

(19) World Intellectual Property Organization
International Bureau



(43) International Publication Date
6 November 2008 (06.11.2008)

PCT

(10) International Publication Number
WO 2008/132724 A1

- (51) International Patent Classification:
G06F 3/01 (2006.01) H04N 13/00 (2006.01)
G06F 3/03 (2006.01)
- (21) International Application Number:
PCT/IL2008/000530
- (22) International Filing Date: 17 April 2008 (17.04.2008)
- (25) Filing Language: English
- (26) Publication Language: English
- (30) Priority Data:
60/924,003 26 April 2007 (26.04.2007) US
60/935,426 13 August 2007 (13.08.2007) US
- (71) Applicant (for all designated States except US): MAN-TISVISION LTD. [IL/IL]; 12 Yehudit Blvd., 67016 Tel Aviv (IL).
- (72) Inventors; and
- (75) Inventors/Applicants (for US only): GORDON, Eyal

[IL/IL]; 10 Zamanhoff St., 64373 Tel Aviv (IL). BITTAN, Gur Arie [IL/IL]; 48 Tarshish St., 60850 Shoham (IL).

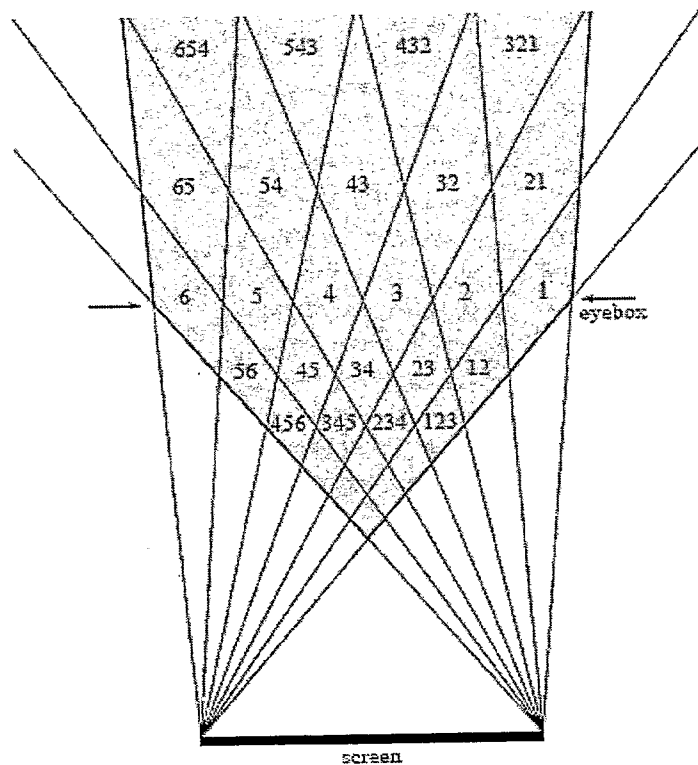
(74) Agent: FRIEDMAN, Mark; 7 Jabotinsky St., 52520 Ramat Gan (IL).

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

(84) Designated States (unless otherwise indicated, for every kind of regional protection available): ARIPO (BW, GH, GM, KE, LS, MW, MZ, NA, SD, SL, SZ, TZ, UG, ZM, ZW), Eurasian (AM, AZ, BY, KG, KZ, MD, RU, TJ, TM), European (AT, BE, BG, CH, CY, CZ, DE, DK, EE, ES, FI,

[Continued on next page]

(54) Title: A METHOD AND APPARATUS FOR THREE DIMENSIONAL INTERACTION WITH AUTOSTEROSCOPIC DISPLAYS



(57) Abstract: A method and apparatus for interactive human computer interface using a self-contained single housing autostereoscopic (AS) display configured to render 3D virtual objects into fixed viewing zones. The system contains an eye location tracking system for continuously determining both a viewer perceived three dimensional space in relation to the zones and a 3D mapping of the rendered virtual objects in the perceived space in accordance with a viewer eyes position. Additionally, one or more 3D cameras determines anatomy location and configuration of the viewer in real time in relation to said display. An interactive application that defines interactive rules and displayed content to the viewer is present. Furthermore, an interaction processing engine receives information from the eye location tracking system, the anatomy location and configuration system, and the interactive application to determine interaction data of the viewer anatomy with the rendered virtual objects from the AS display.

Fig 1

WO 2008/132724 A1



FR, GB, GR, HR, HU, IE, IS, IT, LT, LU, LV, MC, MT, NL,
NO, PL, PT, RO, SE, SI, SK, TR), OAPI (BF, BJ, CF, CG,
CI, CM, GA, GN, GQ, GW, ML, MR, NE, SN, TD, TG).

— *before the expiration of the time limit for amending the
claims and to be republished in the event of receipt of
amendments*

Published:

— *with international search report*

5

A METHOD AND APPARATUS FOR THREE DIMENSIONAL
INTERACTION WITH AUTOSTEROSCOPIC DISPLAYS

FIELD AND BACKGROUND OF THE INVENTION

10

The present invention relates to an apparatus and a method for three dimensional interaction between a viewer and a virtual scene displayed on an autostereoscopic display, and more particularly but not exclusively, to the touching of and/or control of objects in the autostereoscopic scene.

15

Stereoscopic systems have gained in popularity in recent years as the capabilities of computer processing power have grown enormously along with advances in three dimensional screen technologies. Autostereoscopic (AS) display technology allows the viewer to experience the sensation that 3D objects are floating in front of him without the use of any visual aids. As a result of display optics, the area in front of the display is divided into viewing zones, wherein each zone is typically several centimeters wide. Fig 1 shows a typical AS display along with the viewing zones, where zones numbered 1-6 show the optimal viewing distance from the screen, called the eyebox. A viewer's position in front of the display is defined by any two adjacent zones, each eye falling in a separate viewing zone. Each eye in each zone perceives a different image from the AS display, and the two images together
25 give the viewer the 3D sensation, called stereo parallax.

20

Furthermore, with a multi-viewing-zone screen, as in Fig 1, the viewer may experience limited motion parallax as well as stereo parallax. Motion parallax gives the viewer the sensation of actually seeing a different angle of the object as he moves his position in relation to the object. That is to say, as the viewer moves laterally in front of the screen, he views different pairs of images for each pair of adjacent viewing zones. As a result, he may perceive the virtual object displayed on the AS display at various angles as he moves laterally in front of the screen.
30

To illustrate, we refer to Figs 2A and 2B. In Fig 2A, the viewer's two eyes are found in viewing zones 1 and 2 in relation to AS display 14. As a result, he perceives virtual coke bottle 16 as having a particular mapping of virtual 3D coordinates. This mapping is denoted by $([x_{i1}], [y_{i1}], [z_{i1}])$ in the figure and represents a set of 3D coordinates in space that comprise the virtual coke bottle. In Fig 2B, we see the viewer has moved and his eyes now are found in viewing zones 2 and 3. As a result, he perceives the virtual coke bottle at a different angle, giving him the feeling of motion parallax. Furthermore, the set of 3D coordinates in space that comprise the perceived virtual coke bottle has changed, and is now $([x_{i2}], [y_{i2}], [z_{i2}])$. So therefore, for each pair of viewing zones in front of the AS display, a mapping exists of perceived coordinates of the displayed virtual object(s). It is understood that the motion parallax described with reference to the viewing zones is often not seamless and the motion parallax only occurs when the viewer jumps zones. That is to say, movement within the same zones may not be seen as motion parallax in less sophisticated systems.

Referring back to Fig 1, as the viewer approaches close to the screen, his left eye may be found in diamond 345 and his right eye in diamond 234, for instance. These zones represent mixtures of three images each. In such a case, each of the viewer's eyes will see a mixture of three different images, and the viewing experience is hampered. The same phenomenon will occur if he steps farther away from the screen, outside of the eyebox diamonds numbered 1-6. The viewing zones 1-6 are therefore the optimal viewing zones from the AS display in which the viewer perceives clear stereoscopic images. Although only x-direction division of the screen resolution is illustrated for simplicity in the figures, viewing zones may be allocated according to both x and y direction. As such, content perceived by a viewer may change in accordance with both horizontal and vertical position in relation to the display.

As stated, the multiple viewing zones are a result of display optics, typically achieved by use of lenticular lenses or parallax barriers over the flat panel display screen. Although the zones are necessary to allow for perception of a different image in each eye, the drawback is the significant decrease in overall screen resolution. For instance, in Fig 1, for a 6 viewing-zone display, the resolution for any one view is 1/6 of the overall screen resolution. One approach to compensate for the decrease in

screen resolution is the use of adjustable optical filters. Through a tracking mechanism that determines the viewer's eyes position in relation to the display, the optical filters adjust the location of the viewing zones in accordance with the viewer's eyes location. Then, only the appropriate two images for the viewer's current position are displayed. In such a manner, the screen resolution may be divided into two zones and the decrease in resolution caused by multiple zones may be significantly decreased. Such a method is described in US Patent No. 6,075,557.

Autostereoscopic displays ultimately provide a more sensational viewing experience, as the viewer has the sensation that the objects in the displayed scene are floating in front of his eyes without the use of any visual aid or tracking device. Additionally, a virtual hologram experience may give the viewer the perception of 3D into the screen as well as the sensation that objects are floating in front of the screen.

When the objects of a scene are indeed floating in front of one's eyes, there is a natural tendency to want to touch and otherwise interact with the virtual object. However, several problems exist that heretofore have not been solved. Firstly, as explained above, the perceived object has a set of coordinates that are unique to the viewing zone in which the viewer's eyes are found and the viewer's position within the viewing zone. Secondly, known devices for interacting with displayed technology, whether 2D or 3D, generally require cumbersome 'user control' tracking device(s) either held by the viewer or mounted to one or more locations of the viewer's hand. Such systems may include equipment that the viewer must wear, such as gloves and helmets having tracking devices thereupon. Unfortunately, these systems do not allow for ease-of-use, quick response, and simplicity so necessary in interactive environments, such as video games.

Therefore, there is an unmet need for, and it would be highly useful to have an enhanced system and method to allow for interaction in 3D environments to overcome the above drawbacks.

SUMMARY OF THE INVENTION

5 According to one aspect of the present invention there is provided an apparatus for providing an interactive human computer interface to a viewer, the apparatus comprising:

- 10 a) a data storage configured to store three-dimensional virtual environment representation data including at least one three-dimensional virtual object within the virtual environment;
- b) an autostereoscopic (AS) display configured to display simultaneous perspectives of virtual objects of the 3D virtual environment representation data to spatially separated viewing zones located in front of the AS display
- 15 c) an anatomy tracking system including at least one 3D camera, the anatomy tracking system operative to determine respective real-world locations of:
 - i) a viewer's eyes; and
 - ii) an anatomical part on the viewer's body
- d) a 3D registration engine configured to generate a 3D volume comprising
20 3D location data of the viewer's anatomical part and 3D location data of the virtual objects perceived by the viewer in accordance with the viewer's eye location,
- e) an anatomical part-virtual object relation computation engine operative to determine a relation between the virtual object and the anatomical part in accordance with output of the registration engine:
- 25 f) a rule enforcement engine operative to modify the three-dimensional environment representation data in accordance with the determined anatomical part-virtual object relation and the virtual environment data.

According to another aspect of the present invention there is an apparatus for providing an interactive human computer interface to a viewer, the apparatus
30 comprising:

- a) a data storage configured to store three-dimensional virtual environment representation data including at least one three-dimensional virtual object within the virtual environment;

b) an autostereoscopic (AS) display configured to display simultaneous perspectives of virtual objects of the 3D virtual environment representation data to spatially separated viewing zones located in front of the AS display

c) an anatomy tracking system including at least one 3D camera, the anatomy tracking system operative to determine respective real-world locations of:

i) a viewer's eyes; and

ii) an anatomical part on the viewer's body

d) an anatomical part-virtual object relation computation engine operative to determine a relation between a perceived virtual object by the viewer and the anatomical part

e) a rule enforcement engine operative to modify the three-dimensional environment representation data in accordance with the determined anatomical part-virtual object relation and the virtual environment data.

According to another aspect of the present invention there is an autostereoscopic (AS) display configured to display simultaneous perspectives of virtual objects of the 3D virtual environment representation data to spatially separated viewing zones located in front of the AS display, the AS display having an acoustic lens with electrically controlled refractive index to dynamically adapt the viewing zone locations.

