{o}$ of the camera and a position $\hat{q}_i$ of another illuminator is also formed. The camera pupil center $\hat{o}$ of the camera and the cornea center are the same in each plane although the camera pupil center $\hat{o}$ position is known. This will result in 6 equations with 8 unknowns. In *Hennessey*, the gaze detection coordinate system is treated as an auxiliary coordinate system for which a rotation matrix $\hat{R}_i$ can transform points between the auxiliary coordinate systems for each plane and a single world coordinate system such as the third coordinate system which relates the position of the detection area 139 to the illuminators 153. A constraint exists in which the cornea center defined for each glint is the same in the world coordinate system, e.g. $\hat{c}_1 = \hat{c}_2$ and 3 equations result for the different axis components, e.g., $\hat{c}_{1x} = \hat{c}_{2x}$, $\hat{c}_{1y} = \hat{c}_{2y}$, and $\hat{c}_{1z} = \hat{c}_{2z}$, thus providing 9 equations with 8 unknowns. *Hennessey* (p. 90) states to solve numerically for $\hat{c}$ using a gradient descent algorithm. Thus, the position center 164 of the cornea 168 is defined with respect to the positions of the illuminators and the image plane or detection area 139.

[00139] Figure 12 illustrates a method embodiment for determining a pupil center from image data generated by a sensor. In step 642, the one or more processors identify a black pupil area in a number of image data samples of the respective eye and in step 644 averages the black pupil areas in the number of image data samples to adjust for headshake. An assumption may be made that a pupil is a circle and when viewed from an angle is an ellipse. One axis of the

52785-18(S)

47

ellipse, the major axis, remains constant as it represents the diameter of the pupil which does not change, provided the lighting does not change, as pupil size changes with lighting changes.

[00140] The pupil appears as a circle in an image format such as an image frame of a camera having its detection area centered on the optical axis of the display when the pupil is looking straight ahead through the display. As the pupil changes its gaze and moves from the center of the image frame, the pupil appears as an ellipse, as a circle viewed from an angle appears as an ellipse. The width of the minor axis of the ellipse changes with gaze changes. A narrow ellipse to the left of the center of the image frame indicates the user is looking to the far right. A wider ellipse a distance less to the right of the center of the image frame indicates the user is looking left but not far left.

[00141] The center of the pupil is the center of the ellipse. The ellipse is fitted from detected edge points in the image. Because such edge points are noisy and not all of them are on the ellipse, the ellipse fitting process is repeated many times over randomly selected subsets of all edge points. The subset that is most consistent with all the edge points is used to obtain the final ellipse. The processor in step 646 performs an ellipse fitting algorithm on the average black pupil area for determining an ellipse representing the pupil, and in step 648 determines the center of the pupil by determining the center of the ellipse representing the pupil.

[00142] With the center of rotation, the cornea center and the pupil center identified, one can extend a ray from the center of rotation through the cornea and pupil centers to obtain an optical axis for the eye. However, as noted previously, a gaze vector in a human is the visual axis or line of sight from the fovea through the pupil center. Photoreceptors in the fovea region of the human retina are more densely packed than in the rest of the retina. This area provides the highest visual

52785-18(S)

48

acuity or clearness of vision, and also provides stereoscopic vision of nearby objects. After determining the optical axis, a default offset angle may be applied so that the optical axis approximates the visual axis and is selected as the gaze vector.

[00143] Figure 13 illustrates a method embodiment for determining a gaze vector based on the determined centers for the pupil, the cornea and the rotation of the eyeball and which embodiment may be used to implement step 604. In step 652, the one or more processors model an optical axis 178 for the eye as a ray extending from the fixed center of rotation of the eyeball through the determined cornea and pupil centers and in step 654 applies a correction to the modeled optical axis for estimating a visual axis. In step 656, the one or more processors extend the estimated visual axis from the pupil through the display optical system of the see-through, near-eye display into the user field of view.

[00144] In one embodiment, with the fixed positioning of the illuminators as a basis, the effect of different areas of the eye on reflectivity and hence on the amount or intensity of light reflected is used as a basis for gaze detection. Intensity data from either IR or visible light sensors may be used to determine gaze, so the reflectivity data may be based on IR based reflectivity or visible light reflectivity. For illustration, the sclera is more reflective than other areas of the eye like the pupil and the iris. If a user looks to the user's far left, an illuminator 153 located on the frame 115 at the user's far right causes a glint reflection on the right sclera of the user's right eye. PSD 134r or as in Figure 3B, photodetector 152 on the inner right frame near bridge 104 receives more reflected light represented in a data reading while the light from reflection at the other photodetector 152 or position on the PSD when the illuminator 153 nearest the bridge is turned on receives a lower amount of reflected light in a range associated with the black pupil. The reflectivity of the iris may also be captured by camera 134 and stored for the user

52785-18(S)

49

by the processor 210, the processing unit 4 or a mobile device 5 embodying the processing unit 4.

[00145] The accuracy may not be as much as those based on images of the full eye, but may suffice for many applications. Additionally, such a gaze detection may be useful as an auxiliary or backup gaze detection technique. For example, during computationally intensive periods of generating complex virtual images, such a glint based technique relieves some processor overhead. Furthermore, such a glint-based technique can be executed many more times in a time period than an image based technique which processes more data or a computationally intensive but more accurate technique which may be run at a slower rate to recalibrate accuracy of gaze detection periodically. An example of a gaze detection technique which is both image based and more computationally intensive is one for determining a gaze vector with respect to inner parts of the eye based on glint data and pupil image data like the embodiments described in Figures 7 to 13 which may be run at a slower rate to recalibrate accuracy of gaze detection periodically. For example, an embodiment of the more computationally intensive technique based in part on image data may be run at ten (10) times a second while the glint based gaze detection technique may be run at a faster rate of one hundred (100) times per second or even five (500) hundred in some instances.

[00146] Figure 14 is a flowchart illustrating a method embodiment for determining gaze based on glint data. In step 673, data is captured representing each glint intensity value. Based on specular reflectivities of different eye parts, and positions of illuminators, an eyeball part is identified in step 674 based on the intensity value detected for each glint position in a geometrical relationship of the glints. In step 675, a gaze angle is estimated based on the eyeball part associated with each of the glint positions. As described in previous examples, an eyeball part

52785-18(S)

50

may be an iris, a pupil or a sclera of the eyeball. The positions of the illuminators form a geometry for the glints, e.g. a box, a circle, a rectangle, etc. which frame or surround the pupil, at least on two sides. A gaze vector is determined in step 676 based on the gaze angle, and a point of gaze in the 3D user field of view is determined in step 677 based on the intersection of the gaze vectors determined for both eyes,

[00147] As noted above, different methods with different accuracies may be employed at different periodic rates to trade accuracy for speed. A method embodiment based on glint intensity values such as that described in Figure 14 is an example of a technique with a low computational intensity which may be employed. In another example, training gaze data sets may be used for comparison with current pupil position data to determine a gaze vector.

[00148] Using training data sets for gaze determination relies on the assumption that the near-eye display device 2 with respect to the eye has not moved. If movement is detected, the training gaze data sets are to be recalibrated. A lighting change may also be a basis for recalibration.

[00149] A training gaze data set is acquired for each of a set of predetermined gaze directions. For example, training data sets may be obtained for different sections of the display optical system 14 through which the user's pupils gaze at a gaze or pupil angle. In one example, there are nine (9), one for each of the four (4) corners of the display optical system, a middle left side block or area, a middle right side block or area, a top middle block, a bottom middle block, and a center area. In the case of glints, a comparison of intensity values at the four glint positions for current data against training data sets may be used.

[00150] Figure 15a is a flowchart illustrating a method embodiment for

52785-18(S)

51

generating a set of training data sets for a comparison based determination of gaze. The method may be used to determine training sets for gaze angles based glint intensity value data representing pupil positions. The method embodiment is presented in an exemplar loop structure beginning at step 702 and ending at step 714. For each of a number of training gaze data sets, one or more processors of the control circuitry 136, the processing unit 4, the mobile device 5, a networked hub computing environment 12 alone or in combination, generate in step 704 a virtual image at a predetermined different position for the respective training data set for a predetermined time period in the user field of view. As previously discussed, the microdisplay 120 generates virtual images at different positions in the user field of view.

[00151] In step 706, data of each eye is captured during the predetermined time period based on glints. In step 708 from the captured data, the one or more processors determine data representing a pupil position, for example, a set of intensity values from a number of glints. In step 710, a gaze vector is determined based on the pupil position data and the predetermined different position of the virtual image in the user field of view. In the case of pupil and glint data being captured, a gaze vector may be determined based on the cornea center, pupil center and fixed center of eyeball rotation as discussed above with respect to the embodiments of Figures 7 to 13 and the position of the virtual image as a check. In the case of glint only data, the intensity values of the glints may be correlated with stored values reflecting different areas of reflection on the eye and is associated with a gaze vector extending to the virtual image position in the user field of view. The glint values may be checked against a set of values for the expected angle of the pupil viewing the virtual image at the predetermined position. In step 712, the one or more processors store the pupil position data and the gaze vector for the respective training gaze data set and proceeds in steps 714 and 702 to start

52785-18(S)

52

processing the next training gaze data set until the predetermined number of sets is reached.

[00152] Figure 15B is a flowchart illustrating a method embodiment for determining gaze based on the training data sets. In step 722, the at least one sensor captures data of each eye based on reflected light and the one or more processors determine from the captured data in step 724 data representing a current pupil position. In step 726, the one or more processors determine a gaze vector based on comparison of the current pupil position data with one or more training data sets and determines in step 728 a point of gaze based on the gaze vectors for the two eyes, e.g. where the two vectors intersect in a 3D user field of view.

[00153] Figure 15C is a flowchart of an interpolation method embodiment which may be used with the comparison step 726 of Figure 15B. For example, this embodiment may be used when comparing sensor data of the spatial relationship between the glints, for example, PSD data. In step 732, the one or more processors determine one or more distance vectors between the current pupil position data and the pupil position data of at least a subset of the training gaze data sets in accordance with a mapping criteria. On the detection area of a sensor, for example a camera sensor or discrete position sensitive detector, the mapping may be a distance in mm to pixel mapping. For an isotropic PSD, the mapping may be an area on the detector area to a distance in mm.

[00154] The box or other geometric shape of glints provides another example. A distance vector for each current glint from a training gaze data set of glint intensity values indicates a direction of intensity change as the glints are fixed barring movement of the coordinate system.

[00155] In step 734, the one or more processors select the training gaze data set

52785-18(S)

53

with the smallest sum for its one or more distance vectors and in step 736 interpolates a position change vector between the pupil position of the selected training gaze data set to the pupil position of the current data. In step 738, the one or more processors estimate as the current gaze vector the gaze vector of the selected training gaze data set moved by the position change vector

[00156] Particularly when using training data for comparison, movement of the gaze detection coordinate system is a cause for recalibrating the training data sets. One may periodically re-determine the positions of the cornea center and fixed center of rotation to determine whether there has been a change in the spatial relationship between them and the illuminators and at least one sensor.

[00157] Other tests for movement may be performed based on a facial feature with a fixed characteristic in image data. In one embodiment, an eye camera may capture about 5 to 10 mm of area around the visible eyeball portion of the cornea bulge, eye white, iris and pupil so as to capture part of an eyelid and eyelashes. A positionally fixed facial feature like a mole or freckle on skin such as an eyelid or on the bottom rim of the skin encasing the lower eyeball may also be present in the image data of the eye. In image samples, the position of the mole or freckle may be monitored for a change in position. If the facial feature has moved up, down, right or left, a vertical or horizontal shift can be detected. If the facial feature appears larger or smaller, a depth change in the spatial relationship between eye and display device 2 can be determined. There may be a criteria range in the change of position to trigger recalibration of the training gaze data sets due to things like camera resolution, etc.

[00158] In another example, although lighting is a factor which changes the size of the pupil and the ratio of pupil area to visible iris area within the circumference or perimeter of the iris, the size of the perimeter or circumference of the iris does

52785-18(S)

54

not change with gaze change or lighting change; hence, the perimeter or circumference is a fixed characteristic of the iris as a facial feature. Through ellipse fitting of the iris, the one or more processors can determine whether the iris has become larger or smaller in image data in accordance with criteria. If larger, the display device 2 with its illuminators 153 and at least one sensor 134 has moved closer in depth to the user's eye; if smaller, the display device 2 has moved farther away. A change in a fixed characteristic can trigger a recalibration of training data sets.

[00159] Figure 16 is a flowchart illustrating a method embodiment for checking calibration of a gaze determination system. The one or more processors of or in communication with the display device 2 in step 742 check whether a change has occurred to cause re-calibration of training data sets. One of the checks is determining in step 744 whether movement in accordance with a criteria has occurred. The check may be periodically determining a gaze vector in three dimensions as discussed per Figures 7 through 13 and noting the position of the fixed eyeball rotation has changed with respect to one or more gaze detection elements on the see-through, near-eye display device. The criteria may be a distance of movement in any of three dimensions. Based on the result of the determination in step 744 of has movement occurred indicating no movement, the one or more processors determine in step 746 whether a lighting change in accordance with a criteria has occurred. Responsive to a negative determination in step 746, other processing until next scheduled movement check is performed in step 747. If movement was indicated, the movement may have been detected in an image based technique based on a facial feature. Therefore, an optional step 748 may be performed of determining the boundaries for the gaze detection coordinate system as discussed for the embodiments of Figures 7 through 13. Responsive to the movement, a new set of training gaze data sets is generated in step 750.

52785-18(S)

55

Furthermore, if it was determined in step 746, that there was a lighting change which exceeds a threshold or other criteria, the new set of training gazed data sets may also be triggered in step 750.

[00160] Figure 17 is a block diagram of an exemplary mobile device which may operate in embodiments of the technology. Exemplary electronic circuitry of a typical mobile phone is depicted. The phone 900 includes one or more microprocessors 912, and memory 1010 (e.g., non-volatile memory such as ROM and volatile memory such as RAM) which stores processor-readable code which is executed by one or more processors of the control processor 912 to implement the functionality described herein.

[00161] Mobile device 900 may include, for example, processors 912, memory 1010 including applications and non-volatile storage. The processor 912 can implement communications, as well as any number of applications, including the interaction applications discussed herein. Memory 1010 can be any variety of memory storage media types, including non-volatile and volatile memory. A device operating system handles the different operations of the mobile device 900 and may contain user interfaces for operations, such as placing and receiving phone calls, text messaging, checking voicemail, and the like. The applications 1030 can be any assortment of programs, such as a camera application for photos and/or videos, an address book, a calendar application, a media player, an internet browser, games, other multimedia applications, an alarm application, other third party applications, the interaction application discussed herein, and the like. The non-volatile storage component 1040 in memory 1010 contains data such as web caches, music, photos, contact data, scheduling data, and other files.

[00162] The processor 912 also communicates with RF transmit/receive circuitry 906 which in turn is coupled to an antenna 902, with an infrared

52785-18(S)

56

transmitted/receiver 908, with any additional communication channels 1060 like Wi-Fi or Bluetooth, and with a movement/orientation sensor 914 such as an accelerometer. Accelerometers have been incorporated into mobile devices to enable such applications as intelligent user interfaces that let users input commands through gestures, indoor GPS functionality which calculates the movement and direction of the device after contact is broken with a GPS satellite, and to detect the orientation of the device and automatically change the display from portrait to landscape when the phone is rotated. An accelerometer can be provided, e.g., by a micro-electromechanical system (MEMS) which is a tiny mechanical device (of micrometer dimensions) built onto a semiconductor chip. Acceleration direction, as well as orientation, vibration and shock can be sensed. The processor 912 further communicates with a ringer/vibrator 916, a user interface keypad/screen, biometric sensor system 918, a speaker 1020, a microphone 922, a camera 924, a light sensor 926 and a temperature sensor 928.

[00163] The processor 912 controls transmission and reception of wireless signals. During a transmission mode, the processor 912 provides a voice signal from microphone 922, or other data signal, to the RF transmit/receive circuitry 906. The transmit/receive circuitry 906 transmits the signal to a remote station (e.g., a fixed station, operator, other cellular phones, etc.) for communication through the antenna 902. The ringer/vibrator 916 is used to signal an incoming call, text message, calendar reminder, alarm clock reminder, or other notification to the user. During a receiving mode, the transmit/receive circuitry 906 receives a voice or other data signal from a remote station through the antenna 902. A received voice signal is provided to the speaker 1020 while other received data signals are also processed appropriately.

[00164] Additionally, a physical connector 988 can be used to connect the

52785-18(S)

57

mobile device 900 to an external power source, such as an AC adapter or powered docking station. The physical connector 988 can also be used as a data connection to a computing device. The data connection allows for operations such as synchronizing mobile device data with the computing data on another device.

[00165] A GPS transceiver 965 utilizing satellite-based radio navigation to relay the position of the user applications is enabled for such service.

[00166] The example computer systems illustrated in the figures include examples of computer readable storage media. Computer readable storage media are also processor readable storage media. Such media may include volatile and nonvolatile, removable and non-removable media implemented in any method or technology for storage of information such as computer readable instructions, data structures, program modules or other data. Computer storage media includes, but is not limited to, RAM, ROM, EEPROM, cache, flash memory or other memory technology, CD-ROM, digital versatile disks (DVD) or other optical disk storage, memory sticks or cards, magnetic cassettes, magnetic tape, a media drive, a hard disk, magnetic disk storage or other magnetic storage devices, or any other medium which can be used to store the desired information and which can be accessed by a computer.

[00167] Figure 18 is a block diagram of one embodiment of a computing system that can be used to implement the hub computing system of Figures 1A and 1B. In this embodiment, the computing system is a multimedia console 800, such as a gaming console. As shown in Figure 18, the multimedia console 800 has a central processing unit (CPU) 801, and a memory controller 802 that facilitates processor access to various types of memory, including a flash Read Only Memory (ROM) 803, a Random Access Memory (RAM) 806, a hard disk drive 808, and portable media drive 806. In one implementation, CPU 801 includes a level 1 cache 810

52785-18(S)

58

and a level 2 cache 812, to temporarily store data and hence reduce the number of memory access cycles made to the hard drive 808, thereby improving processing speed and throughput.

[00168] CPU 801, memory controller 802, and various memory devices are interconnected via one or more buses (not shown). The details of the bus that is used in this implementation are not particularly relevant to understanding the subject matter of interest being discussed herein. However, it will be understood that such a bus might include one or more of serial and parallel buses, a memory bus, a peripheral bus, and a processor or local bus, using any of a variety of bus architectures. By way of example, such architectures can include an Industry Standard Architecture (ISA) bus, a Micro Channel Architecture (MCA) bus, an Enhanced ISA (EISA) bus, a Video Electronics Standards Association (VESA) local bus, and a Peripheral Component Interconnects (PCI) bus also known as a Mezzanine bus.

[00169] In one implementation, CPU 801, memory controller 802, ROM 803, and RAM 806 are integrated onto a common module 814. In this implementation, ROM 803 is configured as a flash ROM that is connected to memory controller 802 via a PCI bus and a ROM bus (neither of which are shown). RAM 806 is configured as multiple Double Data Rate Synchronous Dynamic RAM (DDR SDRAM) modules that are independently controlled by memory controller 802 via separate buses (not shown). Hard disk drive 808 and portable media drive 805 are shown connected to the memory controller 802 via the PCI bus and an AT Attachment (ATA) bus 816. However, in other implementations, dedicated data bus structures of different types can also be applied in the alternative.

[00170] A graphics processing unit 820 and a video encoder 822 form a video processing pipeline for high speed and high resolution (e.g., High Definition)

52785-18(S)

59

graphics processing. Data are carried from graphics processing unit (GPU) 820 to video encoder 822 via a digital video bus (not shown). Lightweight messages generated by the system applications (e.g., pop ups) are displayed by using a GPU 820 interrupt to schedule code to render popup into an overlay. The amount of memory used for an overlay depends on the overlay area size and the overlay preferably scales with screen resolution. Where a full user interface is used by the concurrent system application, it is preferable to use a resolution independent of application resolution. A scaler may be used to set this resolution such that the need to change frequency and cause a TV resync is eliminated.