According to another aspect of the present invention there is provided a method for providing an interactive human computer interface to a viewer, the method comprising:

a) storing three-dimensional virtual environment representation data including at least one three-dimensional virtual object within the virtual environment;

b) displaying on an autostereoscopic (AS) display simultaneous perspectives of virtual objects in the 3D virtual environment representation data to spatially separated viewing zones located in front of the AS display

c) tracking the anatomy of a viewer to determine respective real-world locations of:

i) a viewer's eyes; and

ii) an anatomical part on the viewer's body

d) generating a registered 3D volume comprising 3D location data of the viewer's anatomical part and 3D location data of the virtual objects perceived by the viewer in accordance with the viewer's eye location,

e) determining a relation between the virtual object and the anatomical part in accordance with the registration,

f) modifying based on interactive rules the three-dimensional environment representation data in accordance with the determined anatomical part-virtual object relation and the virtual environment data.

According to another aspect of the present invention there is provided a system for interactive human computer interface, the system comprising:

a self-contained autostereoscopic (AS) display configured to render 3D virtual objects into neighboring viewing zones associated with the display,

an eye location tracking system, comprising at least one 3D video camera, for continuously determining:

1) a viewer perceived three dimensional space in relation to the display, and

2) a 3D mapping of the rendered virtual objects in the perceived space in accordance with viewer eyes position in relation to the fixed viewing zones

an anatomy location and configuration system, comprising at least one 3D video camera, for continuously determining a 3D mapping of viewer anatomy in relation to the display, and

an interactive application that defines interactive rules and displayed content to the user, and

an interaction processing engine configured to receive information from

1) the eye location tracking system

2) the anatomy location and configuration system, and

3) the interactive application

thereby to determine interaction data of the viewer anatomy with the rendered virtual objects from AS display.

Unless otherwise defined, all technical and scientific terms used herein have the same meaning as commonly understood by one of ordinary skill in the art to which this invention belongs. The materials, methods, and examples provided herein are illustrative only and not intended to be limiting.

5 Implementation of the method and system of the present invention involves performing or completing certain selected tasks or steps manually, automatically, or a combination thereof. Moreover, according to actual instrumentation and equipment of preferred embodiments of the method and system of the present invention, several selected steps could be implemented by hardware or by software on any operating
10 system of any firmware or a combination thereof. For example, as hardware, selected steps of the invention could be implemented as a chip or a circuit. As software, selected steps of the invention could be implemented as a plurality of software instructions being executed by a computer using any suitable operating system. In any case, selected stages of the method and system of the invention could be described as
15 being performed by a data processor, such as a computing platform for executing a plurality of instructions.

20

BRIEF DESCRIPTION OF THE DRAWINGS

The invention is herein described, by way of example only, with reference to the accompanying drawings. With specific reference now to the drawings in detail, it
25 is stressed that the particulars shown are by way of example and for purposes of illustrative discussion of the preferred embodiments of the present invention only, and are presented in order to provide what is believed to be the most useful and readily understood description of the principles and conceptual aspects of the invention. In this regard, no attempt is made to show structural details of the invention in more
30 detail than is necessary for a fundamental understanding of the invention, the description taken with the drawings making apparent to those skilled in the art how the several forms of the invention may be embodied in practice.

In the drawings:

FIG. 1 is a simplified illustration of a multi view autostereoscopic screen along
35 with associated viewing zones.

FIG. 2A is simplified illustration showing the viewing zones associated with a particular viewer location in relation to an autostereoscopic display.

FIG. 2B is simplified illustration showing the viewing zones associated with a different viewer location in relation to an autostereoscopic display.

5 FIG. 3A is a simplified flow chart illustrating a generalized embodiment of the invention.

FIG 3B is a simplified flow chart of the interaction processing engine.

FIG. 4A is a simplified illustration of the viewer interacting with a virtual object in one location in relation to the AS display according to preferred
10 embodiments of the present invention.

FIG. 4B is a simplified illustration of the viewer interacting with a virtual object in another location in relation to the AS display according to preferred embodiments of the present invention.

FIG 4C is a simplified illustration of the processing engines in preferred
15 embodiments.

FIG. 5 illustrates several visual examples of interaction between a real object and a virtual 3D object.

FIG. 6 is a preferred embodiment showing a three dimensional autostereoscopic menu driven interface.

20 FIG. 7 shows a preferred embodiment wherein viewer classification rules enable the system to identify various characteristics of the viewer and even identification and association with a stored historical profile

FIG. 8 shows the generalized embodiment of Fig. 3 with the addition of a personal profile engine.

25 FIG. 9 is a preferred embodiment of the present invention illustrating gesture-based interaction with 3D objects on an AS display.

FIG. 10 illustrates exemplary gesture based interactions.

FIG. 11 is a simplified illustration of an interface with the virtual scene based on a scaled model of the user seen in the virtual space.

30 FIG. 12 is a illustrates a viewer or anatomical part of a viewer and the parallel "virtual viewer" or "virtual anatomical part" being displayed in the viewer perceived space

FIG. 13A is a simplified illustration showing the 3D camera field of view according to preferred embodiments.

FIG. 13B is a simplified illustration showing the process of eye tracking and position finding.

5 FIG. 13C is a simplified illustration showing the process of 3D skeleton modeling.

FIG.14 is a simplified diagram illustrating a lenticular lens based autostereoscopic display used in preferred embodiments of the present invention.

10 FIG 15 is a simplified illustration of the perceived 3D space in relation to the pixels on the AS display in which the viewer may observe and interact with virtual objects and scenes.

FIG. 16 is a simplified illustration of the interaction zone showing the overlap between the perceived space and the camera field of view for an exemplary case.

15 FIG. 17 shows an alternative embodiment in which the viewer's eyes position data obtained from the 3D camera may also be input to a screen steering device.

FIG. 18 shows a network of 3D virtual touch human interface systems.

20 DESCRIPTION OF THE PREFERRED EMBODIMENTS

An apparatus and a method are now disclosed for three dimensional interaction between a viewer and virtual object(s) displayed on an autostereoscopic (AS) display. According to some embodiments, the viewer is able to experience the
25 sensation of touching and/or control of objects perceived on the autostereoscopic display solely through use of an anatomical part of the viewer interfacing with the display.

More specifically, the present embodiments provide an immersive and three dimensional real time interactive environment through the use of one or more 3D
30 cameras. A virtual object(s) is displayed on an AS display and appears to the viewer in a defined region in front of the display determined by the viewer's eyes position in relation to the display, as explained above. In preferred embodiments, the viewer may manipulate the object in various ways and may interact with the object in real time. These manipulations are then reflected in subsequent frames rendered and perceived

by the viewer in the region for virtual objects in front of the display defined by the user's eyes location. The process creates an "immersive" interactive environment between the viewer and the virtual scene perceived on the auto stereoscopic display.

In some embodiments, through 3D data obtained about the viewer's facial features, a viewer may be identified according to his or her stored profile. In such
5 embodiments, the interaction process can be suited towards that particular viewer's historical preference data and other unique characteristics associated with that viewer in particular.

The principles and operation of an apparatus and method according to the
10 present invention may be better understood with reference to the drawings and accompanying description.

Before explaining at least one embodiment of the invention in detail, it is to be understood that the invention is not limited in its application to the details of construction and the arrangement of the components set forth in the following
15 description or illustrated in the drawings. The invention is capable of other embodiments or of being practiced or carried out in various ways. Also, it is to be understood that the phraseology and terminology employed herein is for the purpose of description and should not be regarded as limiting.

All references cited herein are incorporated by reference in their entirety.
20 Citation of a reference does not constitute an admission that the reference is prior art. The present application contains subject matter relating to a PCT application number IL07/001432 filed on Nov 20, 2008, entitled 3D GEOMETRIC MODELING AND 3D VIDEO CONTENT CREATION. This patent application is assigned to MantisVision, Ltd.

25 Reference is now made to Fig 3A, which is a simplified flow chart illustrating a generalized embodiment of the invention. System 200 contains anatomy tracking system 18, object interaction processing engine 42, virtual environment data module 32, and scene presentation engine 44. Anatomy tracking system 18 is comprised of at least one 3D video camera 22 and a 2D/3D image processing unit 26. The 3D
30 camera(s) may be situated in relation to the display in a variety of different configurations. The camera(s) 22, which can preferably acquire depth information in motion, allow for the derivation of a depth map of the viewer positioned at some location typically in front of the display. This depth map is typically in the form of a

three dimensional point cloud. The point cloud may then be input to a face and eyes finding algorithm contained in the image processing unit. The viewer's eyes location determines, as described above, the mapping of the perceived coordinates ($[x_i]$, $[y_i]$, $[z_i]$) of the displayed virtual objects. Furthermore, as described in figures below, the viewer eyes location determines the field of view or perceived space of the virtual scene seen by the viewer. The perceived virtual objects may be either static or dynamic.

Anatomy tracking system 18 is comprised of at least one 3D camera 22 and 2D/3D image processing unit 26. It is understood that multiple 3D cameras may be present. The 3D cameras together with the image processing unit 26 provide 3D coordinates ($[X_i]$, $[Y_i]$, $[Z_i]$) of the viewer's anatomy and/or movement of the anatomy in the camera field of view, typically the hand, arm, and face. Again, this set of data is typically a depth map in the form of a three dimensional point cloud. Furthermore, the 3D camera preferably is a 3D video camera capable of 3D motion capture, and as such, a point cloud for each frame in a captured video sequence is preferably acquired. The series of point clouds may then be input to a human finding engine and skeleton registering unit contained in image processing unit 38. Such skeleton registering units or 3D classifiers are known in the art and are not discussed herein. Additionally, the anatomy tracking system provides location data of the user's eyes.

The object interaction processing engine 42 accepts the output from the anatomy location tracking system as well as virtual environment data module 32. This virtual environment data is typically stored on a storage device and represents three dimensional representations displayed to the user according to techniques known in the art relating to AS displays.

Reference is now made to Fig 3B, which is an illustration of the internal components of interaction processing engine 42. The anatomical part-virtual object registration engine 34 accepts and processes information from both the anatomy tracking system 18 and virtual environment data module 32. Information concerning the eyes location from the anatomy tracking system 18 together with the virtual environment data from module 32 enables the registration engine 34 to determine the perceived virtual object 3D coordinates for the viewer's eye position ($[x_i]$, $[y_i]$, $[z_i]$). Additionally, the perceived AS display field of view as seen by the viewer according

to his eyes location is determined (see Figs 4A-4B below). The information from the anatomy tracking system 18 pertaining to movement of the viewer's anatomical parts such as his or her hand, arm, and/or head, enables a 3D mapping based on the acquired point cloud coordinates ($[X_i]$, $[Y_i]$, $[Z_i]$). These sets of data are registered to
5 determine a single 3D volume by the registration engine.

This registered data set from the registration engine 32 is then sent to the anatomical part-virtual object relation computation engine 34. In some embodiments, the relation computation engine determines if a collision occurs between the anatomical part and the virtual objects seen by the viewer. This would indicate that
10 the viewer has "touched" the perceived virtual object(s).

This information from relation computation engine 36 is then sent to the rule enforcement engine 38. This engine receives interactive application rules from 30 contained in the virtual environment data module 32. The interactive application rules module 30 contains all rules for determining subsequent AS display frames as a
15 function of user input. The rule enforcement engine then determines what the next presented scene on the AS display should be. For instance, through registration between the data sets (registration engine 34) a collision may be detected (relation computation engine 36), in which case the next frame displayed on the AS display may show the perceived virtual object, as a result of the "virtual touch", moved to a
20 different location (rule enforcement engine 38).

Alternatively, rule enforcement engine 38 may decide that the virtual object should disappear as a result of the "virtual touch". In still further embodiments, other resultant actions may occur to the virtual object as a result of interactivity between the viewer and the virtual object. The rule enforcement engine output is thus the output
25 of the interaction processing engine 42. This output is sent to scene presentation engine 44 to render the next multiple frame content to be displayed on the AS display.

Data storage of the virtual environment data may be implemented in any combination of volatile (for example, RAM, registers, etc) and/or non-volatile memory (for example, flash memory, magnetic medium such as disk, etc). Data
30 storage may be provided in a single location and/or may be distributed among multiple locations, for example, in a local and/or wide-area computer network.

In other embodiments, the viewer's hand acts as a "gravitational" source on the virtual object. In such embodiments discussed below, various gestures made by the

viewer and captured by the 3D camera or 3D camera array signify various applications on the virtual object. In such a case where the viewer's anatomy is not in a 3D volume in which the virtual mapping is displayed, registration between the 3D data sets need not be carried out, since the relative locations between the virtual object and viewer anatomy data sets is not taken into account. Rather, the object interaction processing unit, based on the detected 3D movement of the viewer's anatomy, determines the effect on the currently displayed scene. In still further embodiments, the interaction may combine both "touch" and "gesture" applications. In such cases, the interaction processing engine only contains the relation computation engine 36 and the rules enforcement engine 38.