[00171] An audio processing unit 824 and an audio codec (coder/decoder) 826 form a corresponding audio processing pipeline for multi-channel audio processing of various digital audio formats. Audio data are carried between audio processing unit 824 and audio codec 826 via a communication link (not shown). The video and audio processing pipelines output data to an A/V (audio/video) port 828 for transmission to a television or other display. In the illustrated implementation, video and audio processing components 820-828 are mounted on module 214.

[00172] Figure 18 shows module 814 including a USB host controller 830 and a network interface 832. USB host controller 830 is shown in communication with CPU 801 and memory controller 802 via a bus (e.g., PCI bus) and serves as host for peripheral controllers 804(1)-804(4). Network interface 832 provides access to a network (e.g., Internet, home network, etc.) and may be any of a wide variety of various wire or wireless interface components including an Ethernet card, a modem, a wireless access card, a Bluetooth module, a cable modem, and the like.

[00173] In the implementation depicted in Figure 18 console 800 includes a controller support subassembly 840 for supporting four controllers 804(1)-804(4). The controller support subassembly 840 includes any hardware and software

52785-18(S)

60

components needed to support wired and wireless operation with an external control device, such as for example, a media and game controller. A front panel I/O subassembly 842 supports the multiple functionalities of power button 812, the eject button 813, as well as any LEDs (light emitting diodes) or other indicators exposed on the outer surface of console 802. Subassemblies 840 and 842 are in communication with module 814 via one or more cable assemblies 844. In other implementations, console 800 can include additional controller subassemblies. The illustrated implementation also shows an optical I/O interface 835 that is configured to send and receive signals that can be communicated to module 814.

[00174] MUs 840(1) and 840(2) are illustrated as being connectable to MU ports "A" 830(1) and "B" 830(2) respectively. Additional MUs (e.g., MUs 840(3)-840(6)) are illustrated as being connectable to controllers 804(1) and 804(3), i.e., two MUs for each controller. Controllers 804(2) and 804(4) can also be configured to receive MUs (not shown). Each MU 840 offers additional storage on which games, game parameters, and other data may be stored. In some implementations, the other data can include any of a digital game component, an executable gaming application, an instruction set for expanding a gaming application, and a media file. When inserted into console 800 or a controller, MU 840 can be accessed by memory controller 802. A system power supply module 850 provides power to the components of gaming system 800. A fan 852 cools the circuitry within console 800. A microcontroller unit 854 is also provided.

[00175] An application 860 comprising machine instructions is stored on hard disk drive 808. When console 800 is powered on, various portions of application 860 are loaded into RAM 806, and/or caches 810 and 812, for execution on CPU 801, wherein application 860 is one such example. Various applications can be stored on hard disk drive 808 for execution on CPU 801.

52785-18(S)

61

[00176] Gaming and media system 800 may be operated as a standalone system by simply connecting the system to monitor 16 (Figure 1A), a television, a video projector, or other display device. In this standalone mode, gaming and media system 800 enables one or more players to play games, or enjoy digital media, e.g., by watching movies, or listening to music. However, with the integration of broadband connectivity made available through network interface 832, gaming and media system 800 may further be operated as a participant in a larger network gaming community.

[00177] Although the subject matter has been described in language specific to structural features and/or methodological acts, it is to be understood that the subject matter defined in the appended claims is not necessarily limited to the specific features or acts described above. Rather, the specific features and acts described above are disclosed as example forms of implementing the claims.

52785-18(S)

62

CLAIMS:

1. A mixed reality display system with gaze determination comprising:
 - a see-through, near-eye, mixed reality display device including a respective display optical system for each eye positioned to be seen through by the respective eye and at least one image generation unit for generating at least one virtual image for display in the display optical systems;
 - a respective arrangement of gaze detection elements positioned on the display device for forming a spatial relationship between the gaze detection elements and each eye, the gaze detection elements including
 - a set of illuminators for generating glints on the respective eye, each illuminator positioned on the respective see-through display device at a respective predetermined position and generating illumination about a predetermined wavelength, and
 - at least one respective sensor for capturing reflections from the respective eye and generating data representing the captured reflections, the at least one respective sensor positioned at a predetermined position in relation to predetermined positions of the set of illuminators on the respective see-through display device;
 - a memory for storing software and the data; and
 - a processor communicatively coupled to the at least one respective sensor to receive the data representing the captured reflections and having access to the memory for storing the data, the processor determining a gaze vector for each respective eye based on the data representing the captured reflections and a point of gaze based on the gaze vectors in a three-dimensional (3D) user field of view;

52785-18(S)

63

wherein data representing the captured reflections includes glint intensity data and wherein for each eye the processor determining the gaze vector comprises:

at least some of the time using a first technique that comprises determining the center of the cornea and the center of the pupil based on image data of the respective eye; and

at least some of the time using a second different, less computationally intensive technique based on the glint intensity data independent of the image data of the respective eye.

2. The system of claim 1 wherein

the second technique determines the gaze vector based on determining a position for each glint with respect to at least one eyeball part based on the glint intensity data and the predetermined positions of the set of illuminators and the at least one respective sensor.

3. The system of claim 2 wherein the values of the glint intensity data are correlated with stored values for specular reflectivities of different eye parts.

4. The system of claim 3 wherein the different eye parts comprise an iris or a pupil or a sclera or a combination of two or more of an iris, pupil and sclera.

5. The system of claim 1, wherein:

the 3D user field of view includes both real and virtual objects.

6. The system of claim 1 further comprising:

wherein the illuminators are infra-red (IR) illuminators which transmit about a predetermined wavelength;

at least one light sensor is an IR camera for generating image data of the glints and a pupil; and

52785-18(S)

64

wherein the first technique involves determining the gaze vector including a center of a cornea and a center of a pupil based on the image data of glints and the pupil and the spatial relationship.

7. The system of claim 6 further comprising:

wherein the processor determining the gaze vector using the first technique is performed at a first rate; and

the processor determining the gaze vector by the second technique is performed at a second rate faster than the first rate.

8. The system of claim 6 further comprising:

the processor checks whether movement of the display device has occurred with respect to the eye based on a change in the spatial relationship between the gaze detection elements and each eye.

9. The system of claim 6 wherein the IR camera is aligned with an optical axis of the respective display optical system.

10. A method for determining gaze in a see-through, near-eye mixed reality display system comprising:

determining boundaries of a three dimensional gaze detection coordinate system based on positions of glints detected on a user eye, positions on the near-eye display system of illuminators for generating the glints, and a position of at least one sensor for detecting the glints;

determining a gaze vector for each eye based on reflected eye data including the glints;

determining a point of gaze based on the gaze vectors for the two eyes in a three-dimensional (3D) user field of view including real and virtual objects; and

identifying any object at the point of gaze in the 3D user field of view.

11. The method of claim 10, wherein:

52785-18(S)

65

determining boundaries of the gaze detection coordinate system based on positions of the glints detected on the user eye, positions on the near-eye display system of illuminators for generating the glints, and the position of at least one sensor for detecting the glints further comprises:

determining a position of a center of a cornea of each respective eye with respect to the positions of the illuminators and the at least one sensor based on positions of glints detected on the cornea surface of the respective eye;

determining a pupil center from image data generated from the at least one sensor; and

determining a position of a fixed center of eyeball rotation of each eye based on the position of the cornea center and the pupil center.

12. The method of claim 10, further comprising:

determining a position from the fixed center of eyeball rotation of each eye with respect to the respective illuminators and the at least one sensor for determining movement of the display device with respect to the respective eye.

13. The method of claim 11, wherein the at least one sensor determining a pupil center from image data generated from the at least one sensor further comprises:

identifying a black pixel area as a black pupil area in a number of image samples;

average the black pupil areas to adjust for headshake;

performing an ellipse fitting algorithm on the average black pupil area for determining an ellipse representing the pupil; and

determining the center of the pupil by determining the center of the ellipse representing the pupil.

52785-18(S)

66

14. The method of claim 11, wherein determining a position of a center of a cornea of each respective eye with respect to the illuminators and at least one sensor based on positions of glints detected on the cornea surface of the respective eye further comprises:

generating a first plane including points including positions of a first illuminator which generates a first glint, a camera pupil center of a detection area of the at least one sensor, and the first glint on a surface of the cornea;

generating a second plane including points including a position of a second illuminator which generates a second glint, and the camera pupil center of the detection area of the at least one sensor, and the second glint; and

determining the position of the cornea center based on the intersection of the first and the second planes.

15. The method of claim 14, wherein

determining the position of the fixed center of eyeball rotation of each eye based on the position of the respective cornea center and the respective pupil center further comprises estimating the position of the fixed center of eyeball rotation of each eye based on a vector from the determined pupil center through the cornea center having a length based on length approximations between the respective pupil center and the respective cornea center and between the respective cornea center and the respective center of eyeball rotation.

16. The method of claim 15, wherein

determining the gaze vector for each eye based on reflected eye data including the glints further comprises

modeling an optical axis for each eye as a ray extending from the fixed center of rotation of the eyeball through the respective cornea and pupil centers;

52785-18(S)

67

applying a correction to the modelled optical axis for estimating a visual axis; and

extending the estimated visual axis from the pupil through the see-through, near-eye display into the user field of view.

17. The method of claim 16 wherein determining the point of gaze based on the gaze vectors for the two eyes further comprises identifying the point of gaze as the point where the two gaze vectors intersect.

18. The method of claim 10 wherein the processor determines a change in a depth distance of the respective display optical system based on a change of size of a facial feature in the image captured by a camera.

19. The method of claim 18 wherein the facial feature is an iris of the respective eye.

20. The method of claims 10 wherein:

the coordinate system includes a depth axis;

changes along the depth axis trigger redetermination of the boundaries or re-calibration of training data.

21. A mixed reality display system with gaze determination based on glints comprising:

a see-through, near-eye display device including a respective display optical system for each eye positioned to be seen through by the respective eye and at least one image generation unit for generating at least one virtual image for display;

a set of infra-red illuminators for producing glints for each eye, each illuminator positioned on the respective see-through display device at a respective predetermined position and generating infra-red radiation about a predetermined wavelength;

at least one respective sensor for detecting the glints reflected from each eye and generating glint data including intensity data for the glints, the

52785-18(S)

68

at least one respective sensor positioned in relation to one or more respective illuminator positions on the respective see-through display device at at least one predetermined position to detect infra-red reflected radiation about the predetermined wavelength;

a memory for storing software and data including glint data; and

a processor communicatively coupled to the at least one respective sensor to receive the glint data and having access to the memory for storing the glint data, the processor determining one or more glint positions relative to an eye part based on the glint intensity data generated by the at least one respective sensor, the respective predetermined position of each illuminator in the set of infra-red illuminators, and the at least one predetermined position of the at least one respective sensor for detecting the glints reflected from each eye;

the processor determining a gaze vector for each respective eye based on the one or more glint positions relative to an eye part; and

the processor determining a point of gaze based on the gaze vectors in a three dimensional user field of view.

22. The system of claim 21 further comprising:

the processor is communicatively coupled to the at least one image generation unit for controlling the at least one image generation unit for displaying the at least one virtual image at different locations in the user field of view viewable through the respective display optical system;

the at least one respective sensor capturing a calibration glint data set of intensity values for each glint during display of the at least one virtual image at each different location;

the processor storing the calibration glint data sets of intensity values as indicators of calibration gaze vectors; and

52785-18(S)

69

the processor determines the gaze vector of a user based on comparisons of the captured glint data with the calibration glint data sets.

23. The system of claim 21 further comprising wherein the at least one respective sensor is a position sensitive detector and the processor turns each illuminator on and off in a predetermined sequence for capturing one or more data sets of glint intensity values.

24. The system of claim 21 wherein the glint intensity data comprises one or more data sets of glint intensity values and the number of glints include at least two glints, and the processor the processor determining a gaze vector for each respective eye the processor determining one or more glint positions relative to an eye part based on the glint intensity data generated by the at least one respective sensor, the respective predetermined position of each illuminator in the set of infra-red illuminators, and the at least one predetermined position of the at least one respective sensor for detecting the glints reflected from each eye further comprises determining a reflectivity data value for represented by each glint intensity value, and identifying a glint corresponding to the glint intensity value as being located on a sclera, an iris portion or a pupil portion based on the reflectivity data value represented by each glint intensity value.

Title: GAZE DETECTION IN A SEE-THROUGH, NEAR-EYE, MIXED REALITY DISPLAY
Applicant: John R. Lewis
Appl. No.: Unknown
Filing Date:
Docket No.: 01466US0
Atty: Eileen Lehmann
Phone: (415) 369-9660

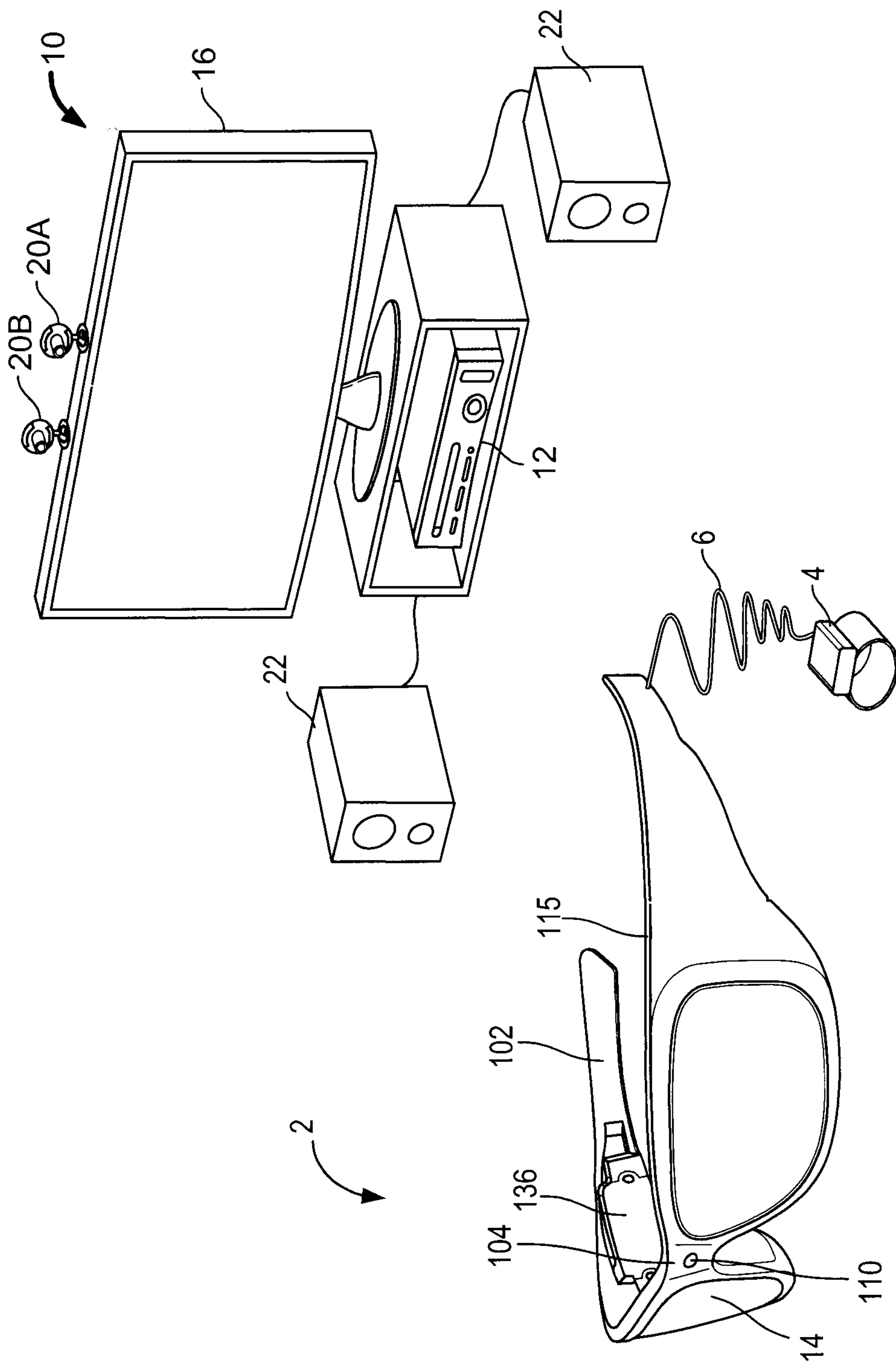


FIG. 1A

Title:	GAZE DETECTION IN A SEE-THROUGH, NEAR-EYE, MIXED REALITY DISPLAY		
Applicant:	John R. Lewis	Docket No.:	01466US0
Appl. No.:	Unknown	Atty:	Eileen Lehmann
Filing Date:		Phone:	(415) 369-9660

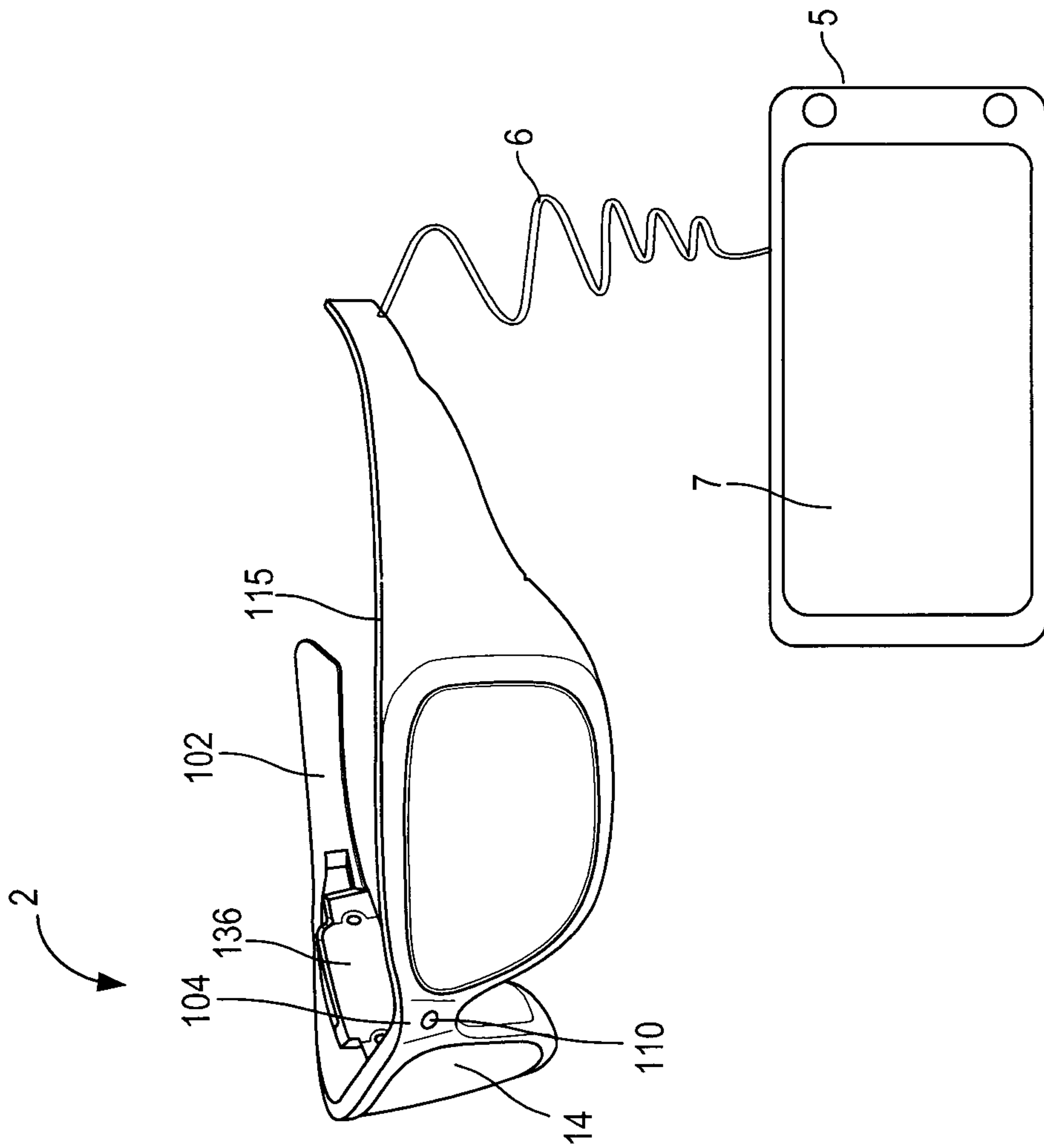


FIG. 1B

Title: GAZE DETECTION IN A SEE-THROUGH, NEAR-EYE, MIXED REALITY DISPLAY
Applicant: John R. Lewis Docket No.: 01466US0
Appl. No.: Unknown Atty: Eileen Lehmann
Filing Date: Phone: (415) 369-9660