The above process thus allows the viewer to touch, manipulate, and move virtual objects with his hand, watch virtual 3D objects fly into the palm of his hand, or control virtual scenes through gestures of his hand that are captured by the 3D camera. The 2D/3D camera system is preferably composed of a structured light active triangulation based imaging apparatus, such as detailed in Appendix 1. Such imaging devices utilize triangulation methods to determine three dimensional coordinates of imaged real world objects. However, there is no limitation on the type of 3D camera, and any type of 3D camera is appropriate, including but not limited to so-called 'time of flight' cameras, 2D cameras utilizing stereo correspondence algorithms, and triangulation based 3D cameras. As noted, the present embodiments preferably utilize 3D video cameras in order to capture 3D motion. The 3D imaging device(s) may be situated in relation to the display in a variety of different configurations, capture depth information in motion, and additionally capture 2D texture data of the imaged scene.

To better understand the present embodiments, a brief description of the 3D imaging system of Appendix 1 is now given. The 3D video camera contains both sensing and projector elements, preferably in a single housing. A pattern containing distinct features is projected onto imaged elements in the camera field of view. The pattern features, once reflected from the imaged elements, are captured in an image on the camera sensor. The location of the features in the image is correlated, through triangulation techniques, to 3D spatial coordinates on the imaged element. The totality of the 3D spatial coordinates comprises a point cloud of 3D points, which give a geometric shape of the imaged elements. The point cloud may further be processed to arrive at a 3D polygonal mesh.

The above process is carried out per frame for every frame of the video sequence, so that 3D modeling over time of imaged objects in both a static and dynamic scene may be implemented. If several cameras are used, then for each frame, triangulation may be carried out separately for each camera and additionally through stereoscopic comparison of the two images and triangulation of common points in both images. In such a case of multiple cameras, multiple point clouds are obtained per frame. These multiple point clouds are then registered to a unified point cloud set.

To further illustrate 3D touch interaction of the present invention, we now turn to Figs 4A and 4B. In the example shown in Fig 4A, the system may include an autostereoscopic (AS) display 14 and an array of 3D cameras 22 optionally situated on the top side of the display. Other configurations of the 3D camera(s) are understood to be possible, and the particular configuration shown is merely for illustration purposes. Shaded region 52 shows the perceived field of view in which the viewer 56 sees virtual objects given the location of his eyes in front of the AS display. That is to say, the viewer only perceives virtual objects and scenes that are within the shaded triangular region 52. This region is defined by light rays from adjacent pixel columns of the AS FLAT PANEL DISPLAY display that are perceived by each eye. The viewer's eye coordinates are denoted $(X^a), (Y^a), (Z^a)$ and $(X^b), (Y^b), (Z^b)$ for the left and right eyes respectively. Fig 15 below further illustrates the perceived space in which virtual objects are seen by a viewer of an AS display. The viewer eyes location leads to a mapping of the virtual object 60 coordinates in 3D space $([x_{i3}], [y_{i3}], [z_{i3}])$ seen by the viewer at this position.

In parallel to the eye tracking, the 3D motion capture video cameras capture data from which the 3D coordinates $([X_m], [Y_m], [Z_m])$ of the viewer's hand and arm 62 in the camera field of view may be derived. Upon processing of this initial point cloud data for each frame, a moving skeleton model is obtained of the viewer's anatomy part. A matching or collision between coordinate sets $([x_{i3}], [y_{i3}], [z_{i3}])$ and $([X_m], [Y_m], [Z_m])$ in a given video frame indicates that the viewer is "touching and/or manipulating" and therefore interacting with the virtual object.

With reference to Fig 4B, the viewer is seen moving laterally to the left in relation to the AS display. As a result, the viewer's perceived field of view of virtual scenes 64 on the AS display is shifted in relation to Fig 4A. Furthermore, the virtual object 60 now has a different mapping of coordinates $([x_{i4}], [y_{i4}], [z_{i4}])$ than that of

Fig 4A, that correspond to the viewer's new location in relation to the display. To interact with the virtual object in 3D space, the viewer now moves his hand and arm to a different location in 3D space within the perceived field of view of virtual objects. The skeleton mapping in this case is now based on a different captured point cloud series, namely $([X_n], [Y_n], [Z_n])$. Again, when a collision occurs between the data representing the 3D virtual object as perceived by the viewer and the data representing the human anatomy movement, some form of interaction occurs.

Reference is made to Fig 4C, which is a simplified illustration of the processing engines in preferred embodiments. 3D cameras 22 send to the anatomy tracking system 18 depth data from which to derive viewer eyes location data. In addition, the depth data is used to derive the viewer anatomy location and movement. The anatomy tracking unit sends data to the object interaction processing engine 42. Engine 42 additionally receives information on the interactive application from virtual environment module 32. The virtual coordinate mapping for virtual object 68 is determined based on the viewer eye location data. The object interaction processing engine registers the sets of virtual coordinate data and the anatomy location data. The next frame in the AS video sequence is based on the interactive rules processed in the interaction processing engine. The virtual environment data module 32 preferably contains the displayed content for current frames and the interactive characteristics and functionality of the displayed program. For instance, the interactive application determines the types of hand movements the viewer may make when he touches the object and whether the displayed object moves, twists, rotates, disappears, the manner in which the object performs these actions, and any other interactive functions.

For example, in a first displayed frame, viewer 56 may grasp virtual coke bottle 68 having initial virtual coordinate mapping $([x_i], [y_i], [z_i])$ at t_1 and rotate his hand and arm 62 to simulate a pouring action. The 3D video cameras 48 capture depth information of the moving hand and arm and send this data to the anatomy tracking system 18. The system 18 then computes the point clouds and subsequent skeleton model of the hand and arm for each frame. For instance, in the figure, the hand and arm have point cloud coordinates $([X_p], [Y_p], [Z_p])$ at t_1 . This engine 18 outputs to the object interaction processing engine 42 a skeleton model showing a hand in a pouring motion. The interaction engine, after processing this information together with the interactive application rules, tells the scene rendering unit 44 to display the next

virtual coke bottle in a tilted position. That is to say, the virtual coke bottle 68 in the next frame at t_2 has a new mapping of perceived coordinates ($[x_j], [y_j], [z_j]$) as a result of the interaction. Box 86 illustrates three different 3D coordinate mappings for three frames in which the virtual coke bottle 68 is displayed. In each successive frame, the perceived coordinates of the virtual coke bottle change in 3D space as the bottle appears to slowly tilt as if it is being poured by the viewer. Box 72 illustrates three different 3D coordinate mappings of the user's hand and arm for three successive time intervals, t_1 - t_3 . It is understood that certain areas of the virtual coke bottle are occluded from the viewer's view when he grasps the bottle, since in essence he is occluding his vision of the AS display at those points in space. Furthermore, the configuration and numbers of processing units in the diagram is for illustration purposes only and is in no way limiting.

Reference is made to Fig 5, which illustrates several visual examples of interaction between a real object, such as the viewer's hand, and a virtual 3D object displayed on an AS screen. Floating virtual beach ball 86 is displayed in the 3D space perceived by the viewer in front of the AS display 90. The viewer may poke the virtual ball with his hand 94. This poking action is detected by the camera(s) and processing devices of the anatomy location and configuration tracking system. A 3D mapping of the viewer's hand is represented by mapping ($[X_j], [Y_j], [Z_j]$). The object interaction processing engine then determines the resultant scene to be displayed based on the virtual environment data and interaction rules as a result of the poking action. In order to simulate a physical beach ball being poked, the virtual beach ball is displayed in a new 3D location in space in the following frame displayed to the viewer. That is to say, frame 1 has virtual object coordinate mapping ($[x_m], [y_m], [z_m]$) and frame 2 has mapping ($[x_n], [y_n], [z_n]$), closer to the AS display. Therefore, the displayed frame 2 on the AS display is a result of interaction between the viewer's hand and the virtual beach ball. It is understood that other actions, such as swinging at the virtual ball and reactions to the above described interaction are possible. The interactive application rule may dictate that the ball disappear, pop out of sight, fade out of sight, or any other desired interactive functionality as a result of the touch interaction or collision between the two sets of 3D coordinates.

Another interactive example shown in Fig 5 is virtual 3D bird 98 perching in the palm of viewer hand 102. As stated above, the viewer's hand is typically modeled

according to a skeleton modeling algorithm utilizing the point cloud coordinates derived from information captured by the 3D camera. Such skeleton modeling algorithms are seen for example in "Markerless Human Motion Transfer" by Cheung et al., "Image Based Markerless 3D Human Motion Capture using Multiple Cues" by Azad, et al., "Markerless Motion Capture from Single or Multi Camera Video Sequence" by Remondino et al., and others. The interactive processing engine unit, based on the interactive application rules and virtual environment data content, recognizes the skeleton model of an outstretched viewer's palm and sends the scene rendering unit 44 the next set of frames so that the virtual bird flies to and then rests on the viewer's hand. The mapping of coordinates of the virtual bird for each frame approaches a collision with the coordinates for the skeleton model of the outstretched hand. Alternatively, as shown, the viewer may attempt to catch a flying virtual bird that appears in the virtual space that he perceives according to his position in relation to the screen. Other examples in the figure show the viewer using his hands to interact with a virtual money bill 108 and a virtual leaf 110. In all of the shown examples, the collision of a set of 3D coordinates representing the human anatomy part with a set of 3D coordinates representing a displayed virtual object indicate touch and/or manipulation of virtual objects. The interactive application rules together with the input of these two sets of coordinates determine the resultant next frame to be displayed. In still a further application, the virtual object may be a direct input device whose interaction rules are derived from existing I/O devices such as a virtual 2D touch screen, multi-touch screen, keyboard, or other known input devices. Collision detection methods for interactive environments are known in the art and thus are not discussed herein. The reader is referred to "Collision Detection In 3D Interactive Environments" by van den Bergen as an exemplary way of performing 3D collision detection.

Fig 6 is a preferred embodiment showing a three dimensional autostereoscopic menu driven interface. The floating balls or nodes 112 appear to the viewer in the 3D space perceived by him in which he may view virtual 3D objects. The floating balls comprise a hierarchical menu. The first stage shows 4 nodes that represent the operating system desired by the viewer. When the viewer "touches" the virtual ball, as described in the above embodiments, a second tier menu appears to him. The second tier represents the desired application software and the third tier the functionality

therein. In the present example the viewer chooses Windows® and Outlook® and then Inbox. It is understood that any hierarchical menu driven interface may be manipulated in such an interactive manner and the details in the figure are merely one example for illustration purposes. In a certain embodiment, the chosen ball becomes larger, thus
5 indicating that it has been selected. This implementation may utilize historical learning algorithms such as seen in the link herein-
<http://www.inference.phy.cam.ac.uk/dasher/>.

Reference is made to Fig. 7 which shows a preferred embodiment wherein viewer classification rules enable the system to identify various characteristics of the
10 viewer and even identification and association with a stored historical profile. A viewer imaging device 250 images the viewer and sends the image to the viewer 3D processing engine 252. The 3D processing engine is controlled by the viewer classification rules module 254. The viewer classification rules module 254 preferably defines the characteristics of the viewer that are to be determined based on the
15 interactive application 256. For instance, the system may want to determine if a viewer is wearing a watch and then offer various products and services associated with various watch types. The system may want to determine the height of the person or his clothing size to present interactive content that is appropriate. Furthermore, an on-line viewer profile is optionally obtained through an on-line viewer database 258
20 containing the historical profile of viewers. The processing of the viewer may be iterative until an appropriate model is obtained, at which point the 3D modeling data is sent to the interaction engine 260. The interaction engine receives the interactive application information 256 as well. The interaction engine then decides upon the 3D displayed content on the display 262 based on the viewer classification and/or
25 identification and the interactive application. With such a user profile, buttons and other virtual objects such as seen in Fig. 7 can be adapted to display to the user a probable selection based on his historical profile. For instance, if he is known to pick Windows often, then in stage 1 he may see the 3D Windows icon enlarged for him for ease of use.

30 In another application, the 3D camera image and skeleton model enable the viewer to appear as a mirror image to himself. In other words, the displayed AS content is a mirror virtual image of the viewer himself. This image may utilize the viewers shape and/or motion in the form of an avatar or an actual image of the viewer.

Such an image can be used to display to the user how he or she would look with various hairstyles, clothing, etc. In a multi-user application, this avatar can be transmitted to from one user to the next for user interaction between them.

Reference is made to Fig 8. Fig 8 is an illustration parallel to Fig. 3 above showing the additional personal profile engine. The personal profile contains 3D camera 272, processing engine unit based on rules 274 and on-line database 276. It is understood that the 3D cameras of the anatomy tracking system may be used for the personal profile engine. The personal profile engine is based on the viewer classification rules that typically call for an iterative analysis process to determine desired characteristics of the viewer in front of the screen. In order to carry out this process, the personal profile engine typically takes input from the interactive application. The application may call for the identification and presentation of content to children that is different than content to adults. The interactive engine 42 receives information from the interactive application rules module as well. The interactive engine 42 then processes all four inputs to arrive at 3D content to be displayed.

Reference is made to Fig 9, which is a preferred embodiment of the present invention illustrating gesture-based interaction with 3D objects on an AS display. In such an embodiment, the viewer gestures typically with his arm and hand to manipulate a 3D virtual object perceived by him in the region 114 in which he may view virtual objects. Various motions of the hand and arm signify interactive functions with displayed 3D virtual objects.

User hand and arm 116 is shown making a pulling gesture to indicate the viewer's desire to pull the virtual beach ball 118 towards him. The viewer's hand and arm have real object mapping ($[X_k], [Y_k], [Z_k]$) in the form of a three dimensional point cloud for each frame. This point cloud is processed by image processing units as described above and the skeleton model of the moving anatomy is determined over time. The virtual beach ball has virtual object mapping ($[x_{i5}], [y_{i5}], [z_{i5}]$) for frame 1 at time 1. The interaction engine identifies the pulling motion from the skeleton model. Moreover, the interaction engine rules may determine that the pulling motion of the hand cause the virtual object to move in the next displayed frame closer to the 3D location of the viewer's hand. As the arrows in the figure indicate, the ball is then displayed in subsequent frames in incrementally closer 3D coordinates to that of the viewer. As seen, frame 2 at time t_2 virtual object coordinates ($[x_{i6}], [y_{i6}], [z_{i6}]$) are

closer to the viewer's hand location in 3D space given by real object mapping ($[X_k]$, $[Y_k]$, $[Z_k]$).

In certain gesture-based embodiments the precise eye location of the viewer may not be necessary for determination of the precise virtual object mapping scene by the viewer, since no actual collision algorithms are used to determine virtual touch. However, a determination of the viewer perceived three dimensional space in relation to the AS display may still be desirable.

Fig 10 shows a variety of other gesture-based interactive functions. All capital coordinate sets denoted $([X], [Y], [Z])$ indicate real objects, while all lower case coordinate sets denoted $([x], [y], [z])$ indicate virtual objects. The viewer may point his finger at the ball, shown as action 120, push the ball, shown as action 122, or pull the ball, shown as action 124. In all of the above cases, the next displayed frame is a function of the interaction application. The interactive application may be a "magnetic" effect, whereby the closer one is to the virtual object the stronger and faster the action. As a result of a specific gesture, the ball may pop, disappear, be dragged to and away from the viewer, among other functions. The viewer may also hold virtual objects in his hand such as virtual wrench 126 to manipulate other virtual objects in the viewer perceived space. Here virtual wrench 126 can be twisted and turned to tighten or loosen virtual screw 128. It is understood that under certain hand positions, the viewer's hand may occlude the virtual tool from appearing to the viewer, and the use of virtual tools may need to be modified to allow viewer visibility. Menu driven interface shown in Fig 6 may additionally be operated in gesture mode wherein the viewer merely gestures to choose the appropriate virtual ball.

Other gesture interactive motions 130 such as thumbs up, thumbs down, finger and hand movements are shown. Gestures may be carried out by a viewer to manipulate objects or send interactive commands without looking at his hands so long as they are within the camera(s) field of view. For instance, a tap of the viewer's left hand downwards may signify a lowering of the volume on the AS display, while upwards may cause an increase in display volume. Such an interaction is similar to touch typing on a computer keyboard, where the viewer is interacting with the computer while not moving his eyes from the screen. Gesture based interaction in 3D environments are discussed in the paper "Multi-Finger Gestural Interaction with 3D Volumetric Displays" by Grossman et.al., which relates to multi finger gestural

interaction with tracking devices for interaction with volumetric displays. It is understood that in contrast, the embodiments of the present invention make use of a 3D camera capable of markerless tracking to capture the 3D motion vector of the viewer's anatomy part at any given time.

5 Reference is made to Fig 11, which shows a preferred embodiment wherein a scaled version of the viewer or moving anatomy part of the viewer is displayed in the viewer perceived space as a "virtual viewer" for interface with other virtual objects. This scaled version 134 is typically miniature in size and sometimes referred to herein as the "virtual viewer" or "ghost". The 3D camera captures either the viewer 138 or an
10 anatomical part of the viewer and, upon image processing of the captured real coordinates ($[X_1]$, $[Y_1]$, $[Z_1]$), a scaled or miniature and preferably semi-transparent version of the viewer and his movements is displayed in the perceived space 142. The moving "ghost" has perceived virtual object coordinates ($[x_{i7}]$, $[y_{i7}]$, $[z_{i7}]$). Additionally, virtual object 146 with 3D perceived coordinates ($[x_{i8}]$, $[y_{i8}]$, $[z_{i8}]$) is
15 displayed in the viewer perceived space. Interaction may now occur between the "virtual viewer" and the virtual object. That is to say, all movements by the real viewer are reflected by parallel movements in the virtual viewer. Thus, instead of the real viewer interacting with the virtual object, the virtual viewer interacts in his place. This interaction may be either touch or gesture-based as discussed in previous
20 embodiments. Also, as in previous embodiments, the interaction engine preferably receives inputs from the anatomy tracking system and determines the interactive functionality. The interaction engine then sends commands to the scene rendering unit.

 Such a "ghost" may be utilized in PDAs or other portable devices with a small 3D display not large enough for accurate interaction between a small displayed virtual
25 object and the viewer's own hand. The miniature version of the viewer may be his entire body or just his hand, or even a three dimensional puppet or ghost-like image that does not resemble the viewer but moves in accordance with the viewer's body movements. The miniature can be in a non proportional scale and may move faster and slower in accordance with the interactive settings. That is to say, the sensitivity to
30 movement of the virtual miniature version of the viewer may depend on speed of the viewer's movement, similar to a computer mouse.

 The present embodiment can also be implemented as a seamless combination of both touch interaction, as described above, and the "ghost" interaction. When the

viewer's anatomy is within the space in which he is capable of viewing virtual objects, the viewer may virtually touch perceived virtual objects in the virtual scene. As the viewer moves his hand out of the 3D space in which he is able to perceive virtual objects, but still within the camera field of view, a seamless transition can occur where
5 the viewer sees the "ghost" interface at a certain distance from the display.

Reference is made to Fig 12, which illustrates a viewer or anatomical part of a viewer and the parallel "virtual viewer" or "virtual anatomical part" being displayed in the viewer perceived space. Viewer 150 sees a "ghost" 154 of himself in miniature form in the perceived virtual object space in front of the display 156. The "ghost"
10 appears to the viewer with a virtual "dashboard" 158 on his back that indicates to the viewer various states and modes of interaction, similar to CAPS LOCK on a standard computer keyboard. Viewer hand 160 is mapped into the perceived virtual object space and may take the form of various shaped virtual hands such as 162 and 164. The virtual hand may additionally manipulate scaled virtual objects such as hammer
15 166 and beach ball 168.

Reference is made to Fig 13A-13C, which illustrates the 3D camera and post processing operations as described in the above embodiments. For ease of understanding and illustration, only one 3D camera is shown in Fig 13A with field of view 170. It is understood that, as shown in the preferred embodiments above, in order
20 to enlarge the field of view multiple 3D cameras and differing configurations of 3D cameras in relation to the AS display may optionally be used. The camera is comprised of a two dimensional camera 172 with vertically attached projector 174. Additional projectors and various configurations of projectors in relation to the AS display may be implemented as well in alternate embodiments. The configuration of projectors for
25 the current embodiment is for illustration purposes only. It is understood, that in order to prevent interference between projectors, particularly when using structured light based 3D cameras, the projections may be separated, for example, by different wavelengths, polarities, or time multiplexing. The camera is typically positioned in a suitable position in front of the viewer and/or attached to the AS display 176.
30 Furthermore, various adjustments can be made to the camera and projector lenses, angles, and positions, aperture settings, as well as the number of cameras and/or the number of projectors to better suit the desired application. For example, in the above discussed figures, several 3D cameras are shown. In the present embodiment, it is

understood that the projector is placed vertically to the camera purely for ease of illustration.

As seen in Fig 13B, the camera captures the viewer's depth map 1773 in the form of a 3D point cloud. The depth map provides the point cloud to a face finding engine or 3D classifier 180. Such 3D classifiers are known in the art, such as described in "Automatic Classification of Objects in 3D Laser Range Scan" by Nuchter et.al. The face finding engine then is used to determine precise eyes location 182 and optionally the direction in space at which the viewer is focusing. This optional element is denoted as eye direction tracker 184. Eye direction tracking may be used to determine the position in space at which the viewer is looking at a given moment in time. The eye position detection is denoted by dashed lines 178 in Fig 13A. The eyes finding engine preferably operates on a continuous basis to determine the viewer perceived 3D space as the viewer, and thus his eyes location, moves position in front of the AS display.

Furthermore, the camera captures the viewer's body or parts thereof in three dimensional positional coordinates, and typically at least the viewer's hand and arm 186 and face 188. Fig 13C shows a typical flow diagram of skeleton modeling based on depth capture. The sampled depth data 190, in addition to optional texture data 192, is input to a viewer finding engine 194 and from there to one or more post processing units to determine his body contour from which a skeleton model is registered 196. As stated, the depth data may be combined with textural information captured simultaneously. Texture may be captured through use of an additional camera or CCD. Such methods for animatable human body reconstruction from data sampled in three dimensions are discussed for example in US Patent No. 6133921. Moreover, the 3D data provide dynamic video images, meaning through 3D motion capture methods the 3D positional data of the viewer's anatomy part is known at any given point in time. The 3D skeleton model may be captured from various types of 3D cameras, such as stereo correspondence, time of flight or triangulation based cameras. As discussed above, the 3D skeleton model of the viewer's anatomical part is used to determine interaction with the virtual scene.

Reference is made to Fig 14, which shows the basic characteristics and functionality of an exemplary automultiscopic/stereoscopic display as used in the previously described embodiments. AS displays may have differing characteristics,

such as screen width, number of views, and optimal viewing distance, but the display principles typically behind such displays is the same- the division of the overall resolution of the screen to a defined number of "viewing zones" through use of some type of optical filter.

5 Display 200 has lenticular lenses 204 placed over a flat panel screen 206 divided into a series of two views 210 and 214. Each pair of views on the flat panel display is covered by a separate half cylindrical shaped lenticular lens. Each view is made up of a pixel column on the flat panel display and each lenticular lens projects both views simultaneously to a viewer. The optical characteristics of the lenticular
10 lens are responsible for the fact that at varying positions in front of such a screen, each eye 216 of a viewer sees one of the two views. By viewing two different images simultaneously, the viewer experiences the three dimensional effect called stereo parallax.

Several methods have been developed, for instance through use of a slanted
15 lenticular lens, to spread the resolution reduction over both x and y components of the screen. Other methods of autostereoscopic display have been developed such as parallax barrier methods. Moreover, as discussed herein, for multiview systems, each zone may display to a viewer therein different content.

Reference is made to Fig 15, which is a simplified illustration of the perceived
20 3D space in relation to the AS display in which the viewer may observe and interact with virtual objects and scenes. The viewer 218 position in relation to the display 220 determines which pixel columns each eye views. The display is typically divided into a repeating series of pixel columns in accordance with the number of viewing zones. In the figure, the display is divided into a series of 4 columns corresponding to 4
25 possible viewing zones. Under each lenticular lens a group of 4 columns exists. Example columns 222 and 224 are shown, and these pixel columns display images seen in the number 2 and 3 viewing zones (not shown). For example, columns 222 project light rays 226 and 228 into the right and left eye respectively of viewer 218 whose eyes lie in zones 2 and 3 (not shown). The entire 3D space in which the viewer
30 may perceive 3D objects from that location with respect to the display is denoted by shaded region 230. Region 230 is the 3D space in front of the screen in which both eyes view light rays from pixels on the screen. It is understood that the light projections shown from the AS display are for illustration purposes. When a viewer

shifts his eye position within a given zone, the viewer typically experiences the visual artifact of a shift in the virtual object. This artifact is a result of the lack of motion parallax within the viewing zone.