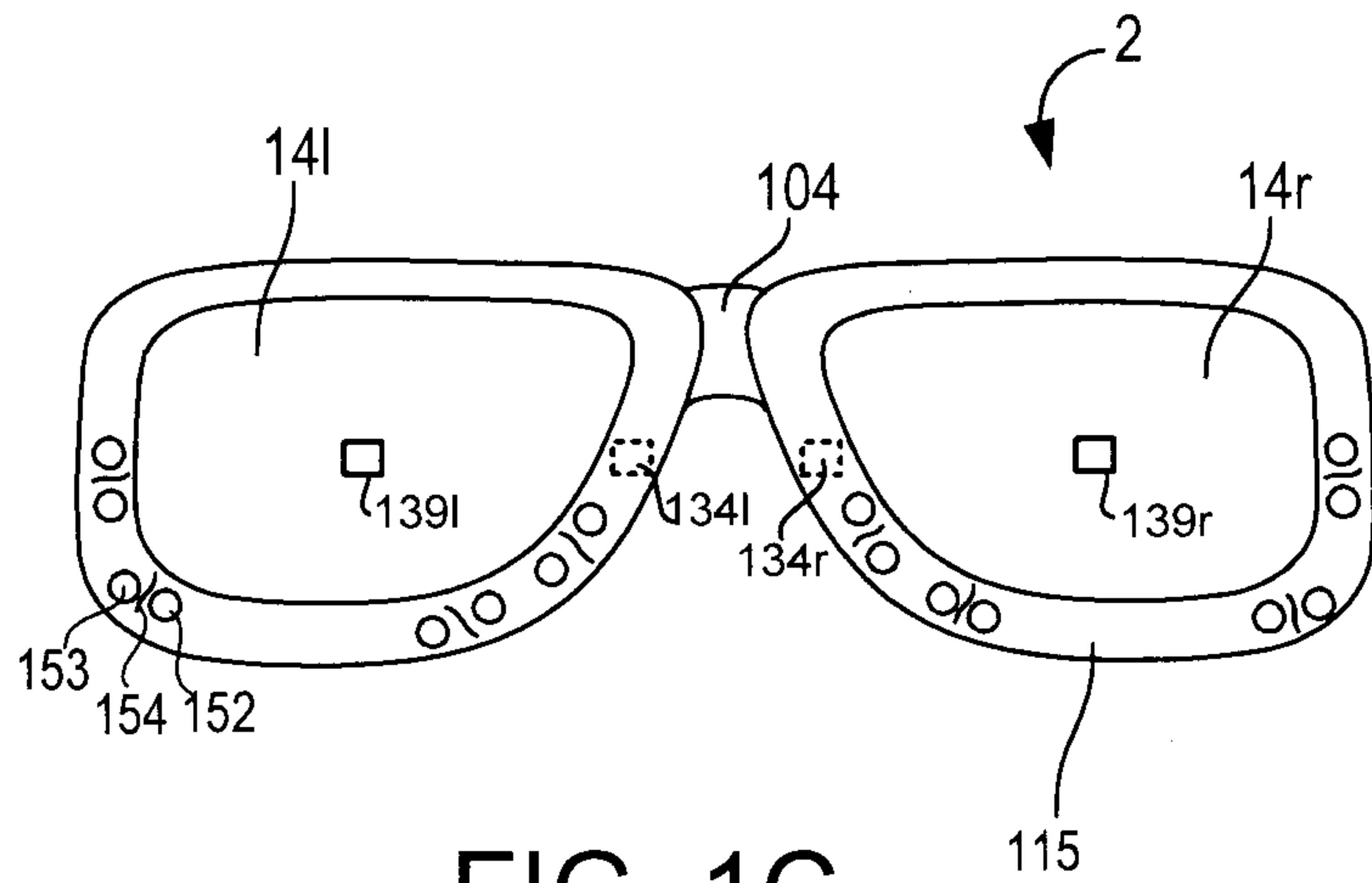


FIG. 1C

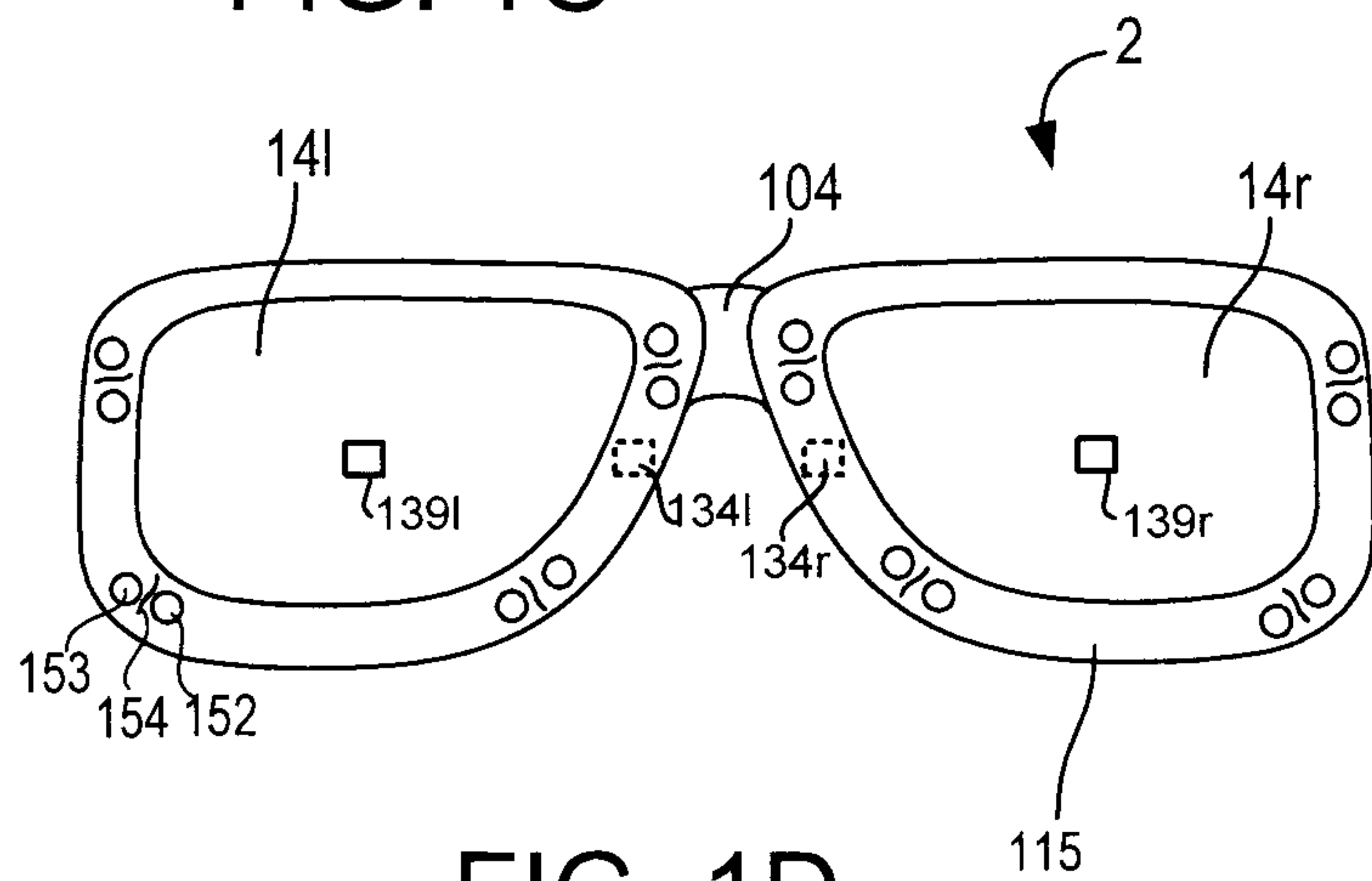


FIG. 1D

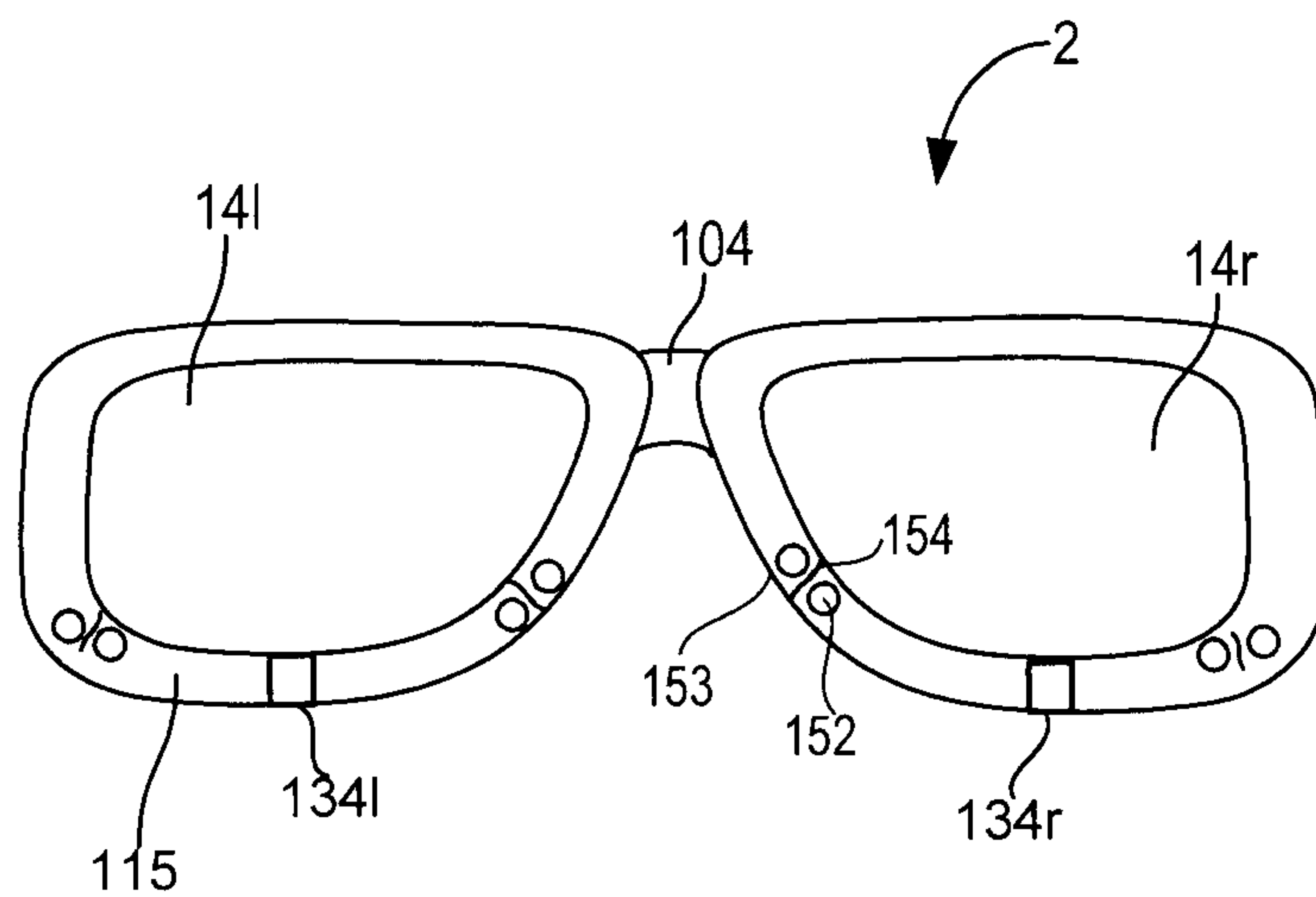


FIG. 1E

Title: GAZE DETECTION IN A SEE-THROUGH, NEAR-EYE, MIXED REALITY DISPLAY
Applicant: John R. Lewis
Docket No.: 01466US0
App. No.: Unknown
Atty: Eileen Lehmann
Filing Date: (415) 369-9660
Phone:

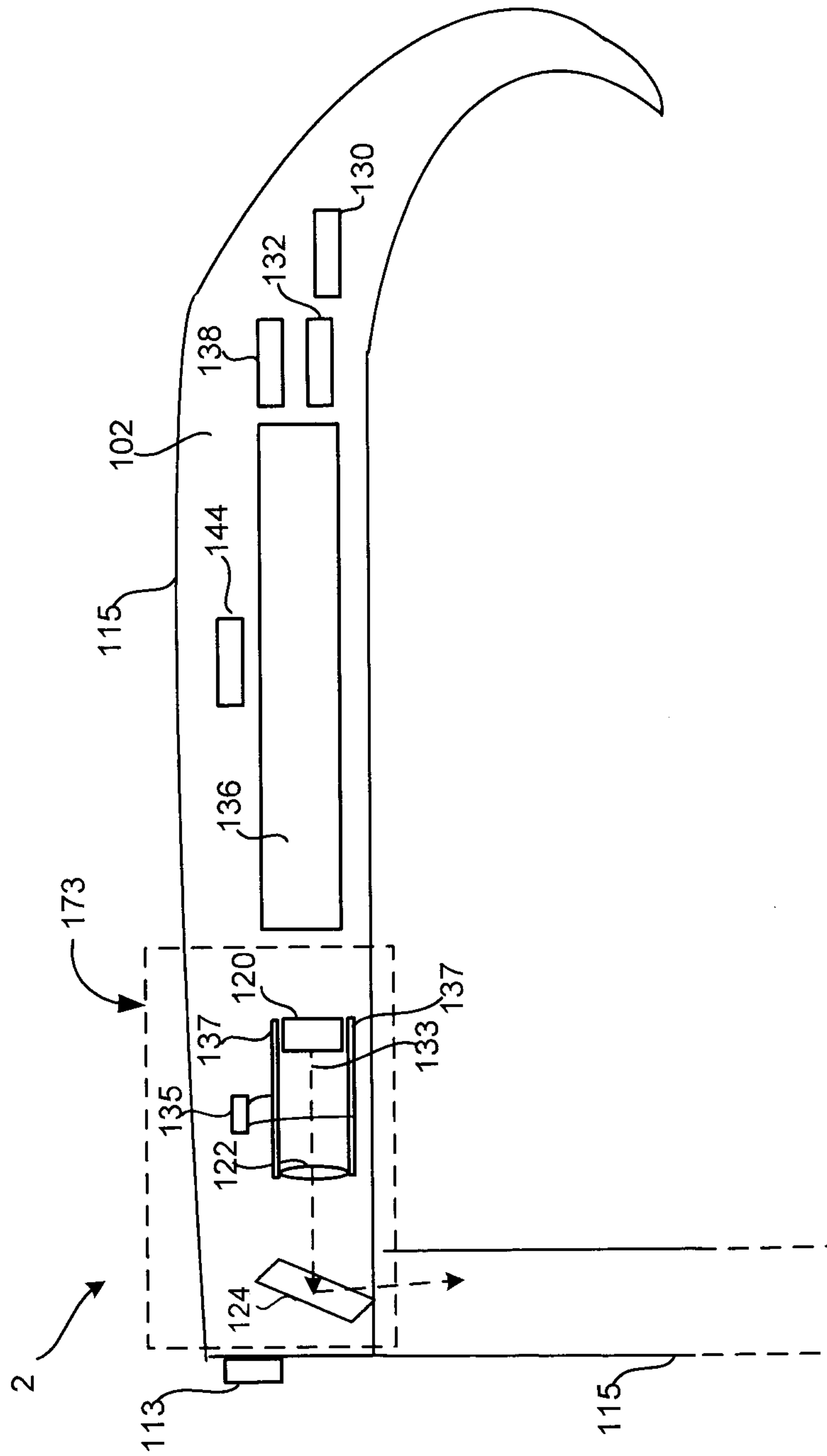


FIG. 2

Title: GAZE DETECTION IN A SEE-THROUGH, NEAR-EYE, MIXED REALITY DISPLAY
Applicant: John R. Lewis Docket No.: 01466US0
Appl. No.: Unknown Atty: Eileen Lehmann
Filing Date: Phone: (415) 369-9660