Reference is made to Fig 16, which is a simplified illustration showing the case where the camera field of view does not cover all of the viewer perceived space in front of the AS display. In such a case, area 232 represents what may be termed an interaction zone. The interaction zone represents the overlap between the 3D camera field of view and the space within which the viewer interacts and sees virtual objects. For "virtual touch" applications, interactivity is determined by the registration between the 3D coordinates of the virtual scene (displayed object(s)) perceived by the viewer and the 3D positional data of the viewer's anatomical part. It is seen that should the viewer hand be outside the camera's field of view but still within the region in which the viewer sees virtual objects, region 234, then even if the viewer perceives the virtual touch the camera does not. As such, no interaction between the viewer and the virtual object is detected in the interaction engine and no interactive application is possible. The interaction zone is thus the totality of 3D points in space in which virtual objects are perceived by the viewer and that is also captured by the camera. It is understood that various layouts and numbers of cameras and projectors may provide for interaction zones of varying volumes.

Reference is made to Fig 17, which show an alternative embodiment in which the viewer's eyes position data obtained from the 3D camera may also be input to a screen steering device 236. In such an embodiment, the screen display's angle and location adjusts to the viewer's precise location through use of a moving optical filter that alters the optical viewing space. The screen then needs only to render the two appropriate views to the viewer's location, and the decrease in screen resolution due to the partition into repeating series of pixel columns for each view is obviated. Such a display device is described for instance in US Patent No. 6,075,557. Furthermore, the exact positioning of the individual viewing zones, seen in Fig 1, can be refined and tuned to the precise viewer location rather than being in a fixed cubic area in front of the AS display.

In other adaptive screen embodiments, a slight rotation and/or translation of the lenticular lens over the flat panel display allows for adjustment of the viewing zones. In AS displays using parallax barrier, the occlusion of light is determined based on the

user's position so as to maximize the zone locations. Various lens types are discussed in "A Survey of 3DTV Displays: Techniques and Technologies" by Benzie, P, et.al. Circuits and Systems for Video Technology, IEEE Transactions on Volume 17, Issue 11, Nov. 2007 Page(s):1647 – 1658. In still other screen types, acoustic optics may be used to control the refractive index. Such an acoustic optical covering may then be used to dynamically adjust the viewing zones. An acoustic lens of this type is discussed in the "Acoustic Lens with Electrically Controlled Refractive Index" by Pappalardo, M from 1980, Ultrasonics Symposium, Volume , Issue , 1980 Page(s): 590 – 593.

Reference is made to Fig 18 which shows a network of 3D virtual touch human interface systems as described herein. Each system is denoted as 200 and has the characteristics described in Fig 3A above. A network of such systems may allow for a multi-user interface wherein one or more users interact with each other based on content created at each viewer's individual station. The scene presentation engine outputs from each system and is sent to other stations over the network. For instance, in a handball game between two users, each user hits the ball in his 3D volume in front of his AS display. The resultant virtual environment information is then sent to the second user over the network. In other words, the second user sees the virtual ball in front of his AS display as a result of interaction of the first user with the ball. In such a way, a multi user game is possible. In the above described multi-user application, a unified 3D volume exists of all user that registers the individual local registered environments into a single registered environment of all users. Multiuser 3D video videoconference is another implementation using such a network of systems.

It is expected that during the life of this patent many relevant devices and systems will be developed and the scope of the terms herein is intended to include all such new technologies a priori.

It is appreciated that certain features of the invention, which are, for clarity, described in the context of separate embodiments, may also be provided in combination in a single embodiment. Conversely, various features of the invention, which are, for brevity, described in the context of a single embodiment, may also be provided separately or in any suitable subcombination.

Although the invention has been described in conjunction with specific embodiments thereof, it is evident that many alternatives, modifications and variations

will be apparent to those skilled in the art. Accordingly, it is intended to embrace all such alternatives, modifications and variations that fall within the spirit and broad scope of the appended claims. All publications, patents, and patent applications mentioned in this specification are herein incorporated in their entirety by reference
5 into the specification, to the same extent as if each individual publication, patent or patent application was specifically and individually indicated to be incorporated herein by reference. In addition, citation or identification of any reference in this application shall not be construed as an admission that such reference is available as prior art to the present invention.

CLAIMS

What is claimed is:

1. An apparatus for providing an interactive human computer interface to a
5 viewer, said apparatus comprising:
- a) a data storage configured to store three-dimensional virtual environment representation data including at least one three-dimensional virtual object within said virtual environment;
 - b) an autostereoscopic (AS) display configured to display simultaneous
10 perspectives of virtual objects of said 3D virtual environment representation data to spatially separated viewing zones located in front of said AS display
 - c) an anatomy tracking system including at least one 3D camera, said anatomy tracking system operative to determine respective real-world locations of:
 - i) a viewer's eyes; and
 - 15 ii) an anatomical part on said viewer's body
 - d) a 3D registration engine configured to generate a 3D volume comprising 3D location data of said viewer's anatomical part and 3D location data of said virtual objects perceived by said viewer in accordance with said viewer's eye location,
 - e) an anatomical part-virtual object relation computation engine operative to
20 determine a relation between said virtual object and said anatomical part in accordance with output of said registration engine:
 - f) a rule enforcement engine operative to modify said three-dimensional environment representation data in accordance with said determined anatomical part-virtual object relation and said virtual environment data.
- 25
2. A system according to claim 1, wherein said virtual environment data contains interactive application rules.

3. A system according to claim 1, wherein said anatomy tracking system determines said viewer's eye position and anatomical part position continuously over time.
- 5 4. A system according to claim 1, wherein said rule enforcement engine is further configured to provide output data in real time to a scene presentation engine for display of subsequent content.
- 10 5. A system according to claim 1, configured to receive input from a plurality of viewers having eye positions in a plurality of adjacent pairs of said viewing zones, said viewers simultaneously interacting with said autostereoscopic display.
- 15 6. A system according to claim 1, wherein said 3D camera comprises an NIR projector and narrowband filter.
7. A system according to claim 1, further comprising a viewer classification module configured to determine viewer identity and characteristics based on a viewer historical profile.
- 20 8. A system according to claim 7, wherein said viewer profile is obtained through an on-line database.
9. A system according to claim 1, wherein said displayed three-dimensional virtual environment representation data is a virtual touch screen.
- 25 10. A system according to claim 1, wherein said displayed three-dimensional virtual environment representation data is a virtual keyboard.
11. A system according to claim 1, wherein said displayed three-dimensional
30 virtual environment representation data is an virtual image of the user.

12. A system according to claim 1, wherein said displayed virtual environment representation data as a result of rule enforcement engine output is sent over a network in a multi-user environment to allow for multi-user interface.

5 13. A system according to claim 1, wherein said displayed virtual environment representation data is continuously changing 3D positional data.

14. A system according to claim 1, further comprising a miniaturized element perceived by said viewer in a viewer perceived space that is determined in accordance
10 with said viewer's eye position, said miniaturized element configured for interaction with said displayed 3D virtual objects.

15. An apparatus for providing an interactive human computer interface to a viewer, said apparatus comprising:

15 a) a data storage configured to store three-dimensional virtual environment representation data including at least one three-dimensional virtual object within said virtual environment;

b) an autostereoscopic (AS) display configured to display simultaneous perspectives of virtual objects of said 3D virtual environment representation data to
20 spatially separated viewing zones located in front of said AS display

c) an anatomy tracking system including at least one 3D camera, said anatomy tracking system operative to determine respective real-world locations of:

i) a viewer's eyes; and

ii) an anatomical part on said viewer's body

25 d) an anatomical part-virtual object relation computation engine operative to determine a relation between a perceived virtual object by said viewer and said anatomical part

e) a rule enforcement engine operative to modify said three-dimensional environment representation data in accordance with said determined anatomical part-
30 virtual object relation and said virtual environment data.

16. A system according to claim 15, further comprising a miniaturized element perceived by said viewer in a viewer perceived space that is determined in accordance

with said viewer's eye position, said miniaturized element configured for interaction with said displayed 3D virtual objects.

17. A system according to claim 16, wherein said miniaturized element is
5 perceived by said viewer when said viewer's interacting anatomy part is outside of said viewer perceived space, and wherein said miniaturized element disappears when said viewer's interacting anatomy part is within said viewer perceived space.

18. An autostereoscopic (AS) display configured to display simultaneous
10 perspectives of virtual objects of said 3D virtual environment representation data to spatially separated viewing zones located in front of said AS display, said AS display having an acoustic lens with electrically controlled refractive index to dynamically adapt said viewing zone locations.

19. A method for providing an interactive human computer interface to a viewer,
15 said method comprising:

a) storing three-dimensional virtual environment representation data including at least one three-dimensional virtual object within said virtual environment;

b) displaying on an autostereoscopic (AS) display simultaneous perspectives
20 of virtual objects in said 3D virtual environment representation data to spatially separated viewing zones located in front of said AS display

c) tracking the anatomy of a viewer to determine respective real-world locations of:

i) a viewer's eyes; and

25 ii) an anatomical part on said viewer's body

d) generating a registered 3D volume comprising 3D location data of said viewer's anatomical part and 3D location data of said virtual objects perceived by said viewer in accordance with said viewer's eye location,

e) determining a relation between said virtual object and said anatomical part
30 in accordance with said registration,

f) modifying based on interactive rules said three-dimensional environment representation data in accordance with said determined anatomical part-virtual object relation and said virtual environment data.

20. A system for interactive human computer interface, said system comprising:
a self-contained autostereoscopic (AS) display configured to render 3D virtual
objects into neighboring viewing zones associated with said display,

5 an eye location tracking system, comprising at least one 3D video camera, for
continuously determining:

1) a viewer perceived three dimensional space in relation to said display, and

2) a 3D mapping of said rendered virtual objects in said perceived space in
accordance with viewer eyes position in relation to said fixed viewing zones

10 an anatomy location and configuration system, comprising at least one 3D
video camera, for continuously determining a 3D mapping of viewer anatomy in
relation to said display, and

an interactive application that defines interactive rules and displayed content
to said user, and

15 an interaction processing engine configured to receive information from

4) said eye location tracking system

5) said anatomy location and configuration system, and

6) said interactive application

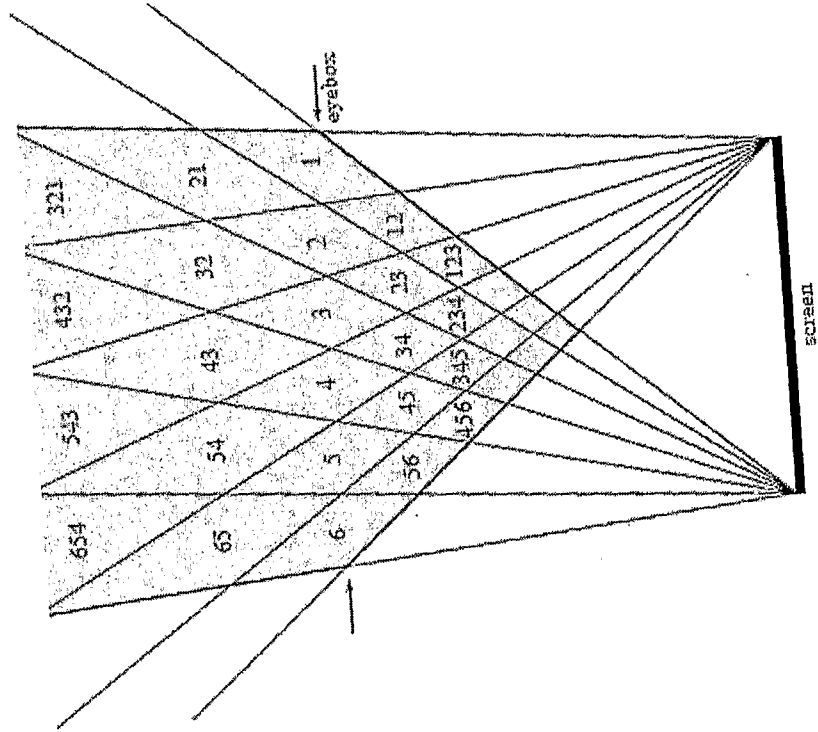
20 thereby to determine interaction data of said viewer anatomy with said rendered
virtual objects from AS display.

21. A system or method for three dimensional interaction with auto stereoscopic
displays substantially as described or illustrated hereinabove or in any of the
drawings.

25

30

Fig 1



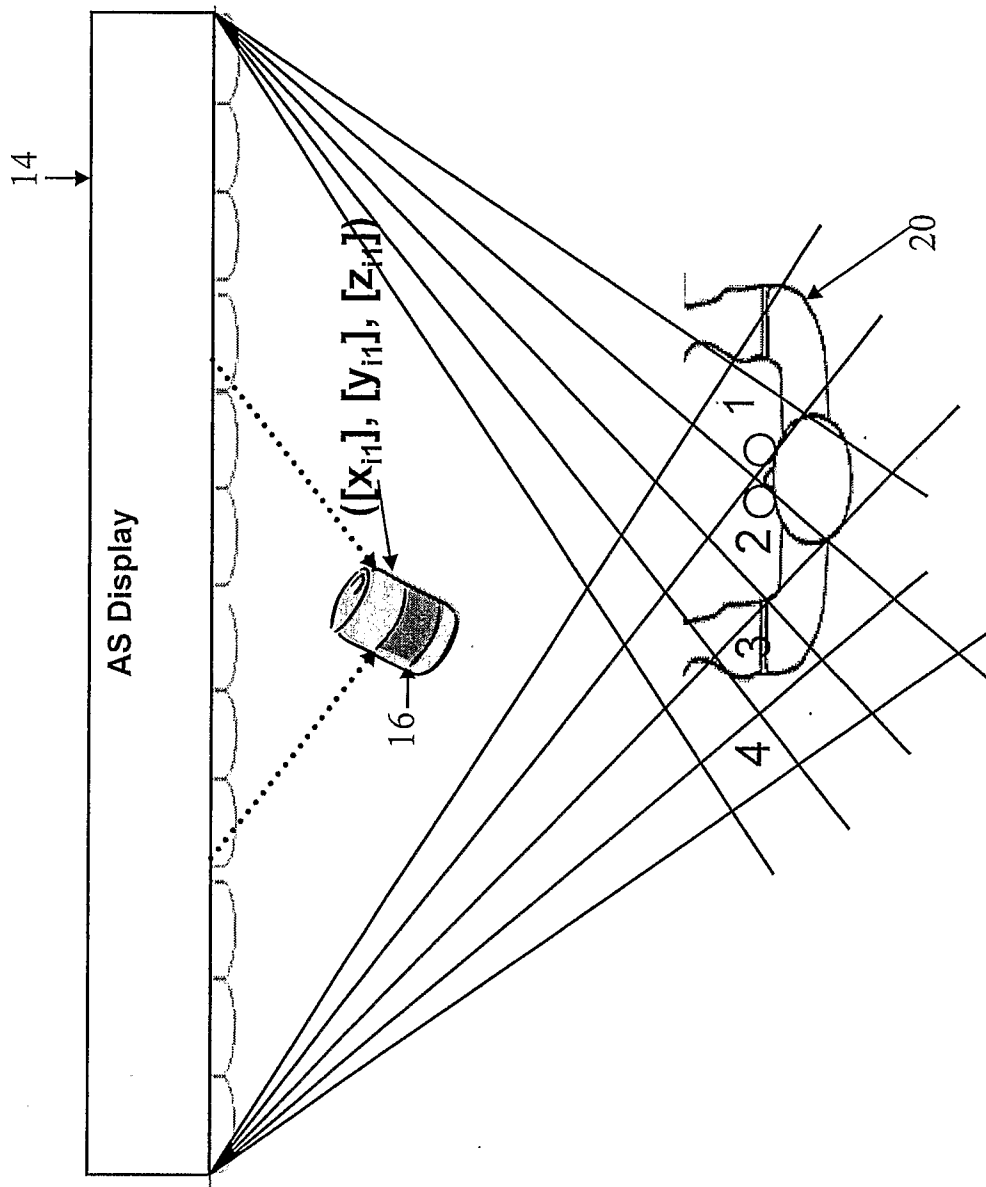


Fig 2A

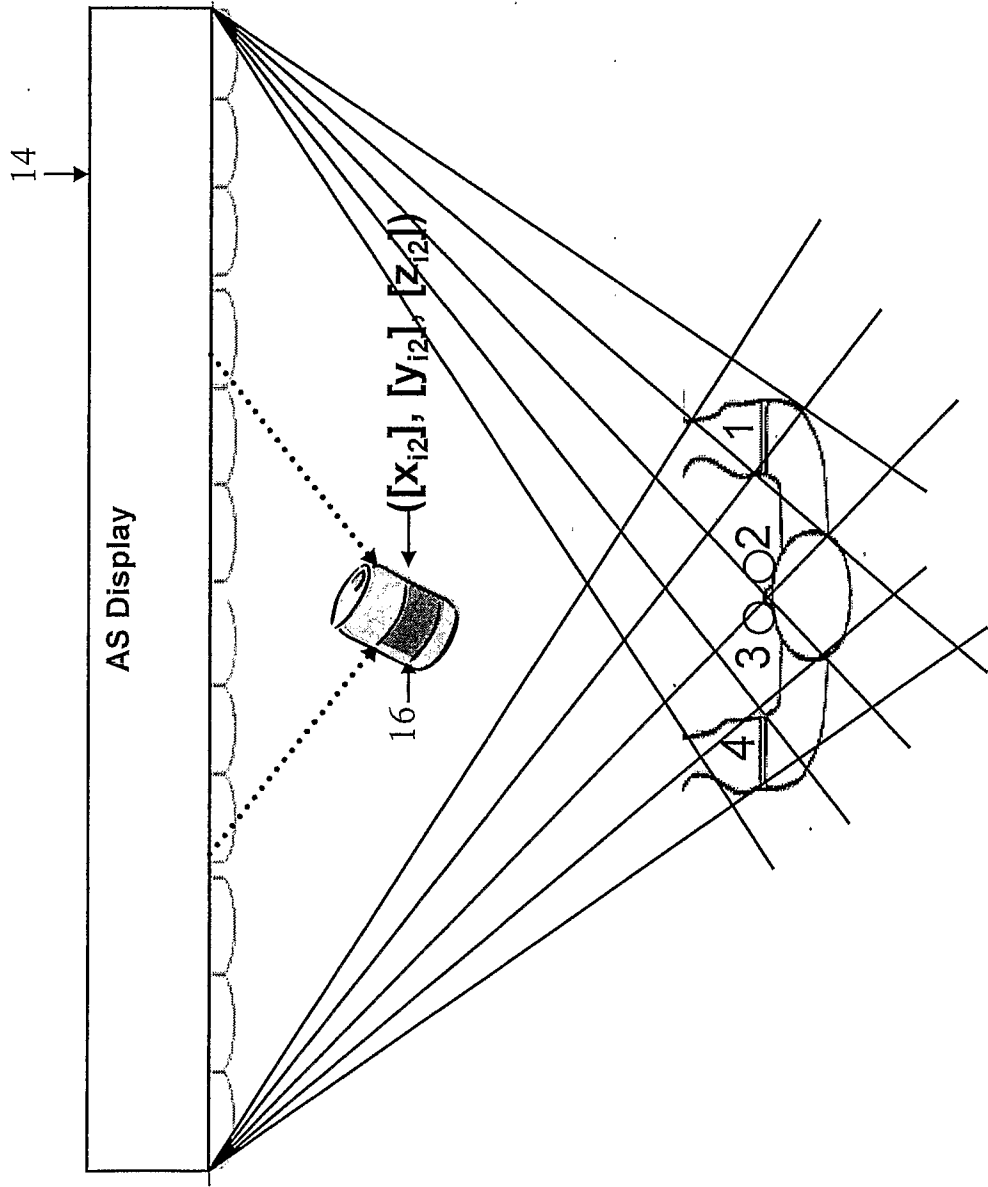


Fig 2B

200

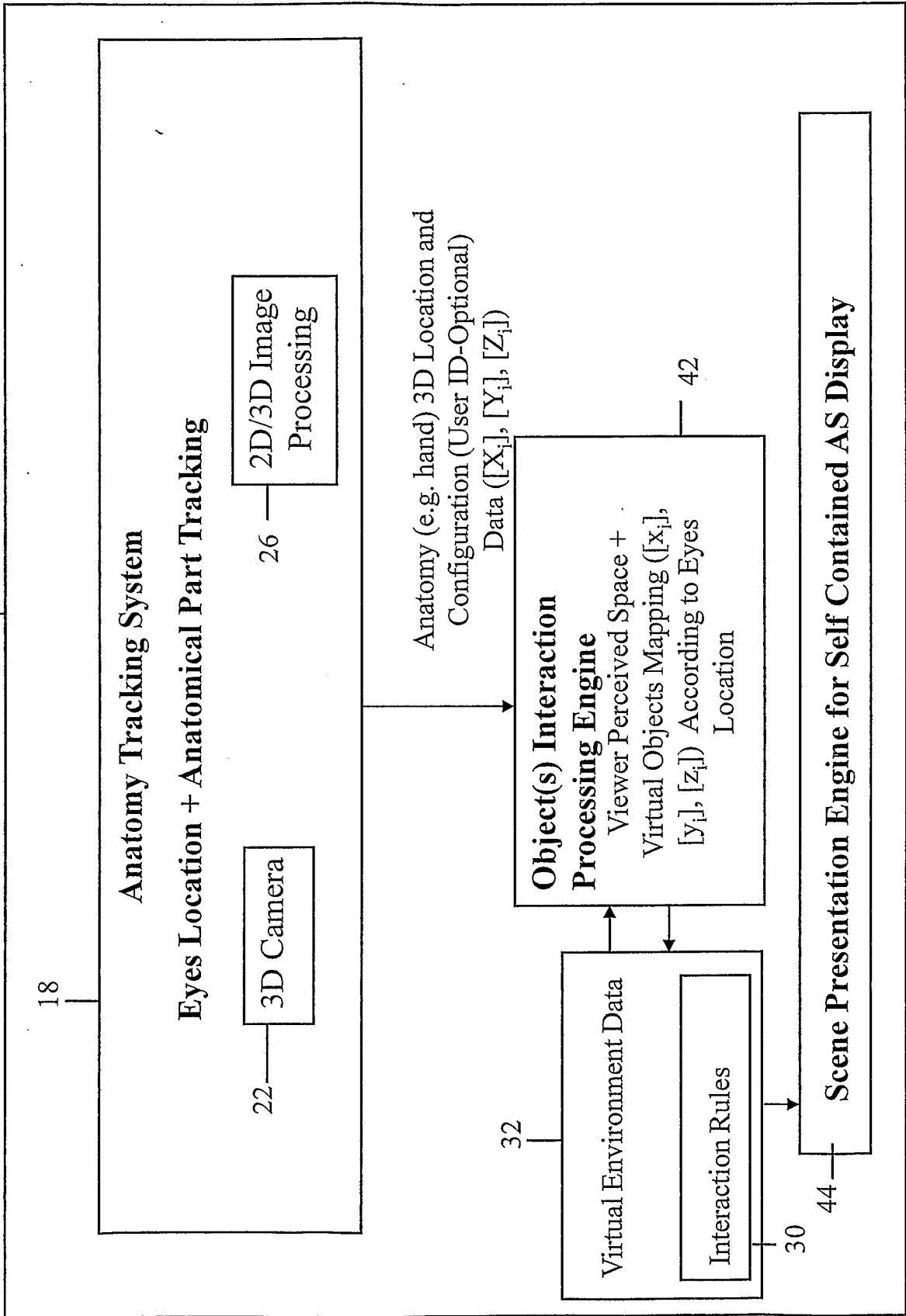


Fig 3A