FIG. 3B

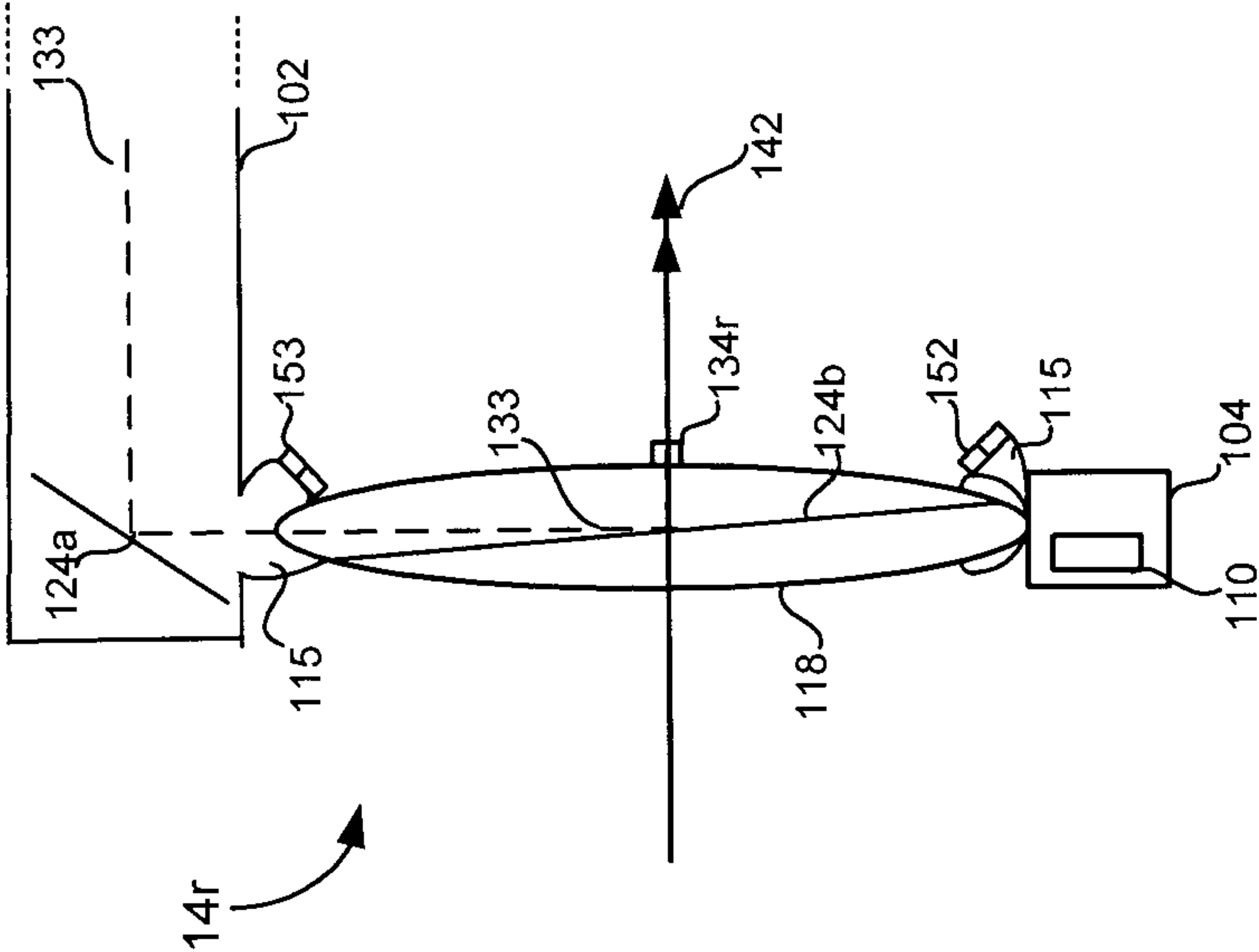
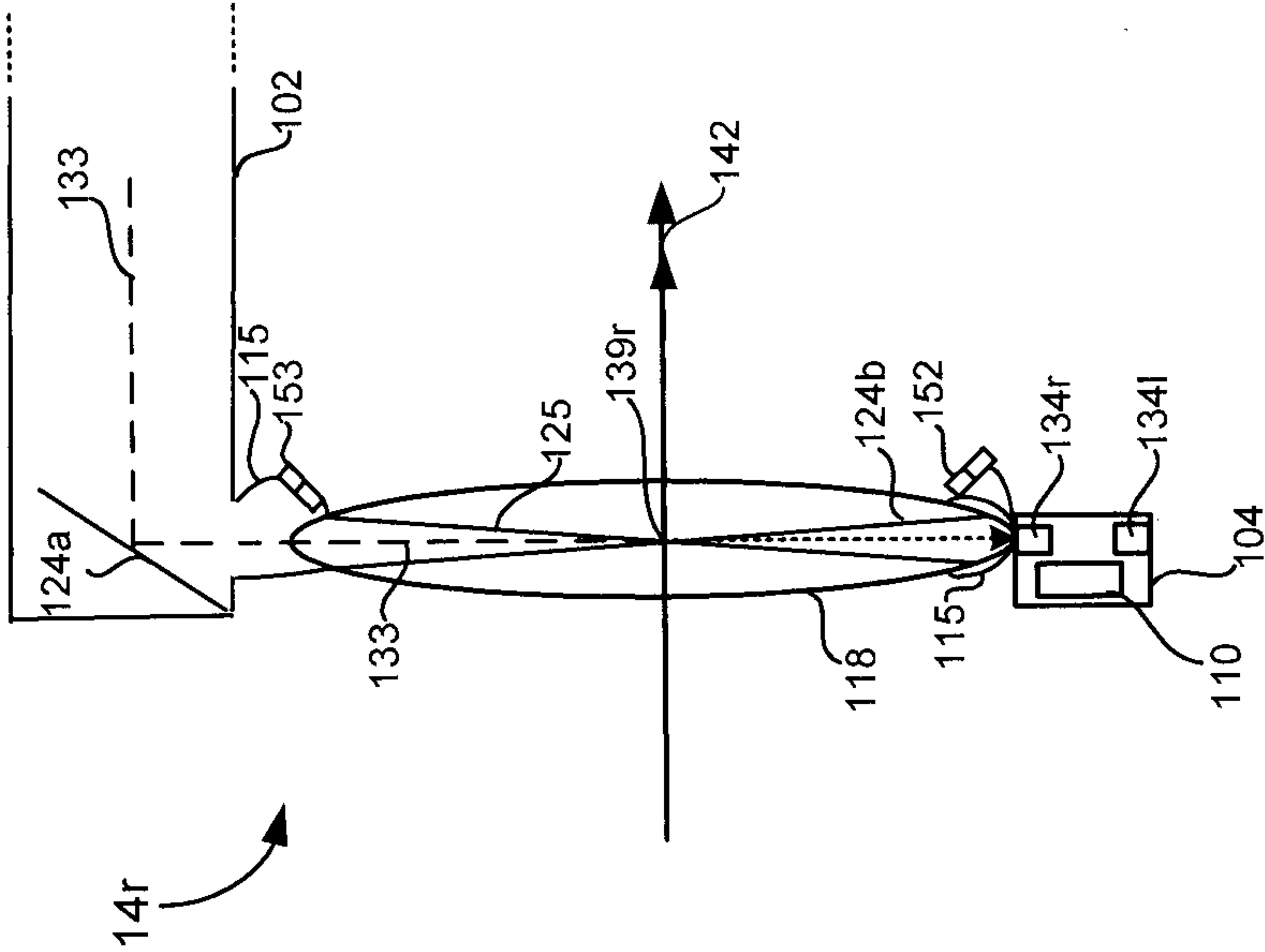


FIG. 3A



Title: GAZE DETECTION IN A SEE-THROUGH, NEAR-EYE, MIXED REALITY DISPLAY
Applicant: John R. Lewis Docket No.: 01466US0
Appl. No.: Unknown Atty: Eileen Lehmann
Filing Date: Phone: (415) 369-9660

FIG. 3D

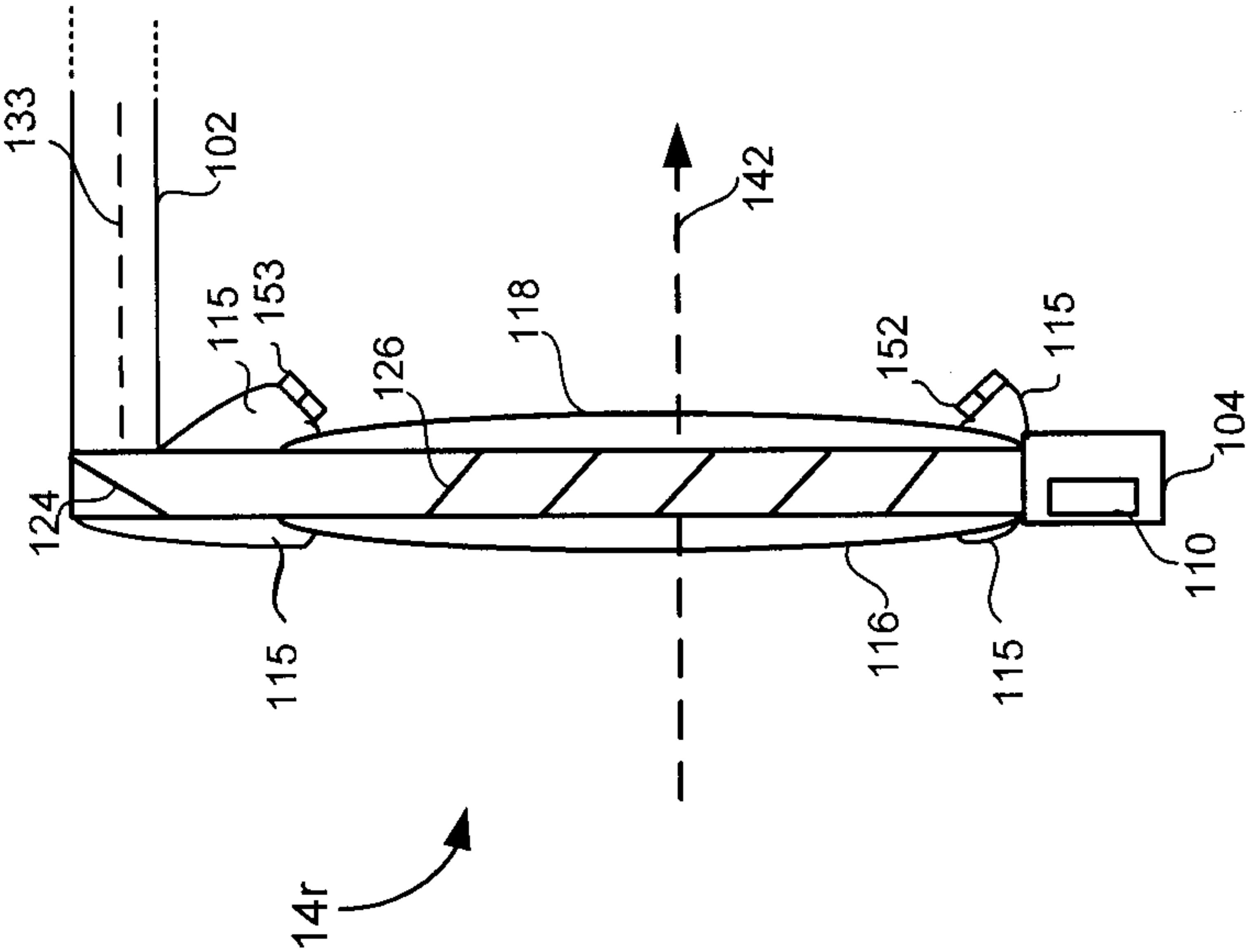
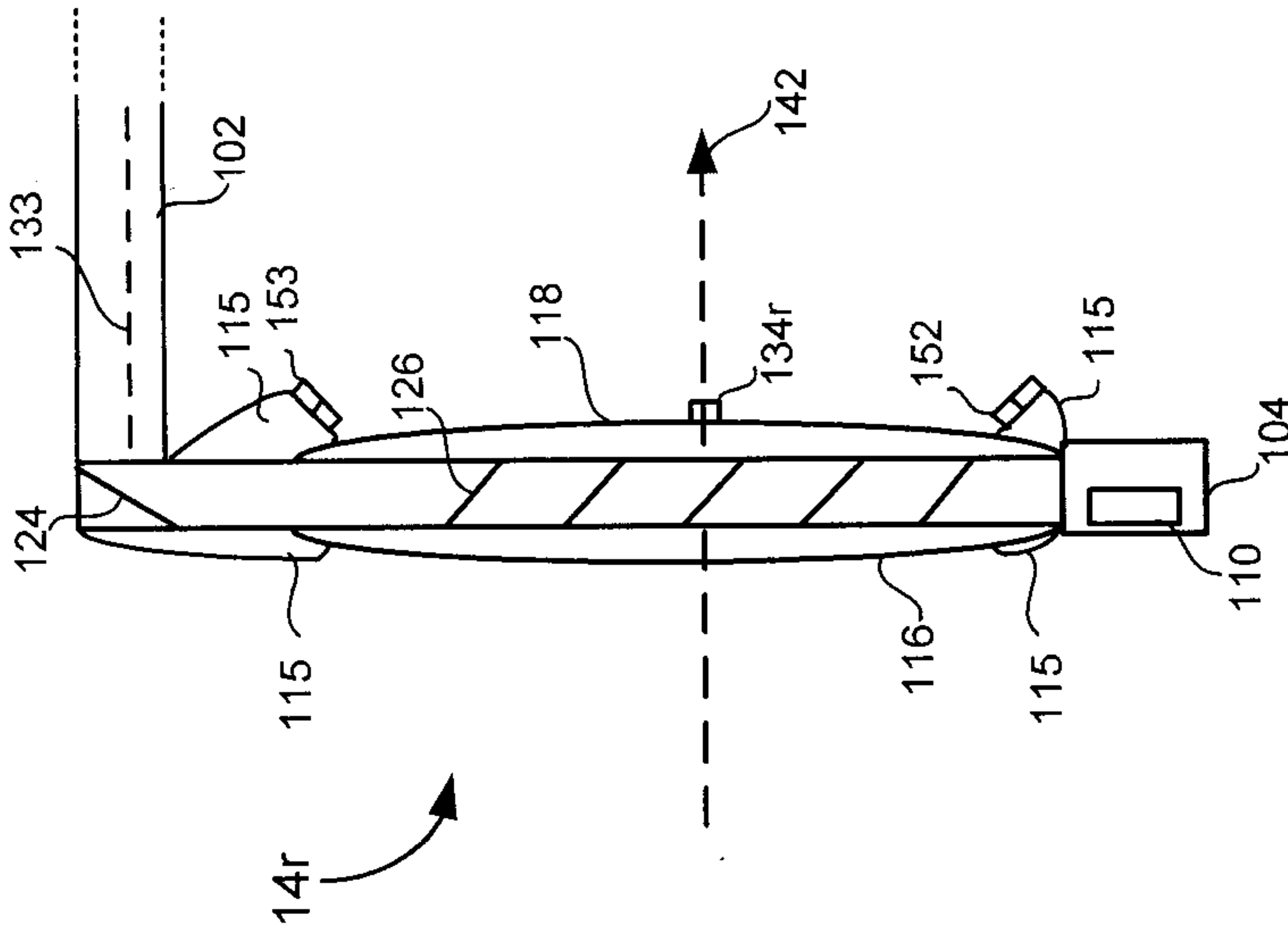
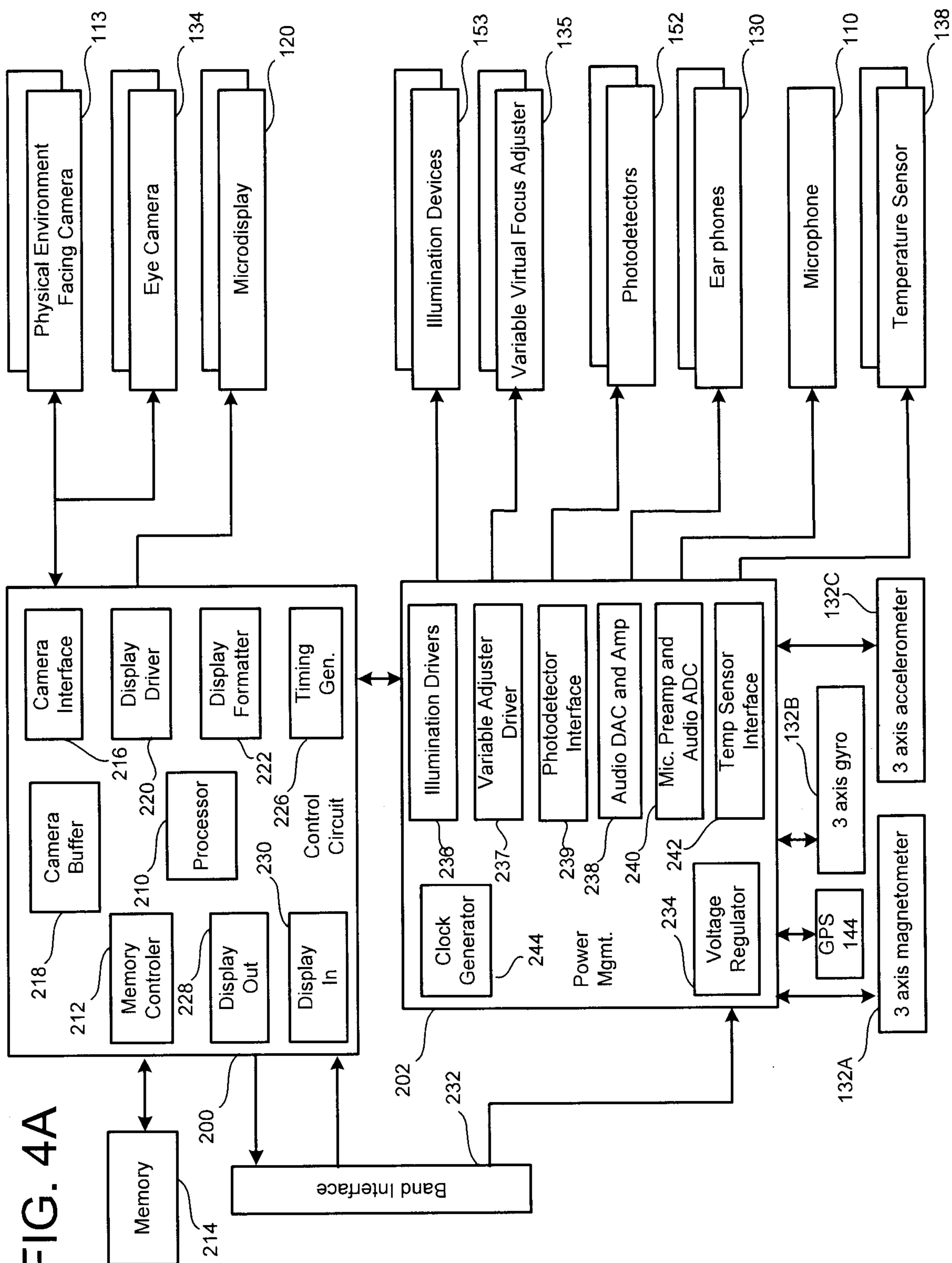