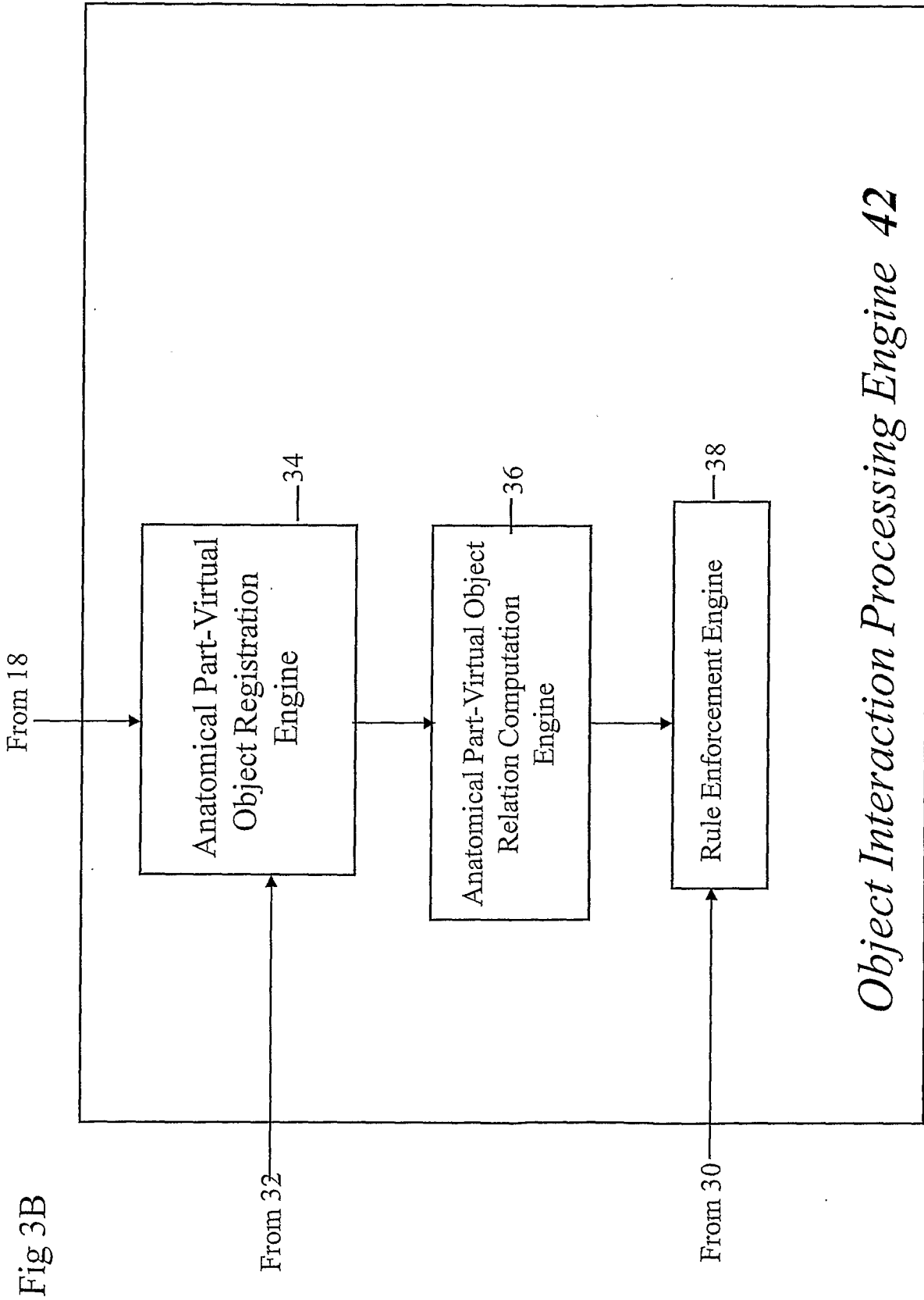


Fig 3B

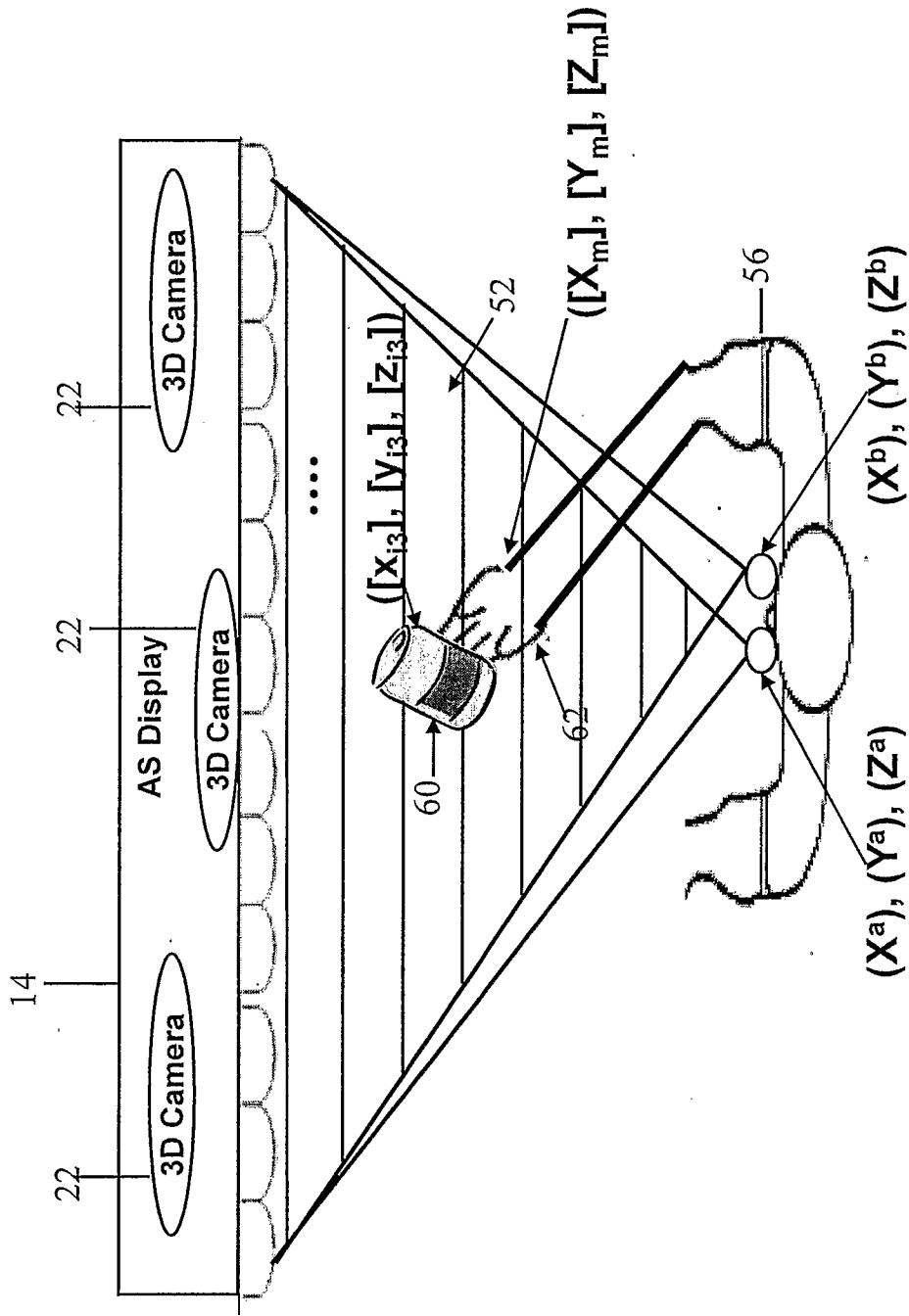


Fig 4A

Fig 4B

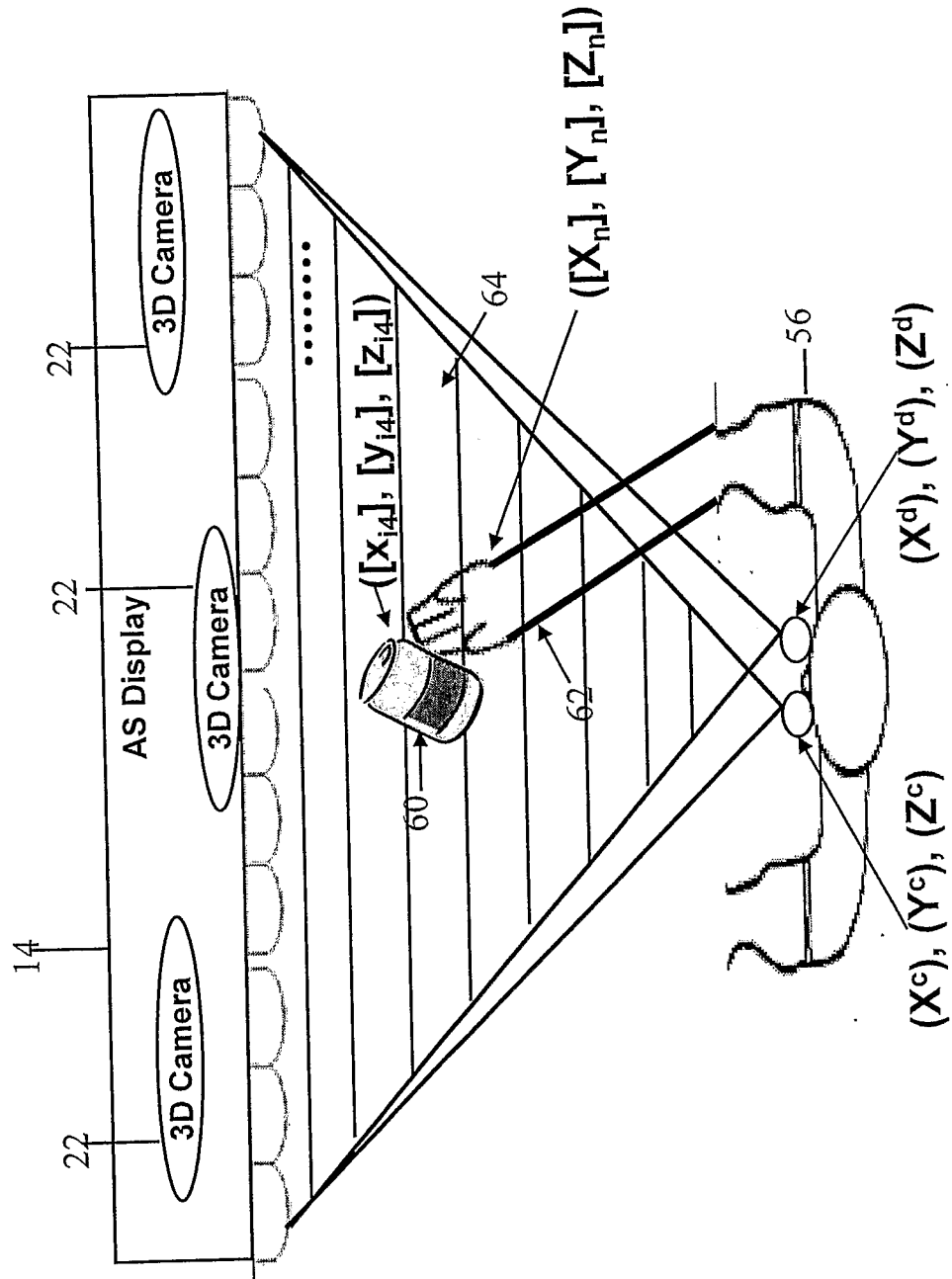


Fig 4C

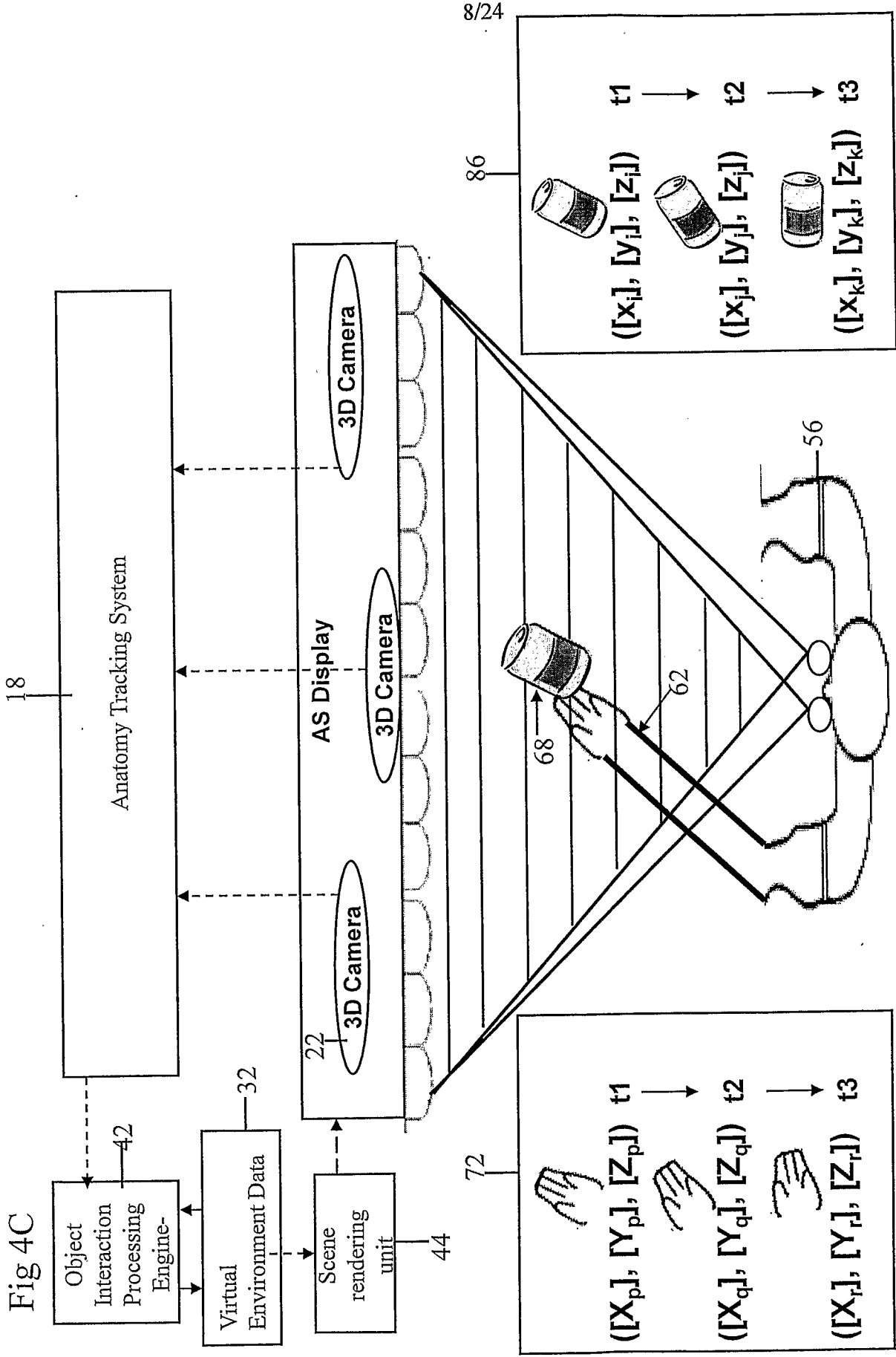