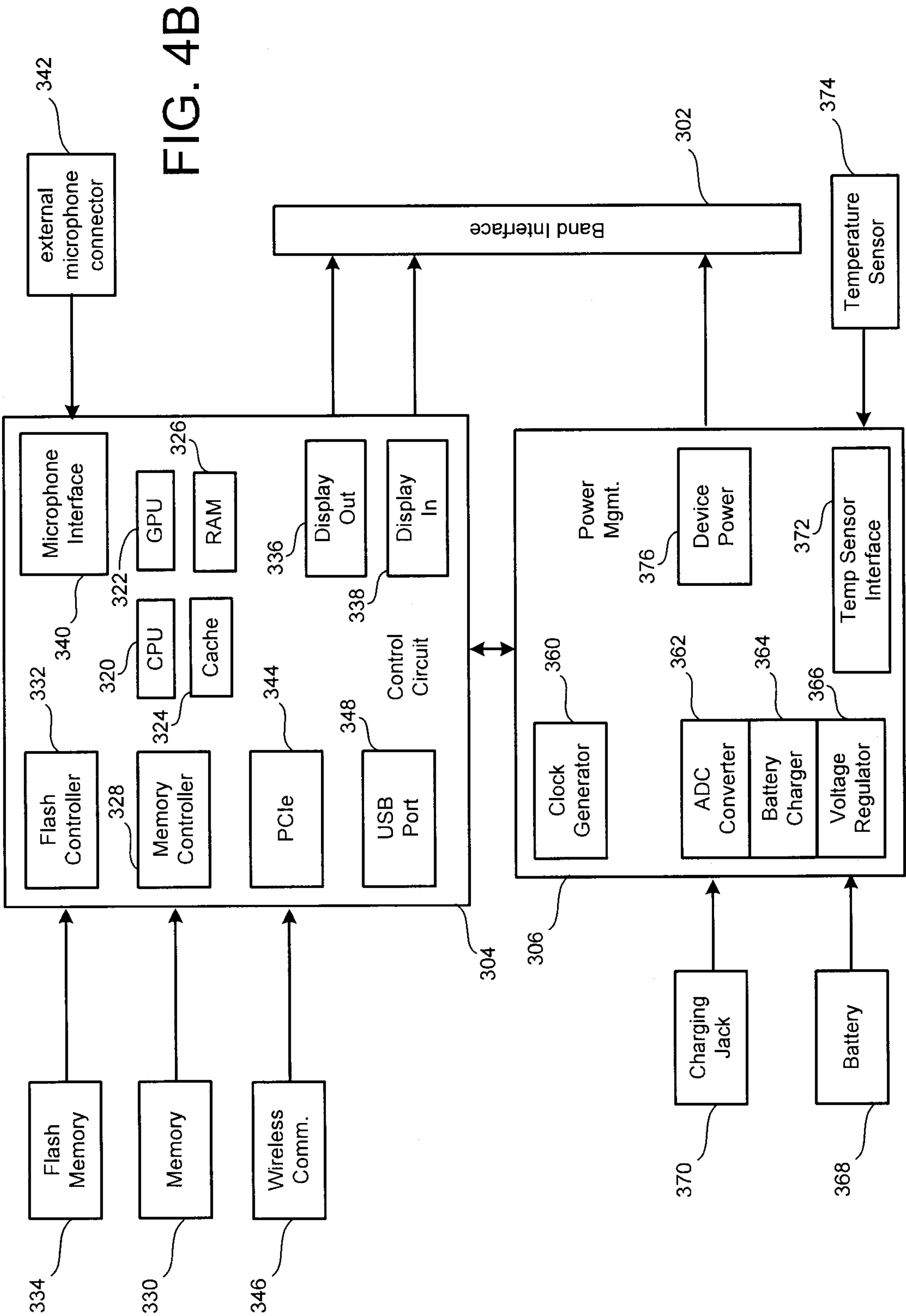
FIG. 3C



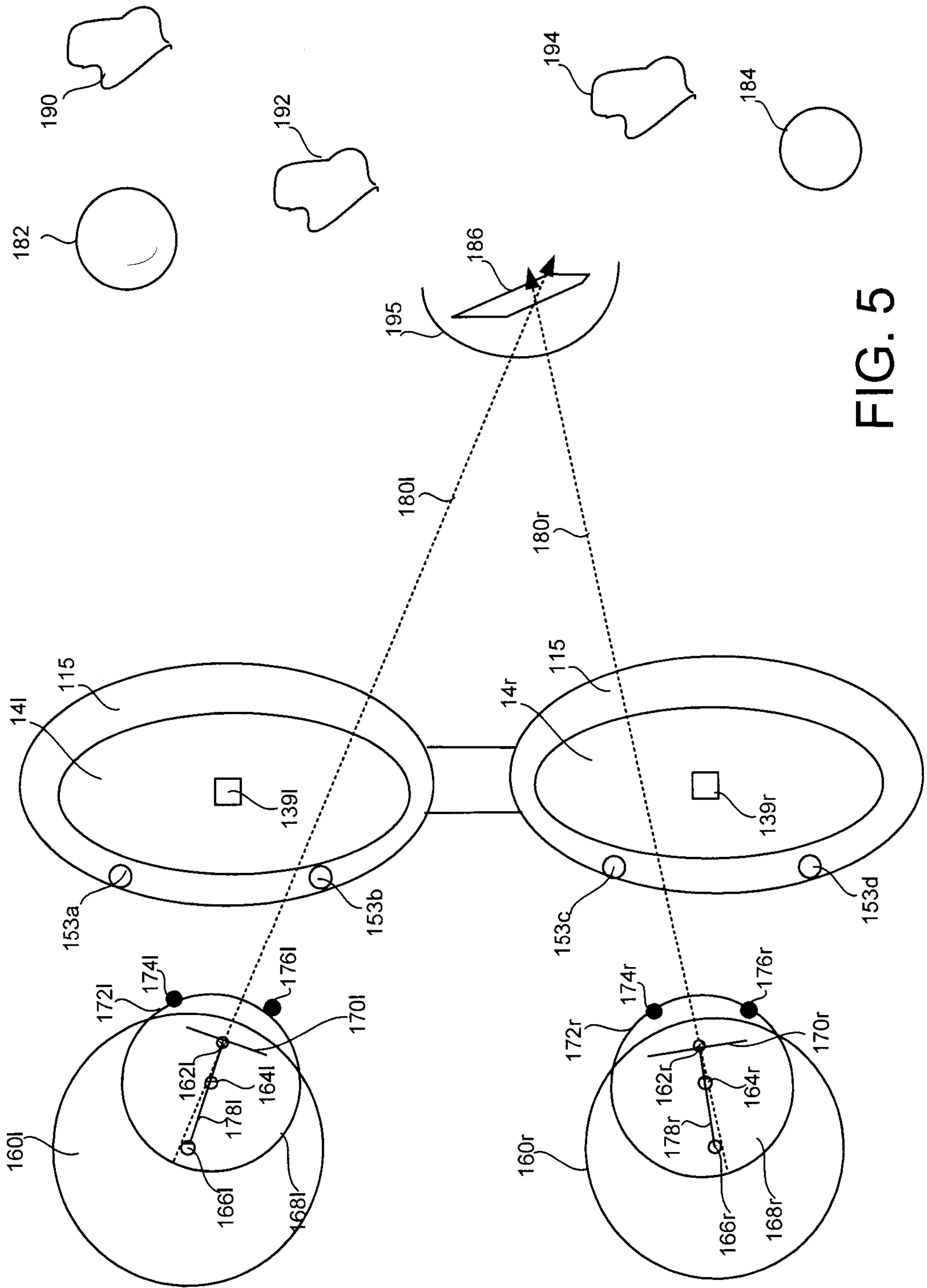
Title: GAZE DETECTION IN A SEE-THROUGH, NEAR-EYE, MIXED REALITY DISPLAY
Applicant: John R. Lewis Docket No.: 01466USO
Appl. No.: Unknown Atty: Eileen Lehmann
Filing Date: Phone: (415) 369-9660

FIG. 4A





Title: GAZE DETECTION IN A SEE-THROUGH, NEAR-EYE, MIXED REALITY DISPLAY
 Applicant: John R. Lewis Docket No.: 01466US0
 Appl. No.: Unknown Atty: Eileen Lehmann
 Filing Date: Phone: (415) 369-9660



Title: GAZE DETECTION IN A SEE-THROUGH, NEAR-EYE, MIXED REALITY DISPLAY
Applicant: John R. Lewis Docket No.: 01466US0
Appl. No.: Unknown Atty: Eileen Lehmann
Filing Date: Phone: (415) 369-9660

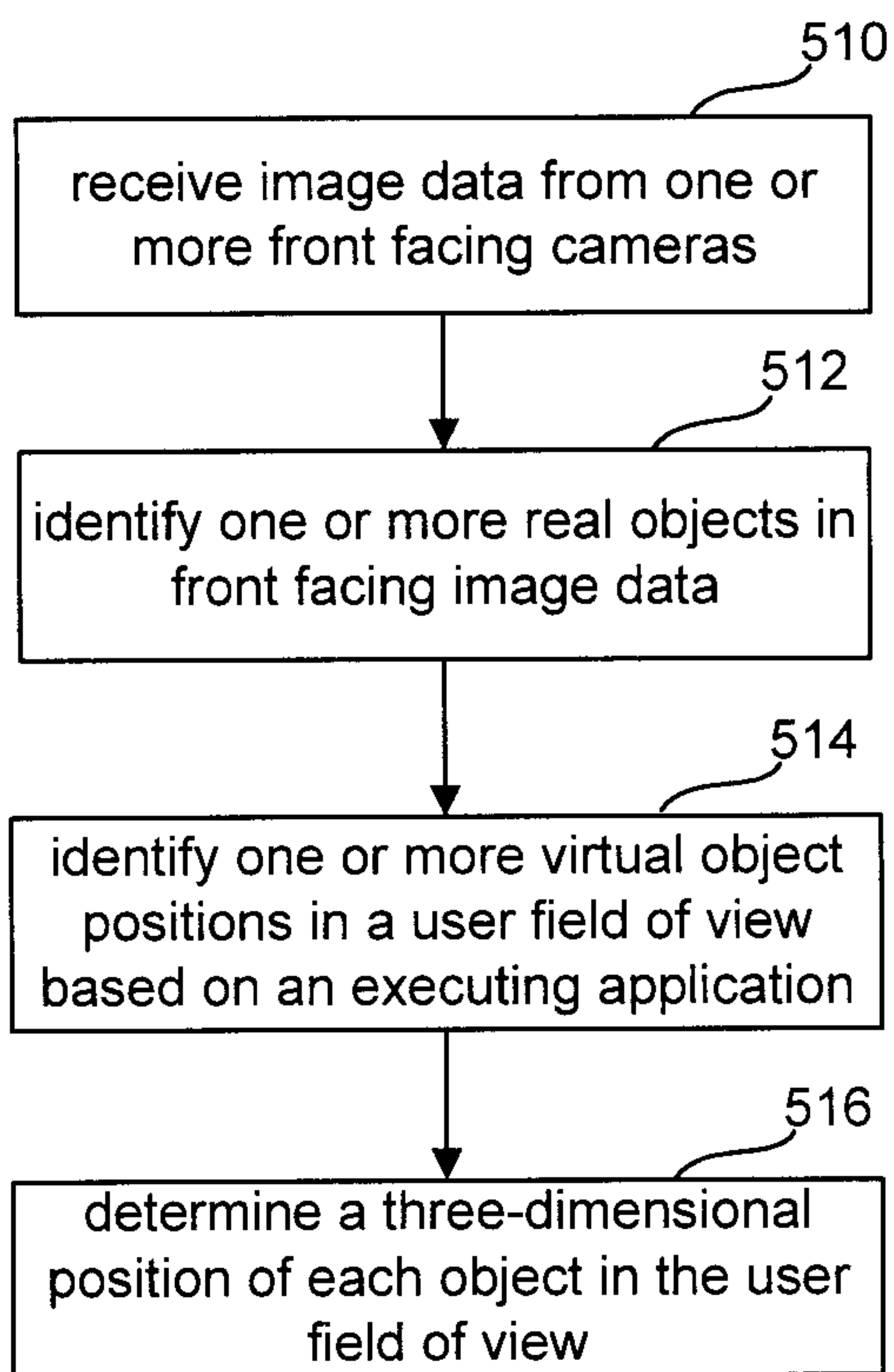


FIG. 6A

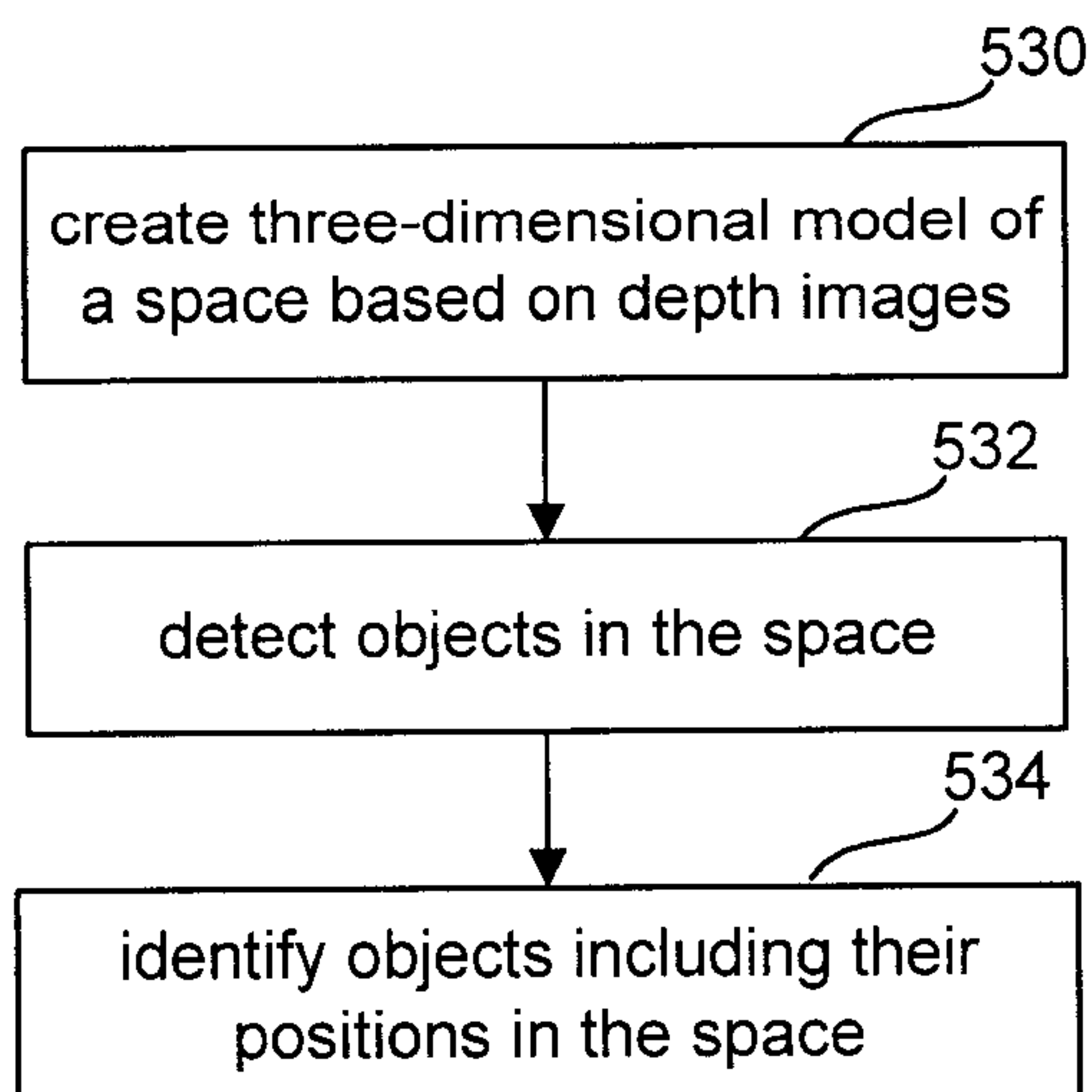


FIG. 6C

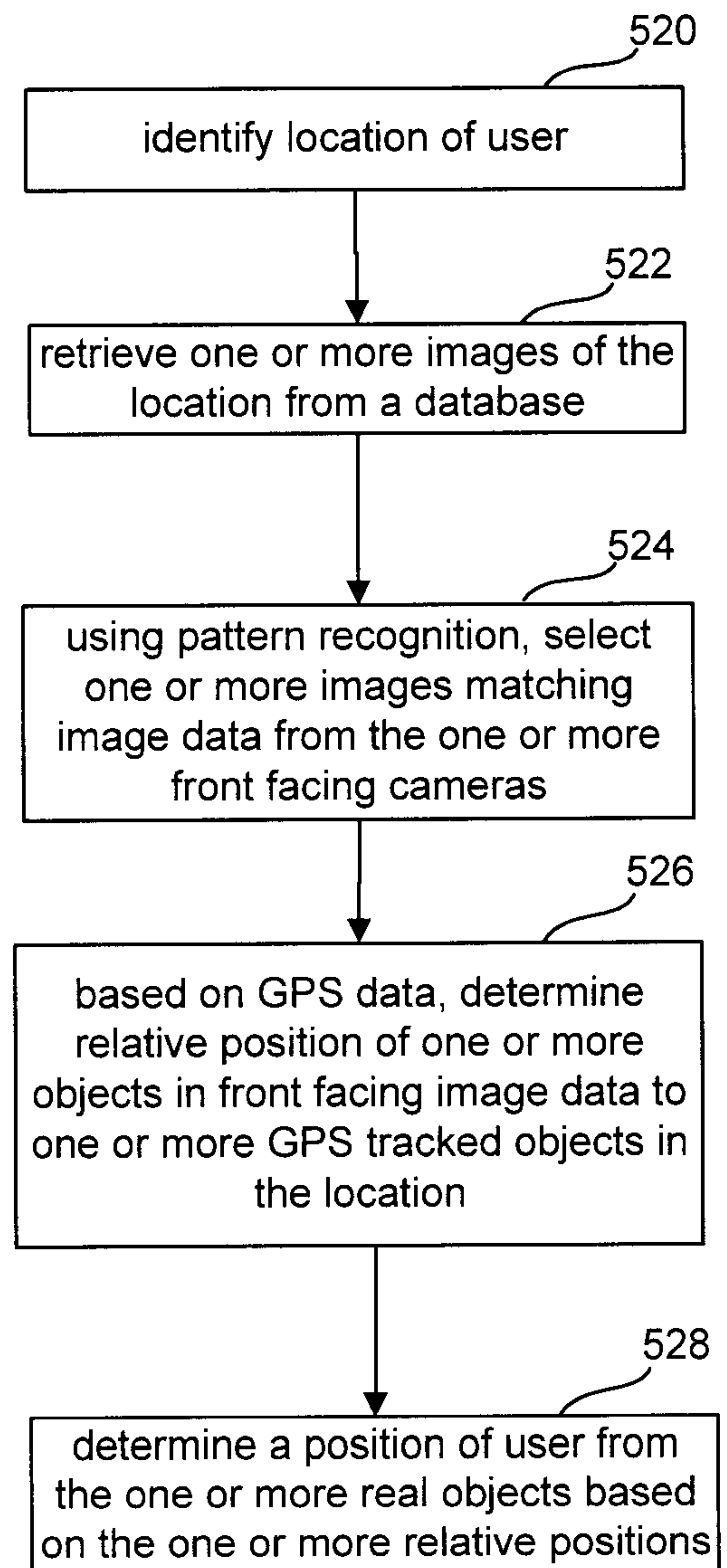


FIG. 6B

Title:	GAZE DETECTION IN A SEE-THROUGH, NEAR-EYE, MIXED REALITY DISPLAY		
Applicant:	John R. Lewis	Docket No.:	01466US0
Appl. No.:	Unknown	Atty:	Eileen Lehmann
Filing Date:		Phone:	(415) 369-9660

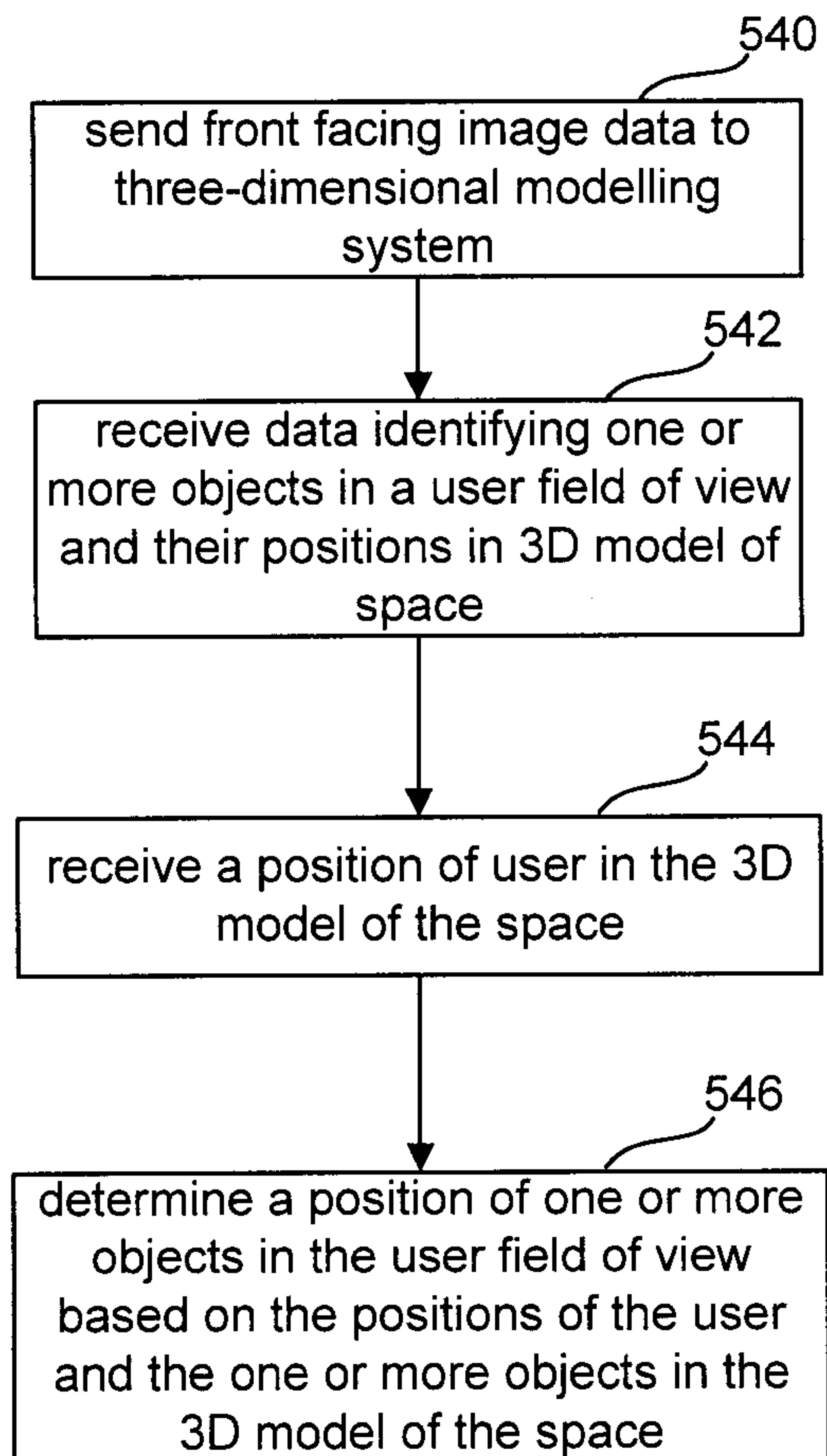


FIG. 6D

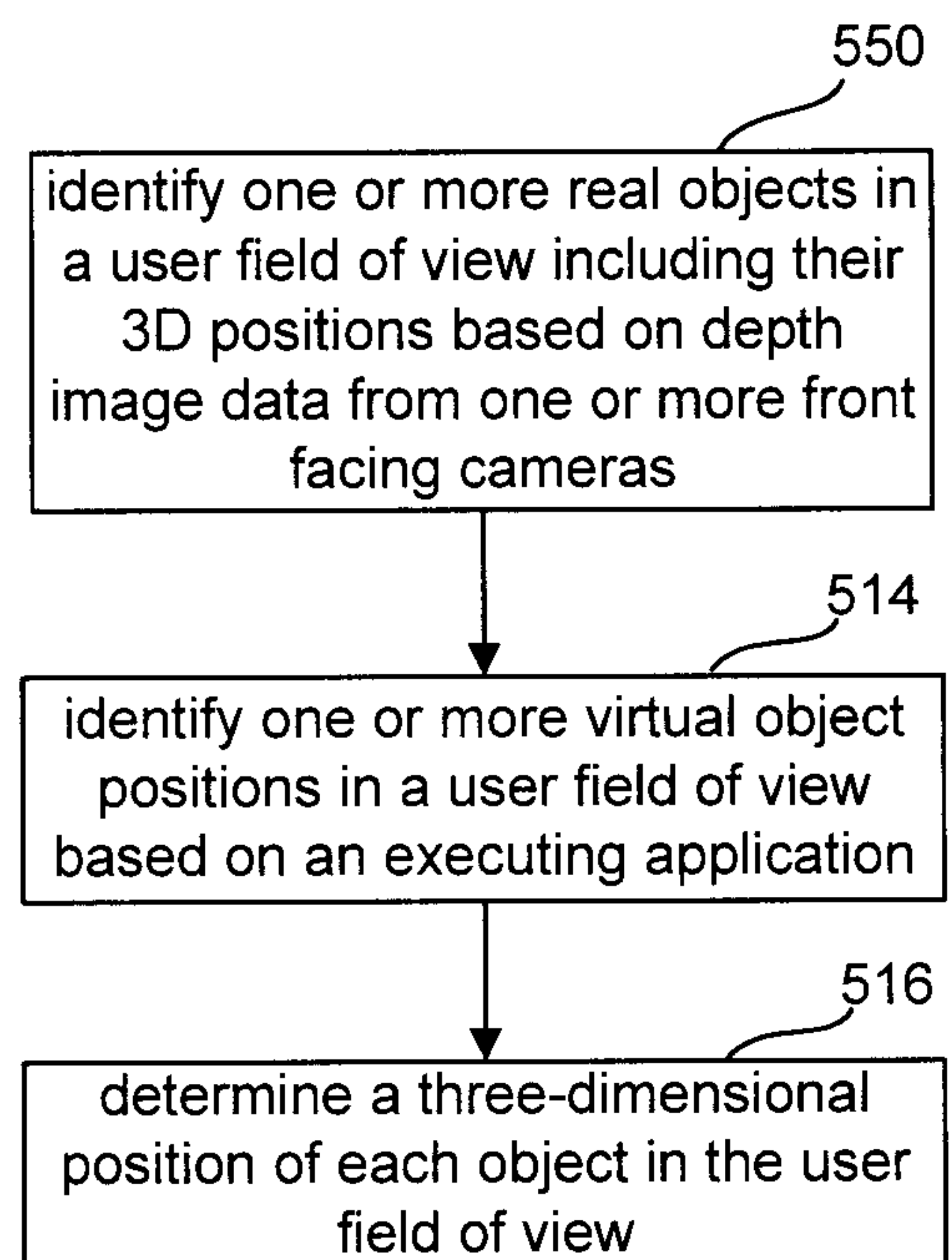


FIG. 6E

Title:GAZE DETECTION IN A SEE-THROUGH, NEAR-EYE, MIXED REALITY DISPLAY

Applicant: John R. LewisDocket No.: 01466US0

Appl. No.: UnknownAtty: Eileen Lehmann

Filing Date:Phone: (415) 369-9660

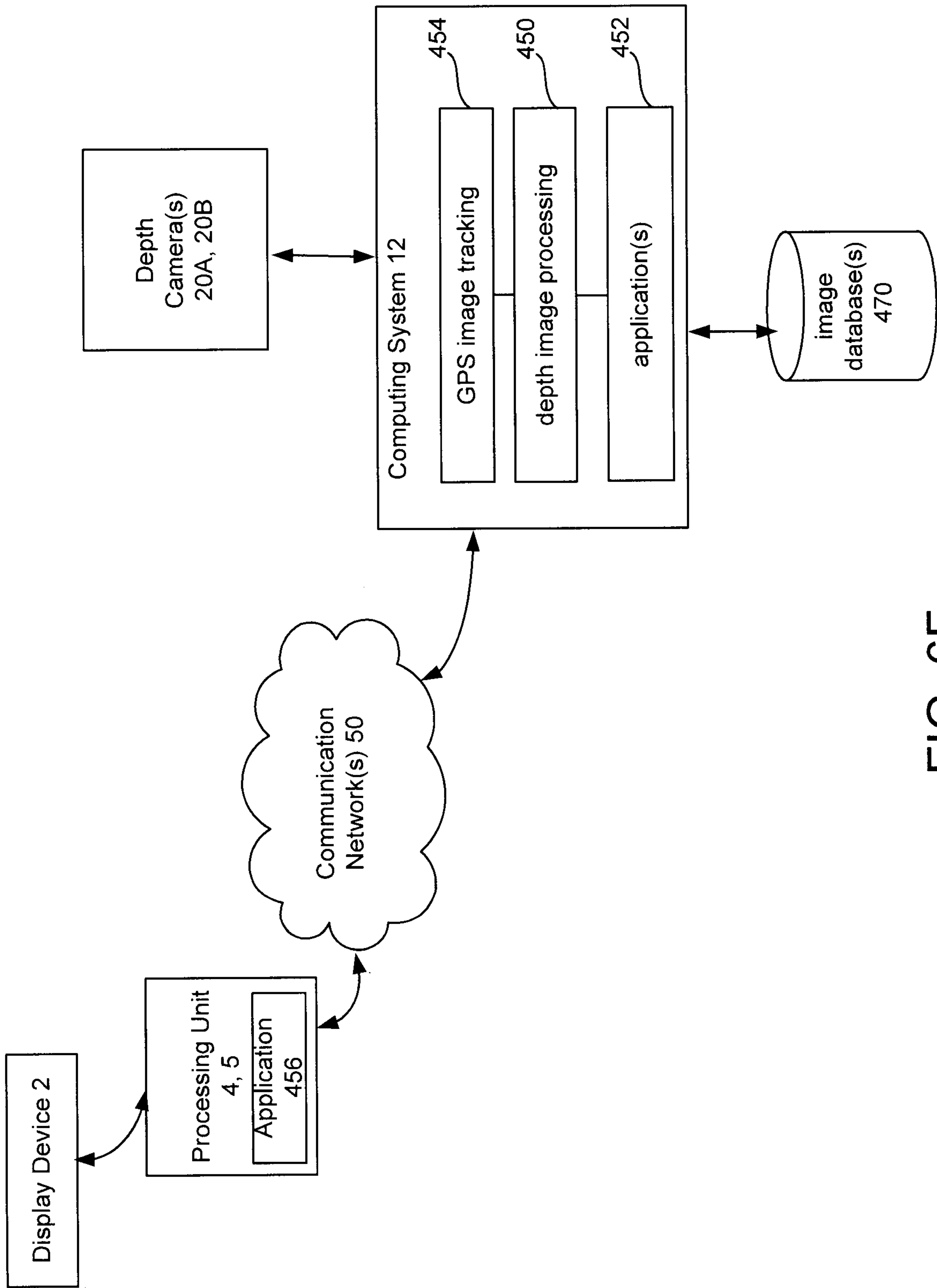
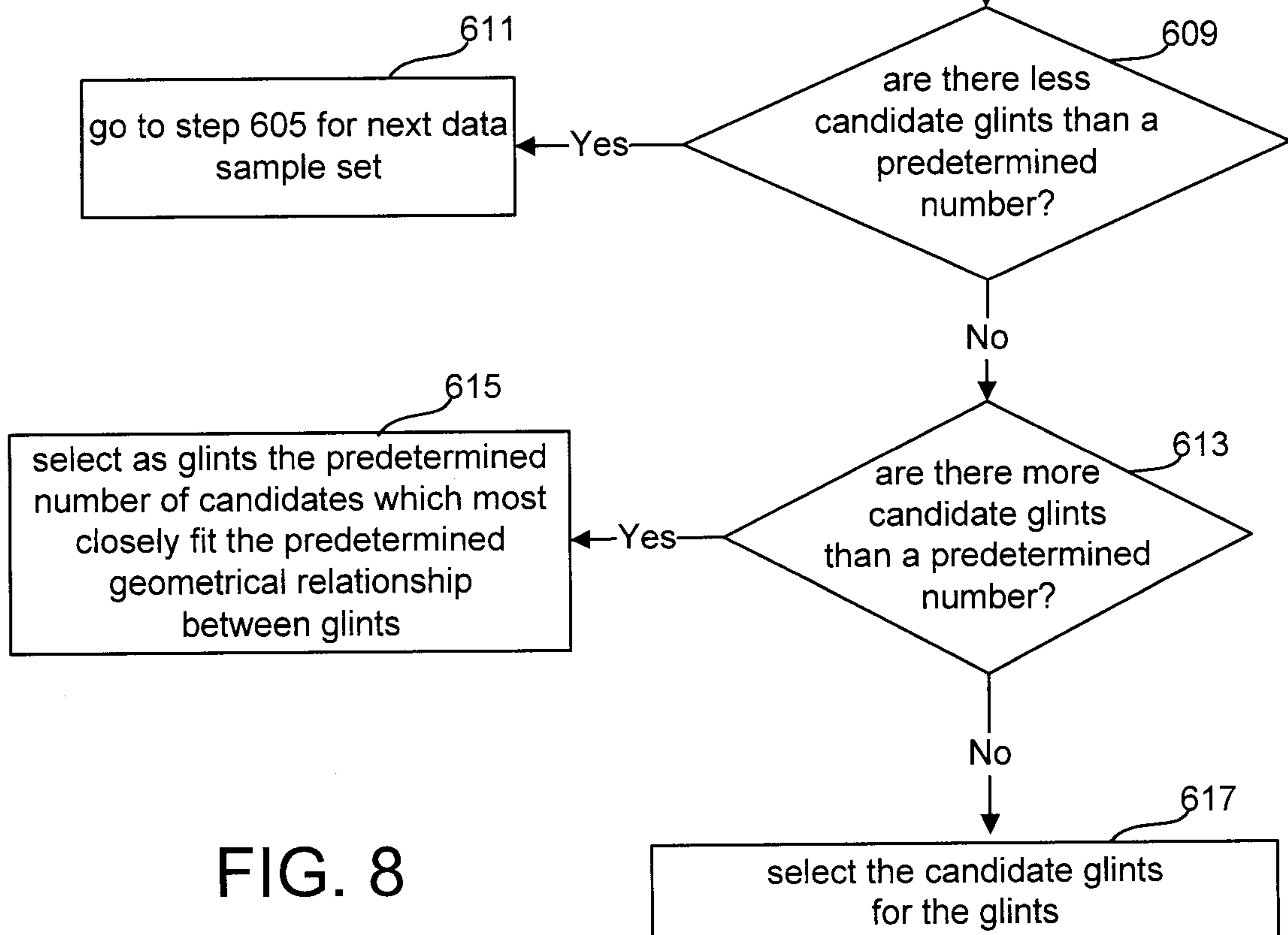
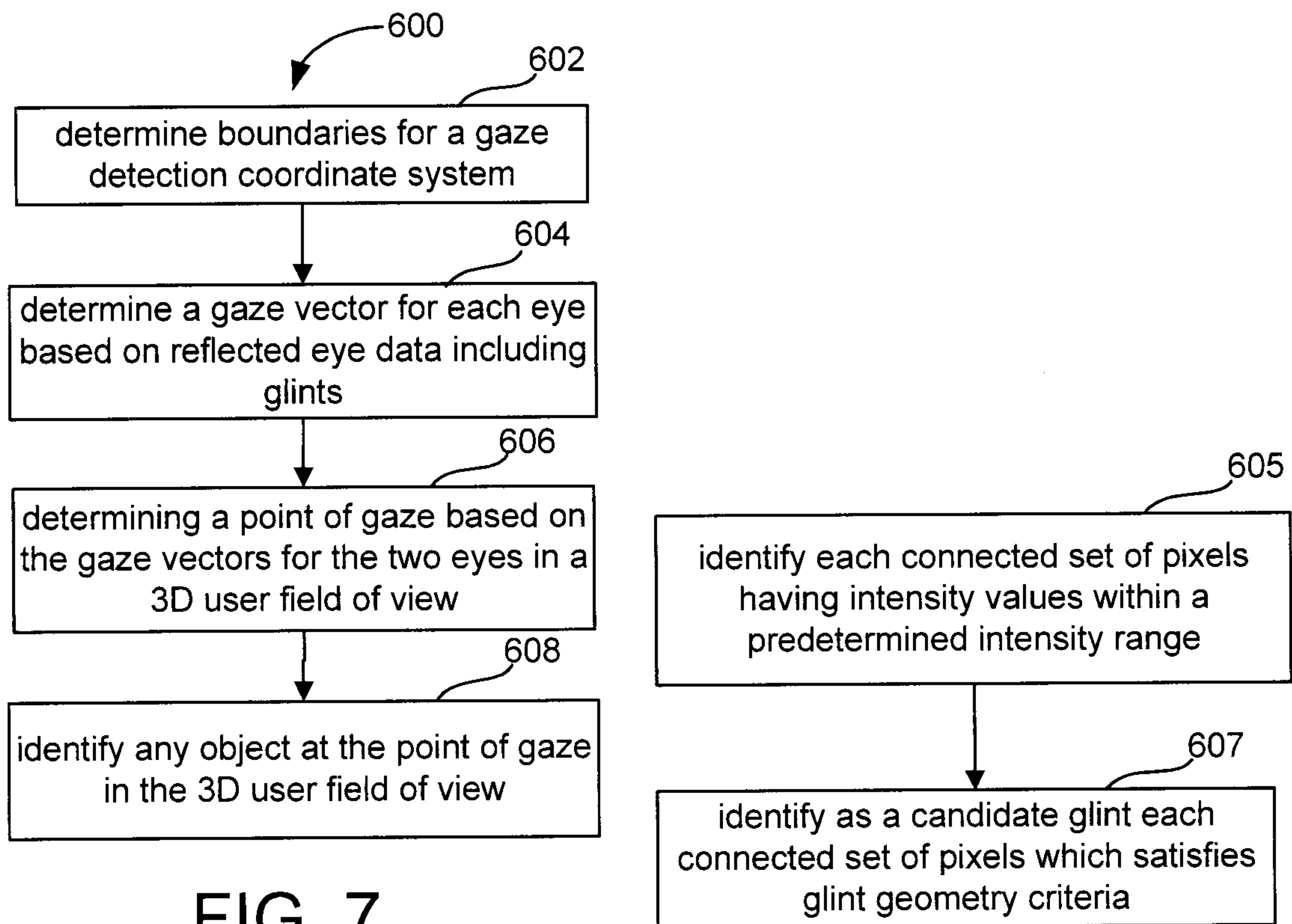


FIG. 6F

Title: GAZE DETECTION IN A SEE-THROUGH, NEAR-EYE, MIXED REALITY DISPLAY
 Applicant: John R. Lewis Docket No.: 01466US0
 Appl. No.: Unknown Atty: Eileen Lehmann
 Filing Date: Phone: (415) 369-9660



Title: GAZE DETECTION IN A SEE-THROUGH, NEAR-EYE, MIXED REALITY DISPLAY
 Applicant: John R. Lewis Docket No.: 01466US0
 Appl. No.: Unknown Atty: Eileen Lehmann
 Filing Date: Phone: (415) 369-9660

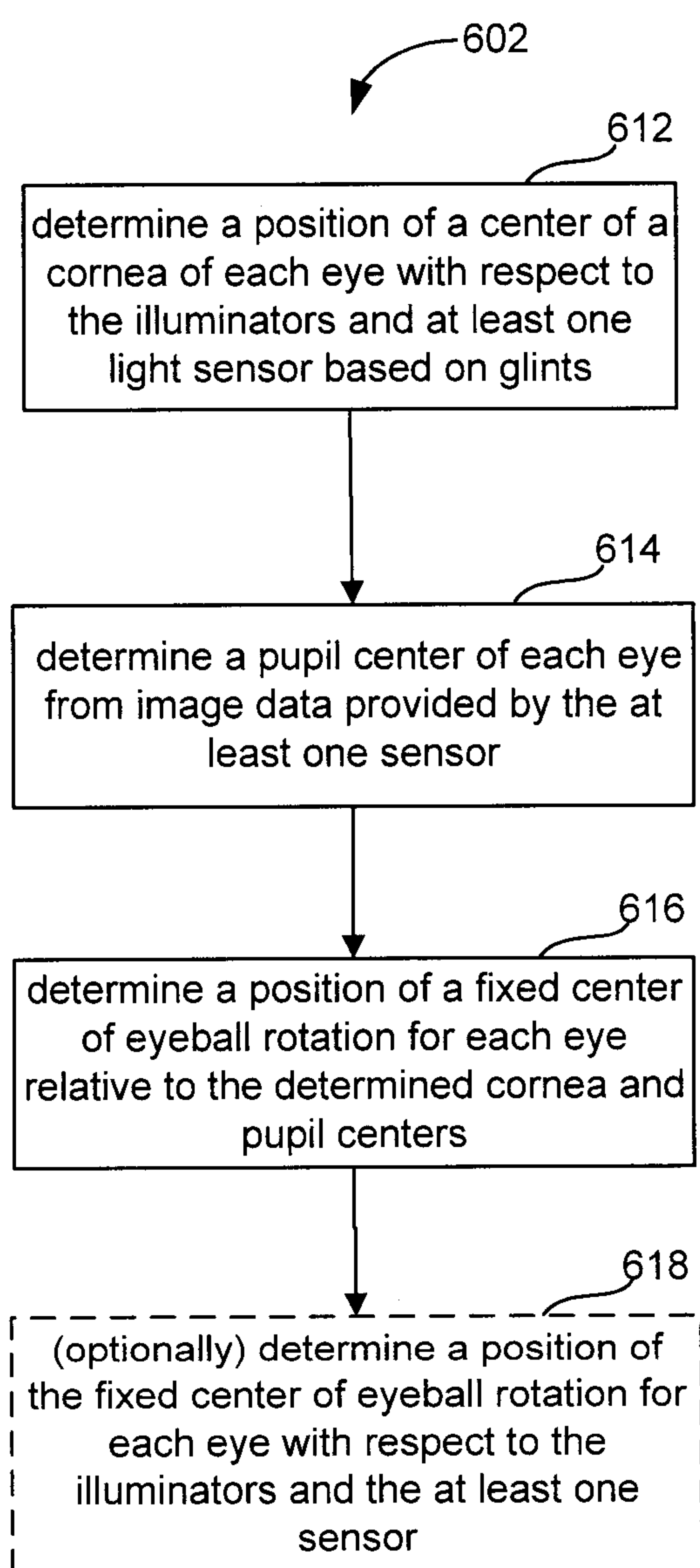


FIG. 9

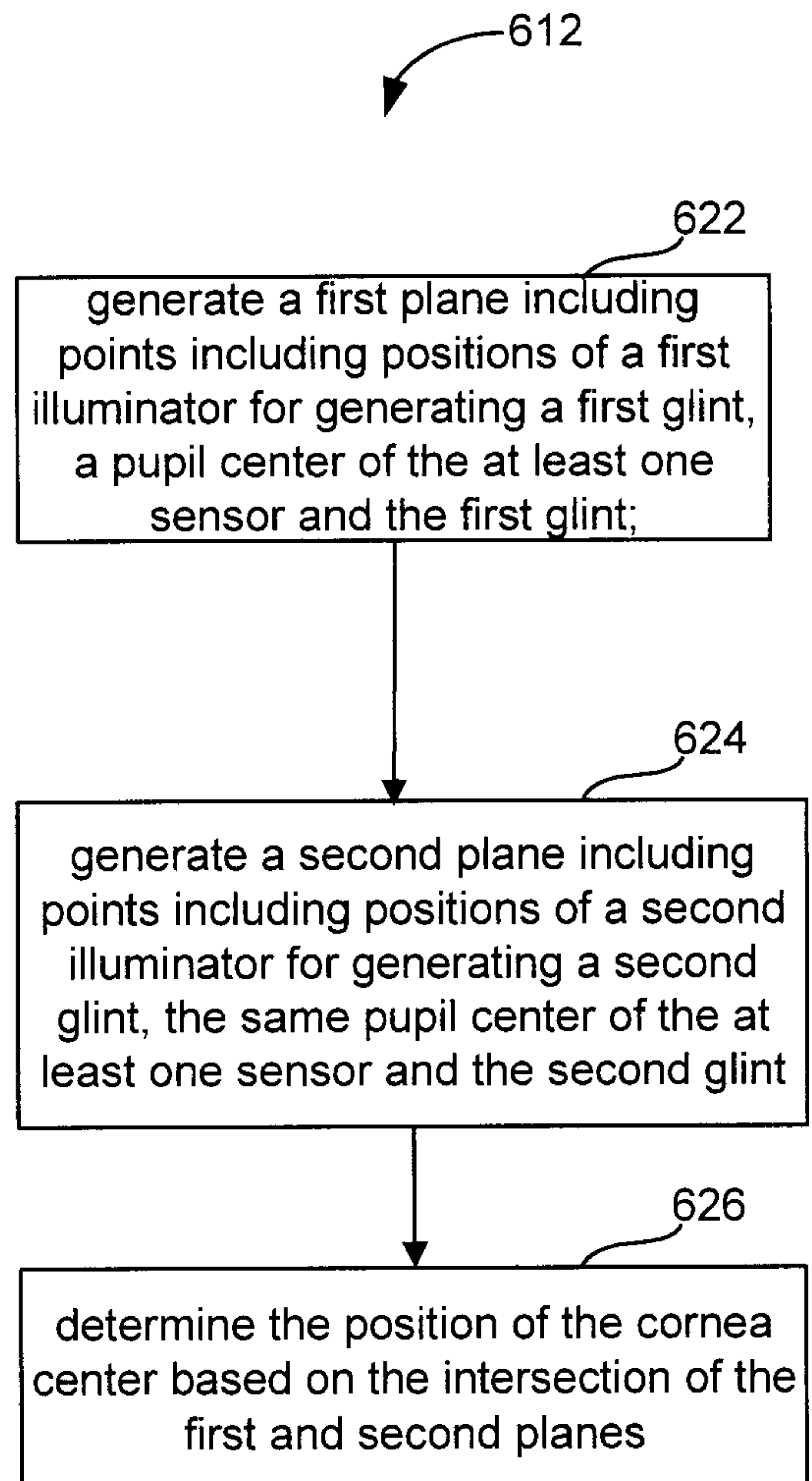


FIG. 10

Title:	GAZE DETECTION IN A SEE-THROUGH, NEAR-EYE, MIXED REALITY DISPLAY		
Applicant:	John R. Lewis	Docket No.:	01466US0
Appl. No.:	Unknown	Atty:	Eileen Lehmann
Filing Date:		Phone:	(415) 369-9660

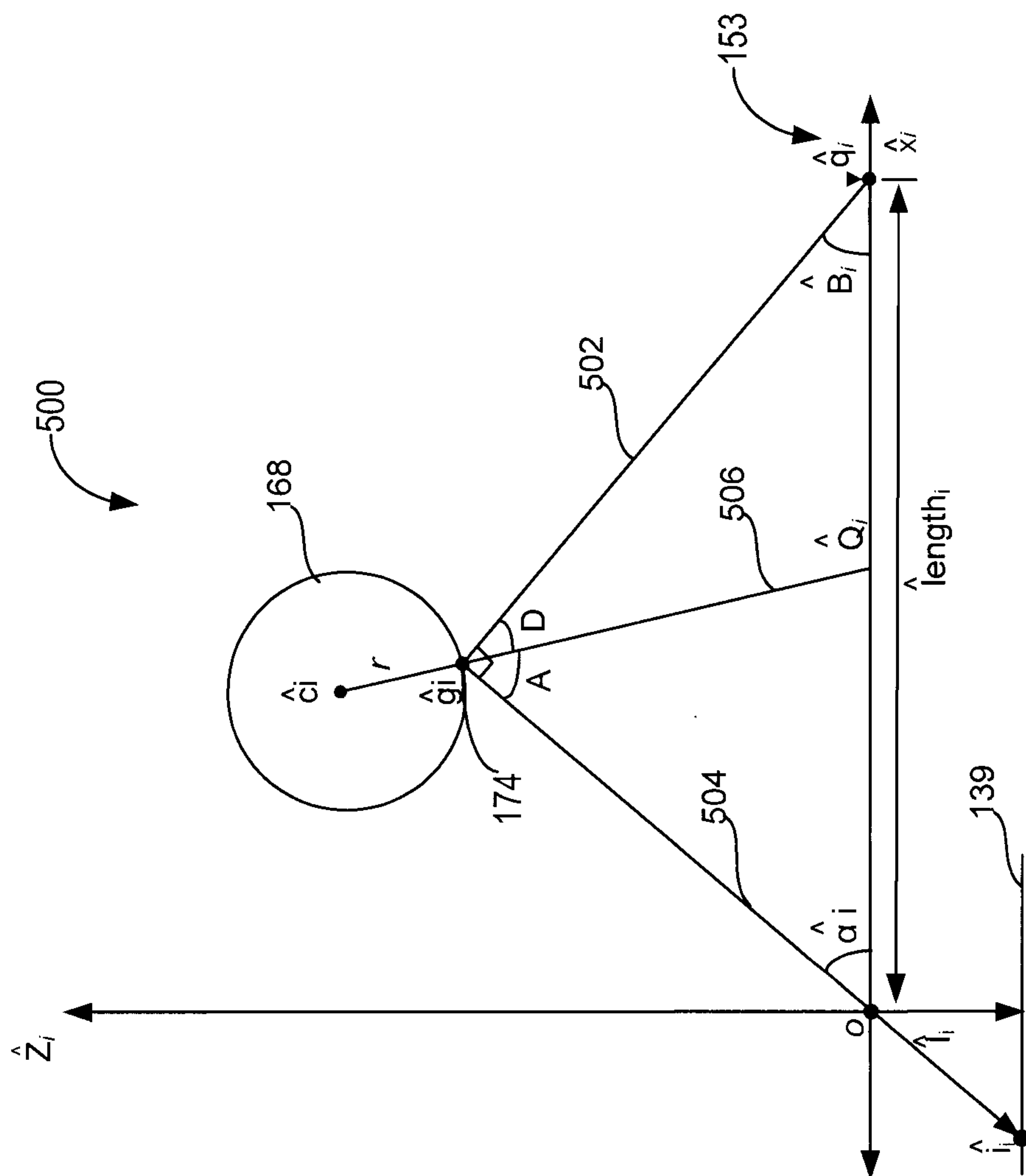


FIG. 11

Title: GAZE DETECTION IN A SEE-THROUGH, NEAR-EYE, MIXED REALITY DISPLAY
Applicant: John R. Lewis Docket No.: 01466US0
Appl. No.: Unknown Atty: Eileen Lehmann
Filing Date: Phone: (415) 369-9660

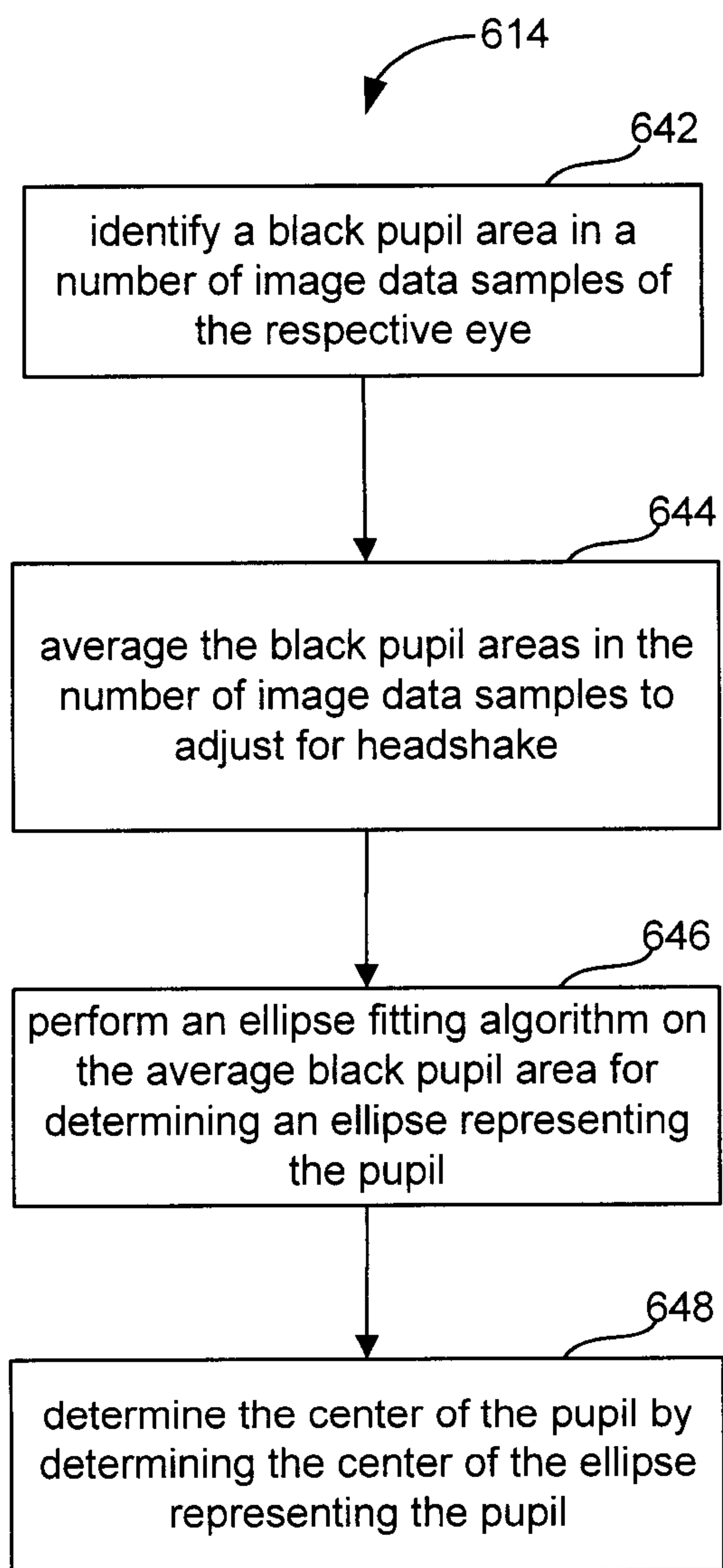


FIG. 12

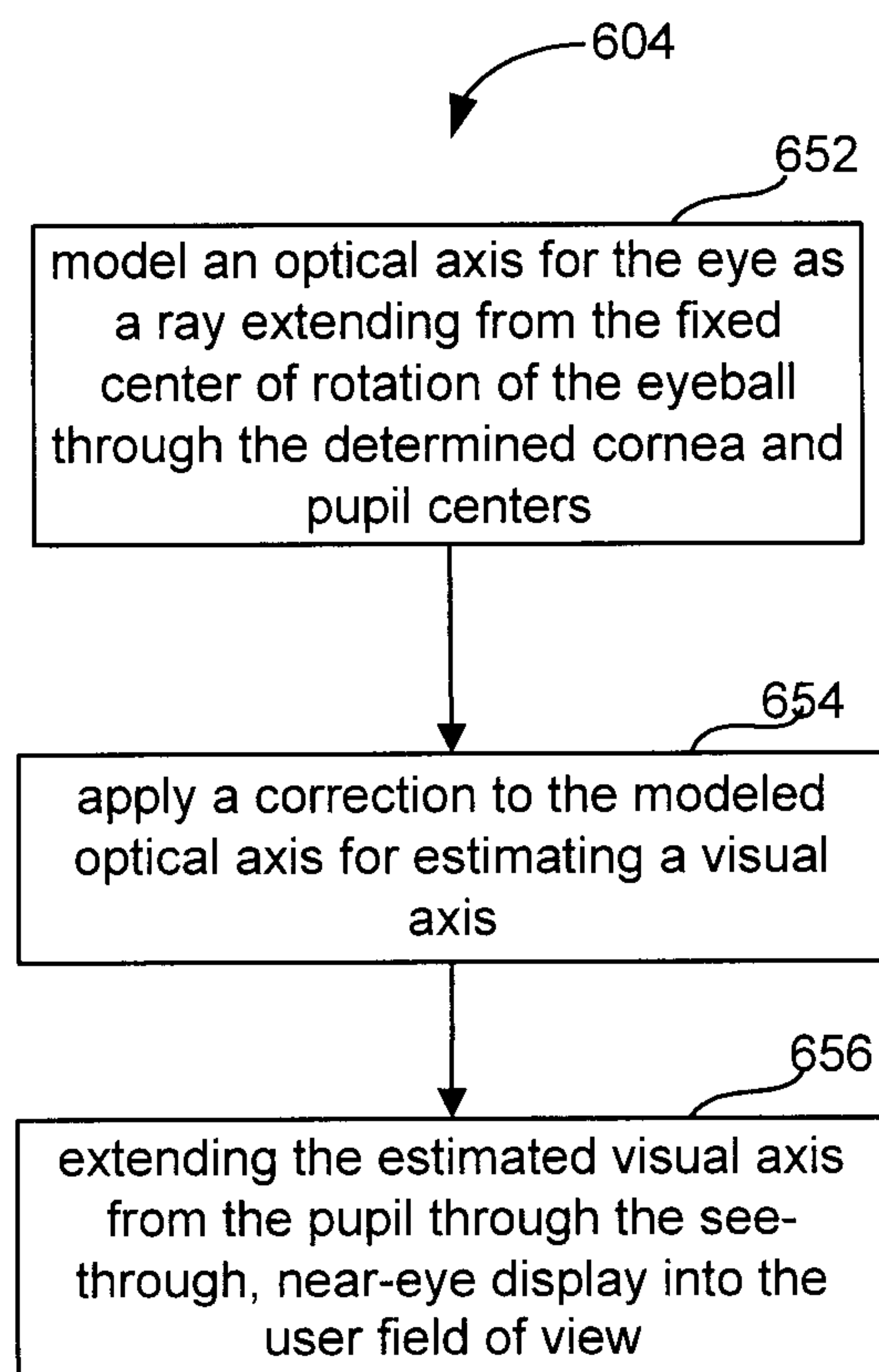


FIG. 13

Title: GAZE DETECTION IN A SEE-THROUGH, NEAR-EYE, MIXED REALITY DISPLAY
 Applicant: John R. Lewis
 Appl. No.: Unknown
 Filing Date:

Docket No.: 01466US0
 Atty: Eileen Lehmann
 Phone: (415) 369-9660

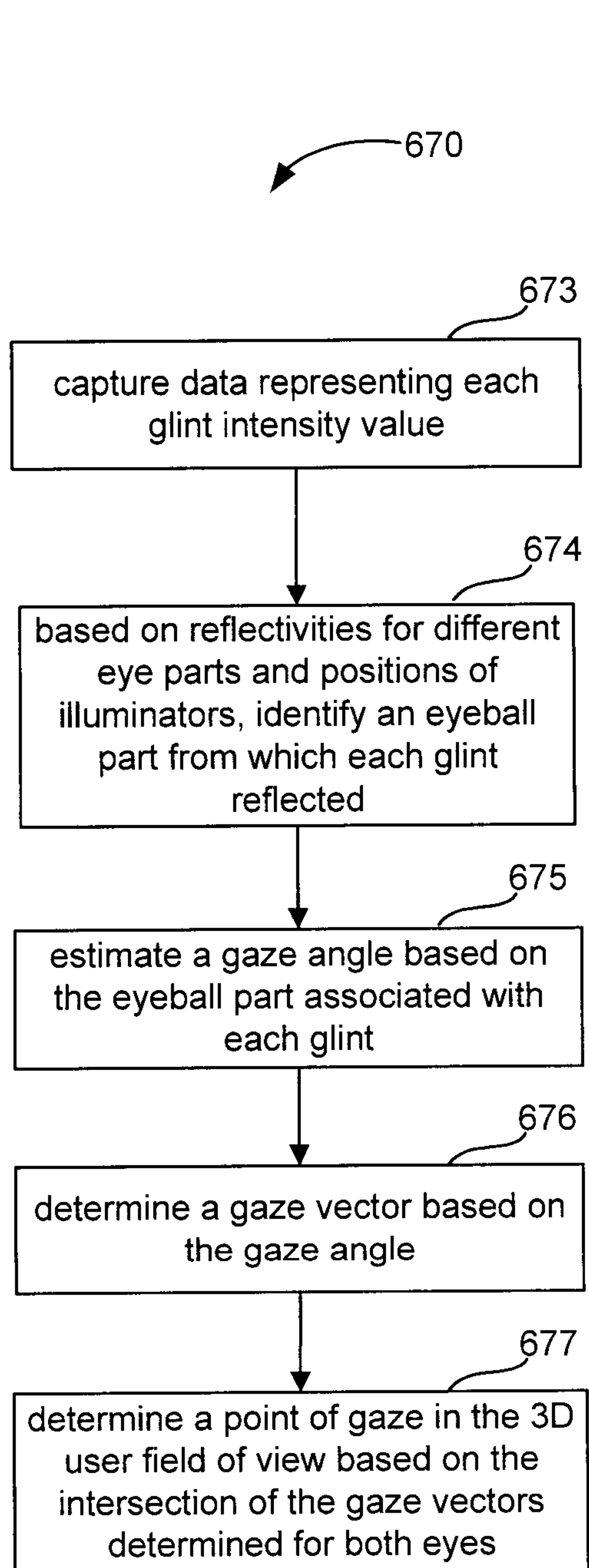


FIG. 14

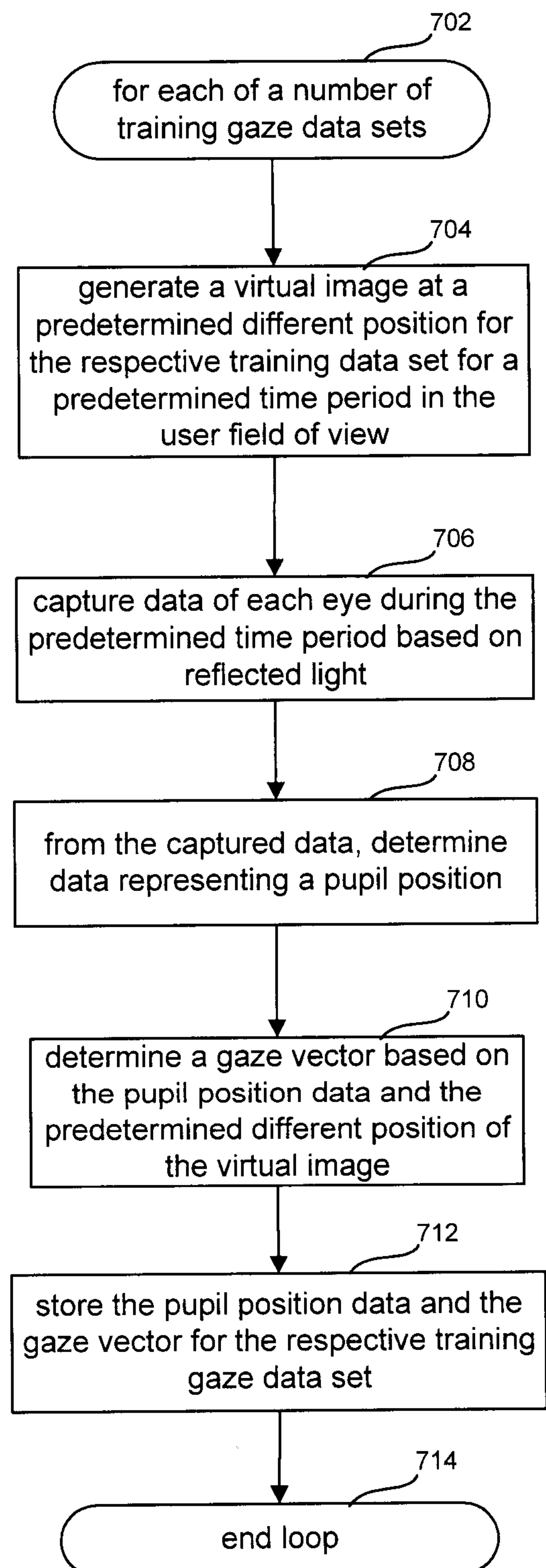


FIG. 15A

Title:	GAZE DETECTION IN A SEE-THROUGH, NEAR-EYE, MIXED REALITY DISPLAY		
Applicant:	John R. Lewis	Docket No.:	01466US0
Appl. No.:	Unknown	Atty:	Eileen Lehmann
Filing Date:		Phone:	(415) 369-9660

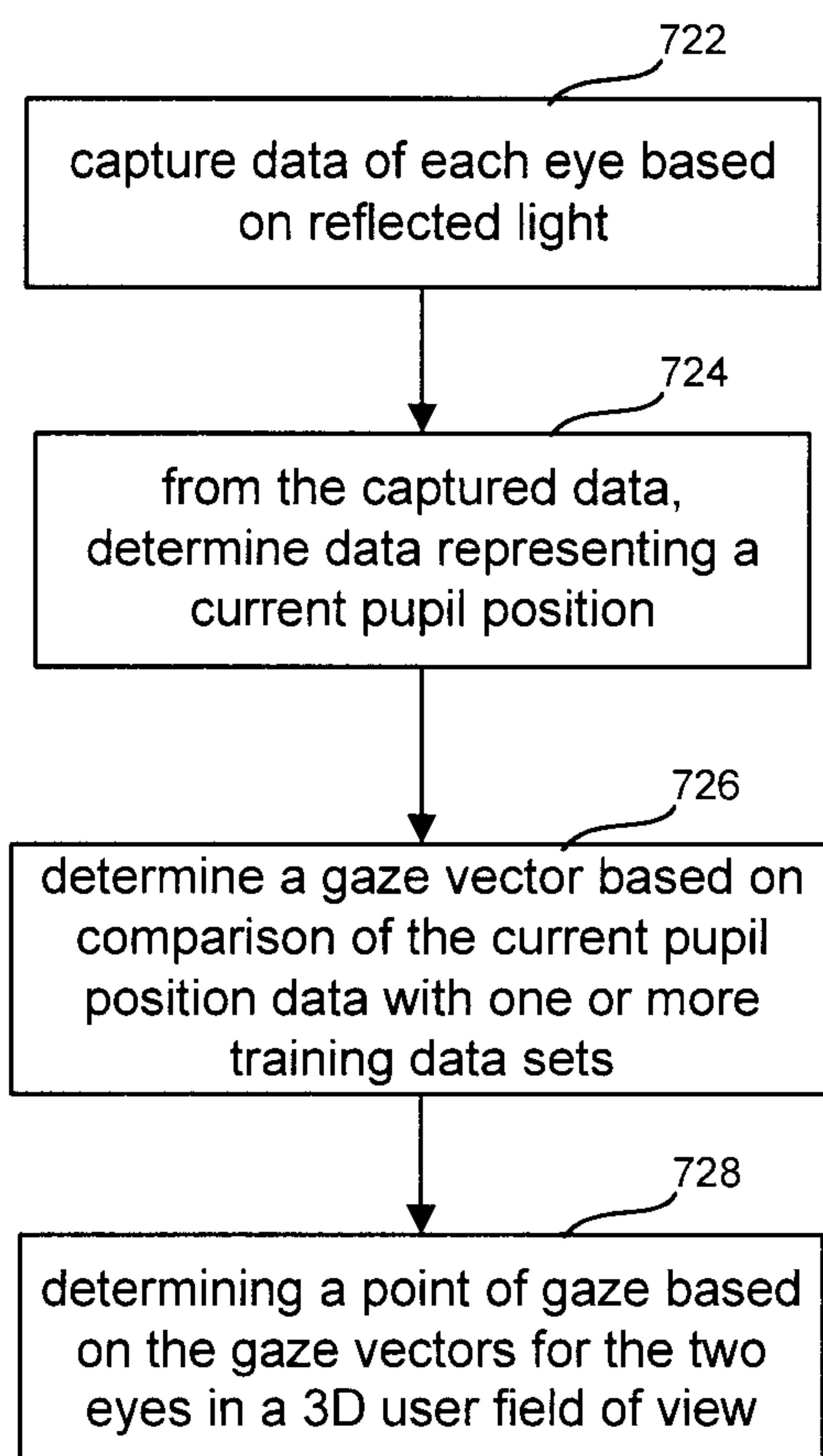


FIG. 15B

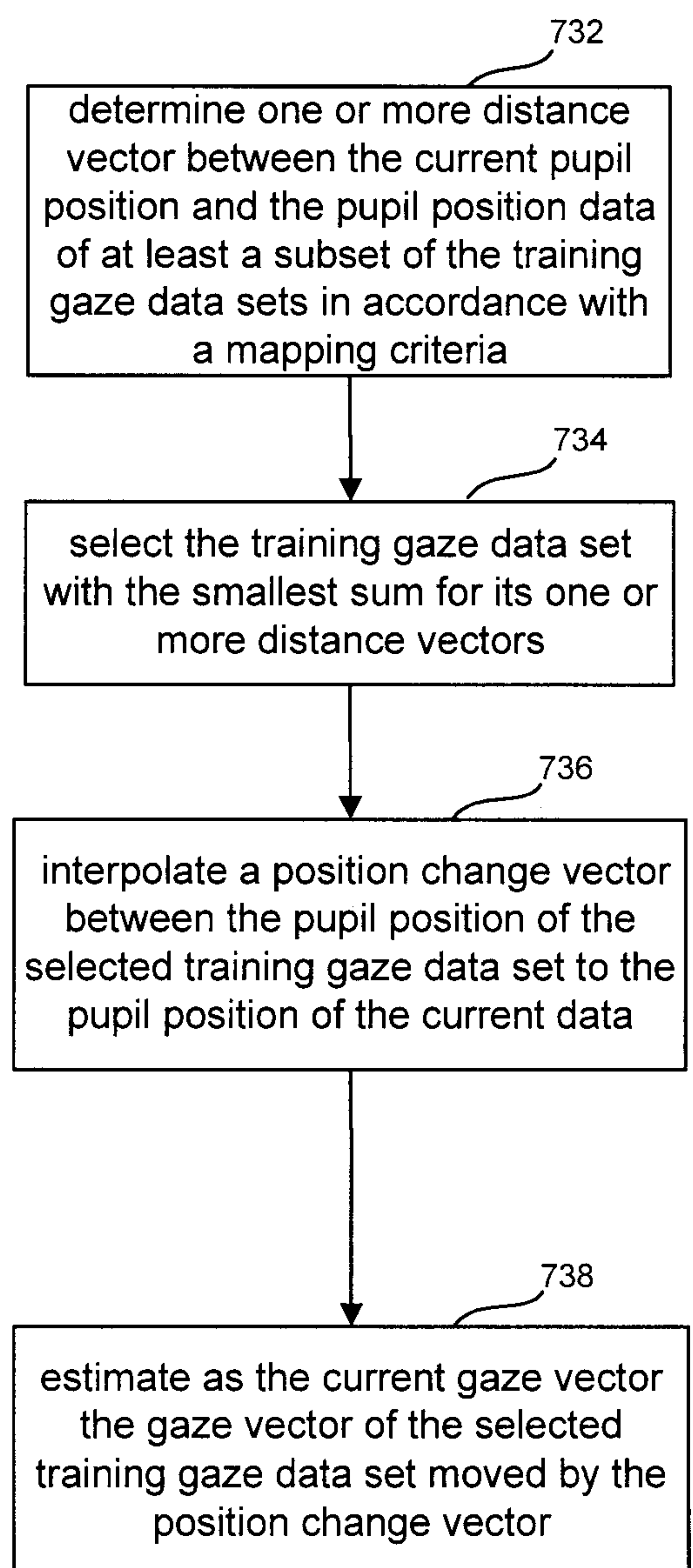


FIG. 15C

Title: GAZE DETECTION IN A SEE-THROUGH, NEAR-EYE, MIXED REALITY DISPLAY
Applicant: John R. Lewis
Appl. No.: Unknown
Filing Date: Docket No.: 01466US0
Atty: Eileen Lehmann
Phone: (415) 369-9660

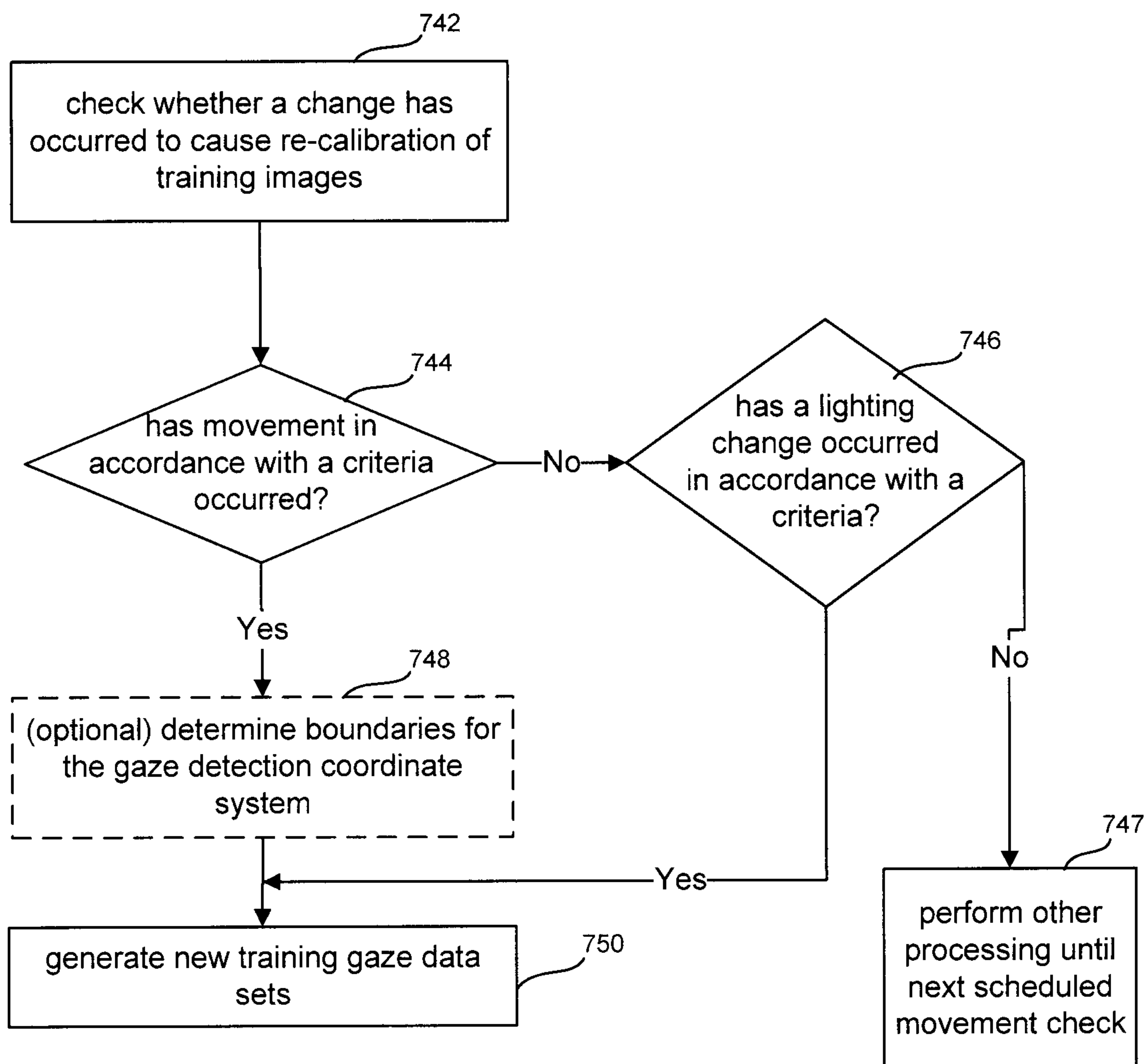
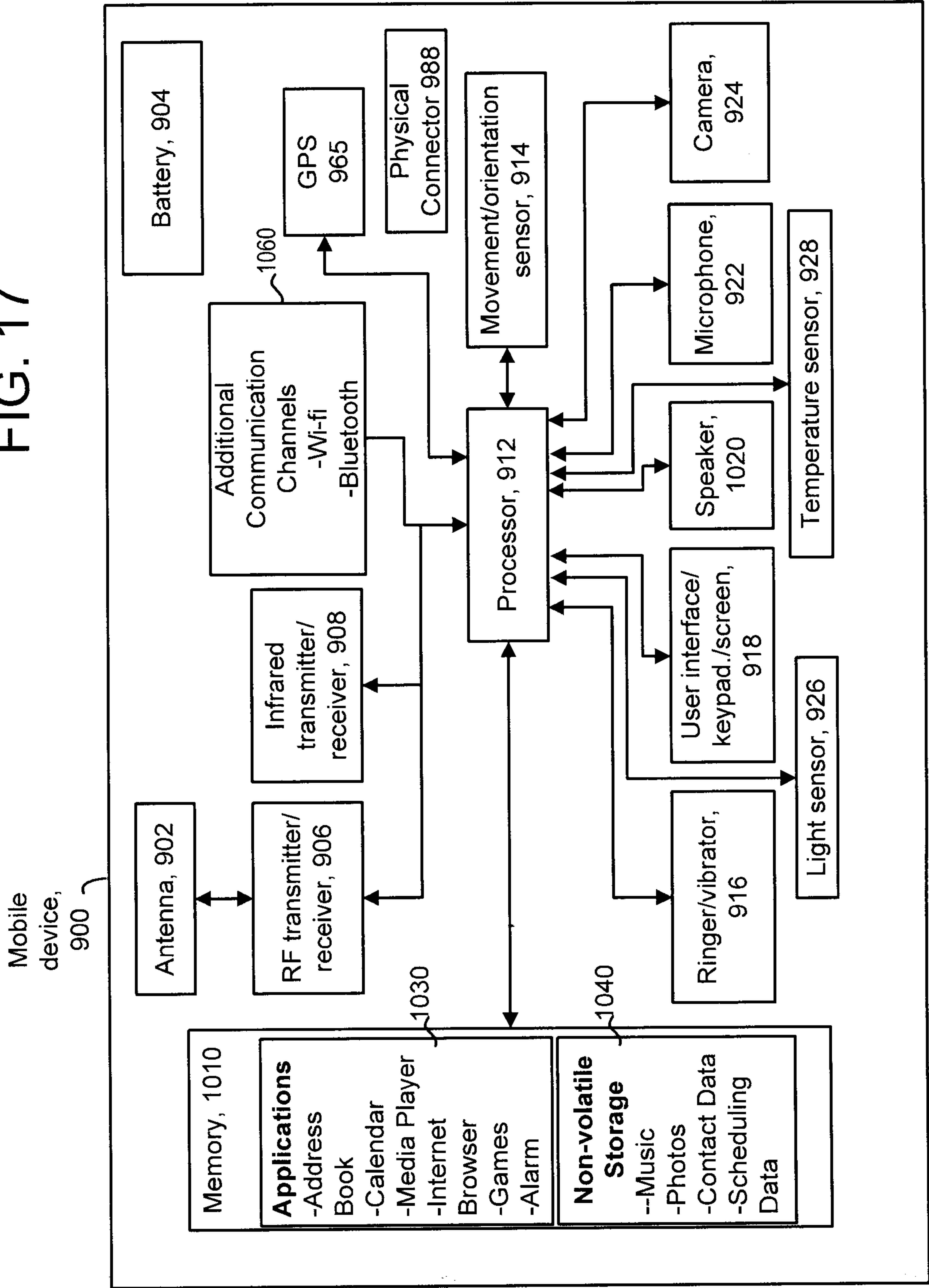


FIG. 16

FIG. 17



Title: GAZE DETECTION IN A SEE-THROUGH, NEAR-EYE, MIXED REALITY DISPLAY
 Applicant: John R. Lewis
 Appl. No.: Unknown
 Filing Date: Docket No.: 01466US0
 Atty: Eileen Lehmann
 Phone: (415) 369-9660

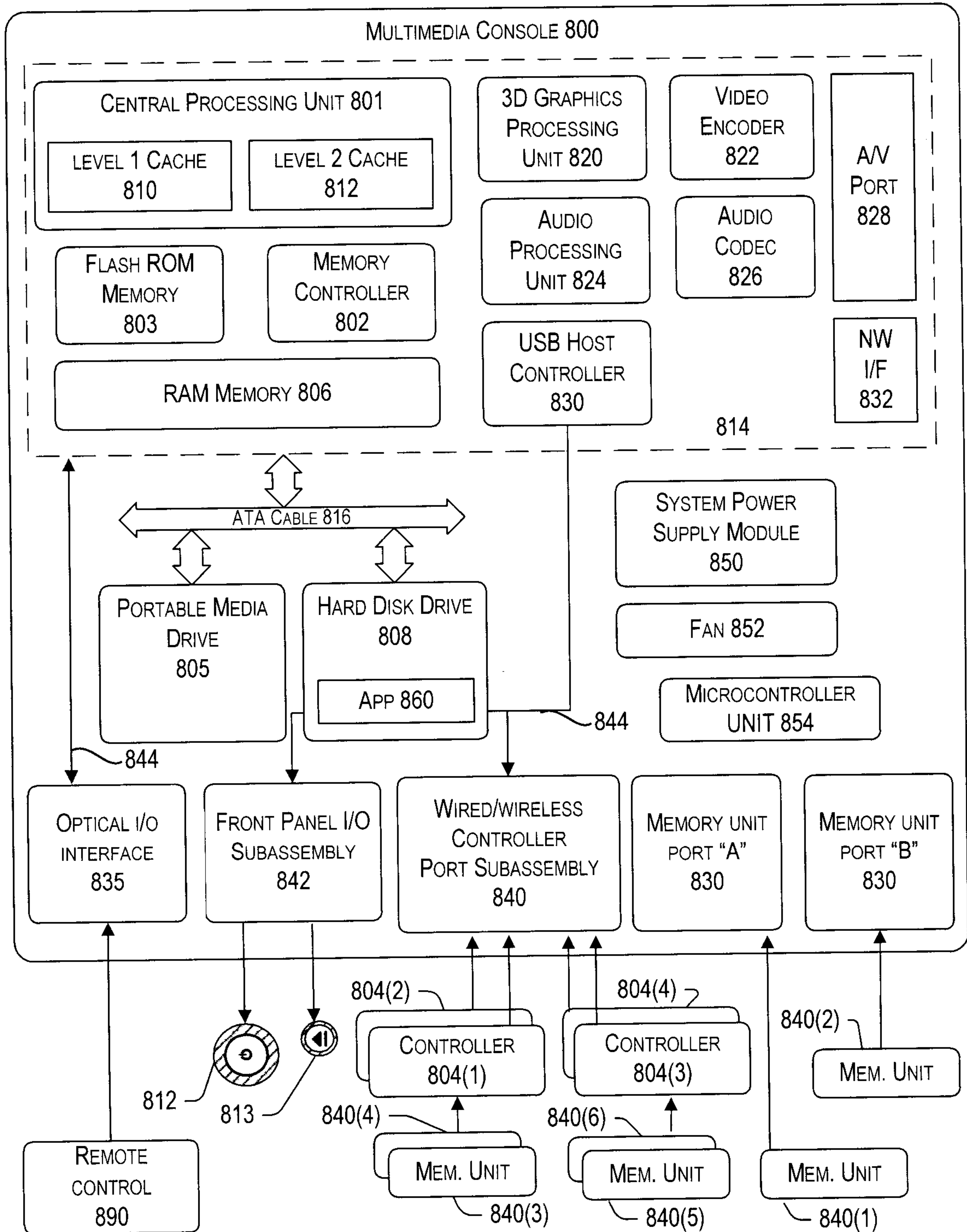


FIG. 18