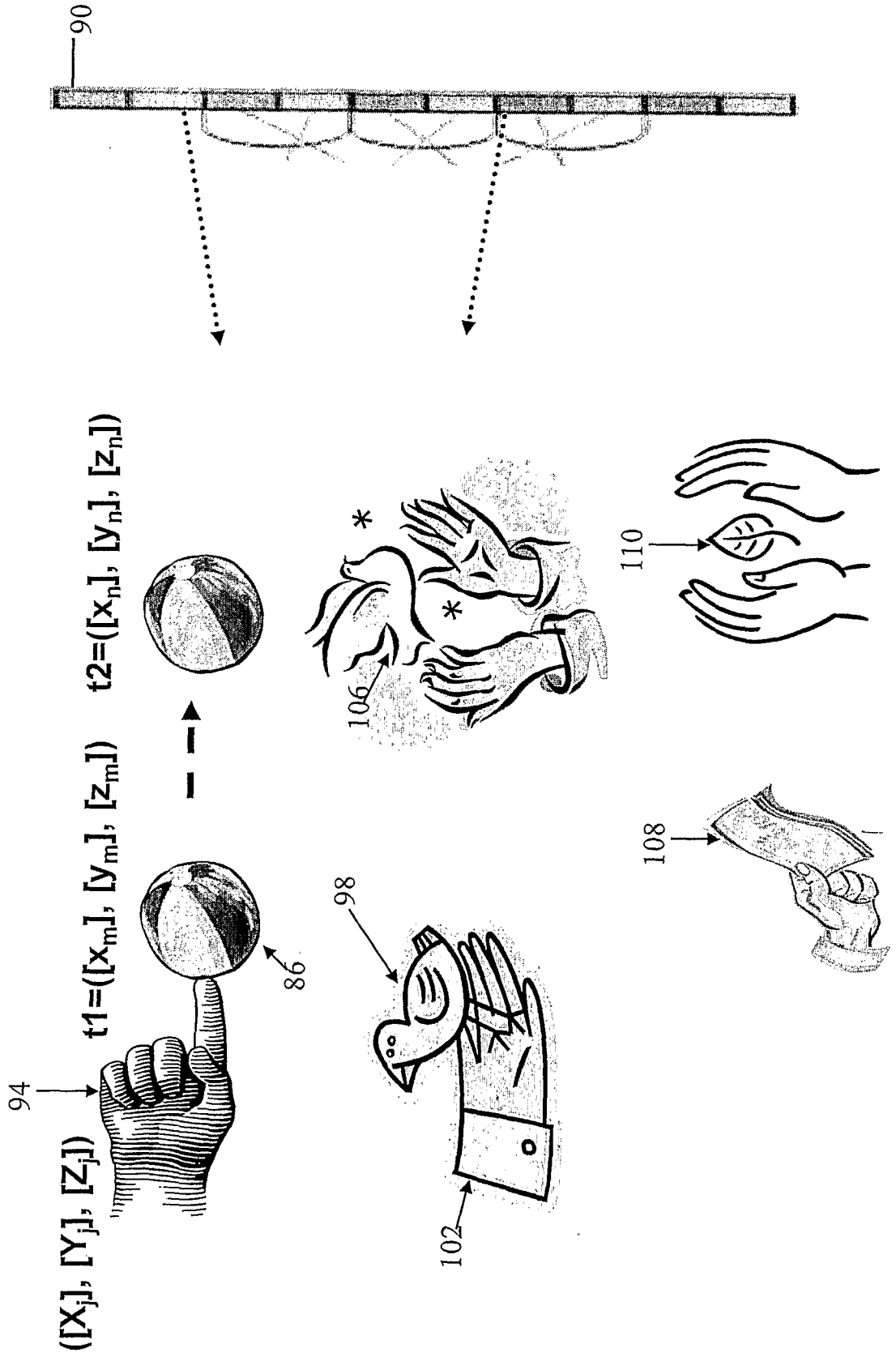


Fig 5



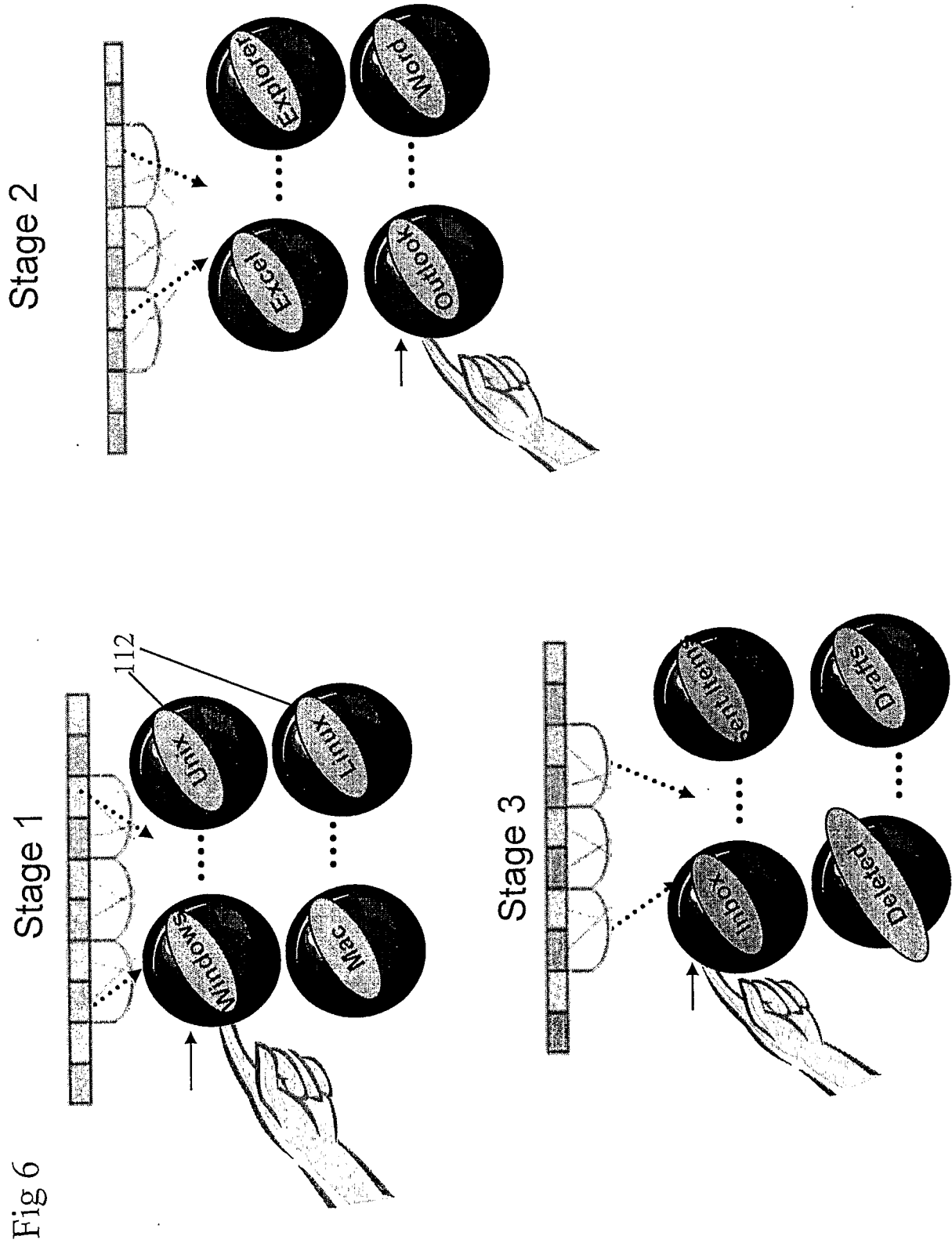


Fig 6

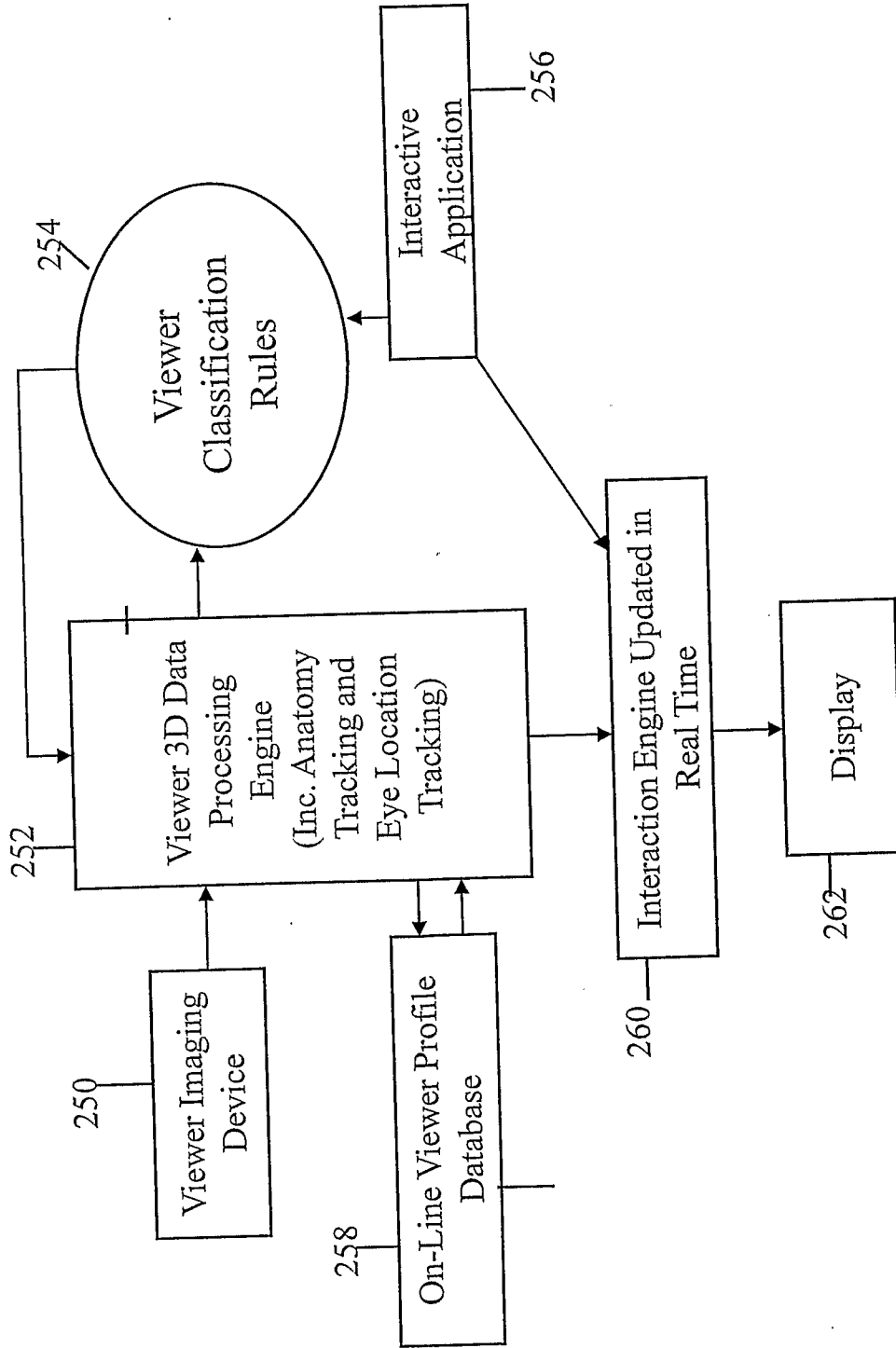


Fig 7

Fig 8

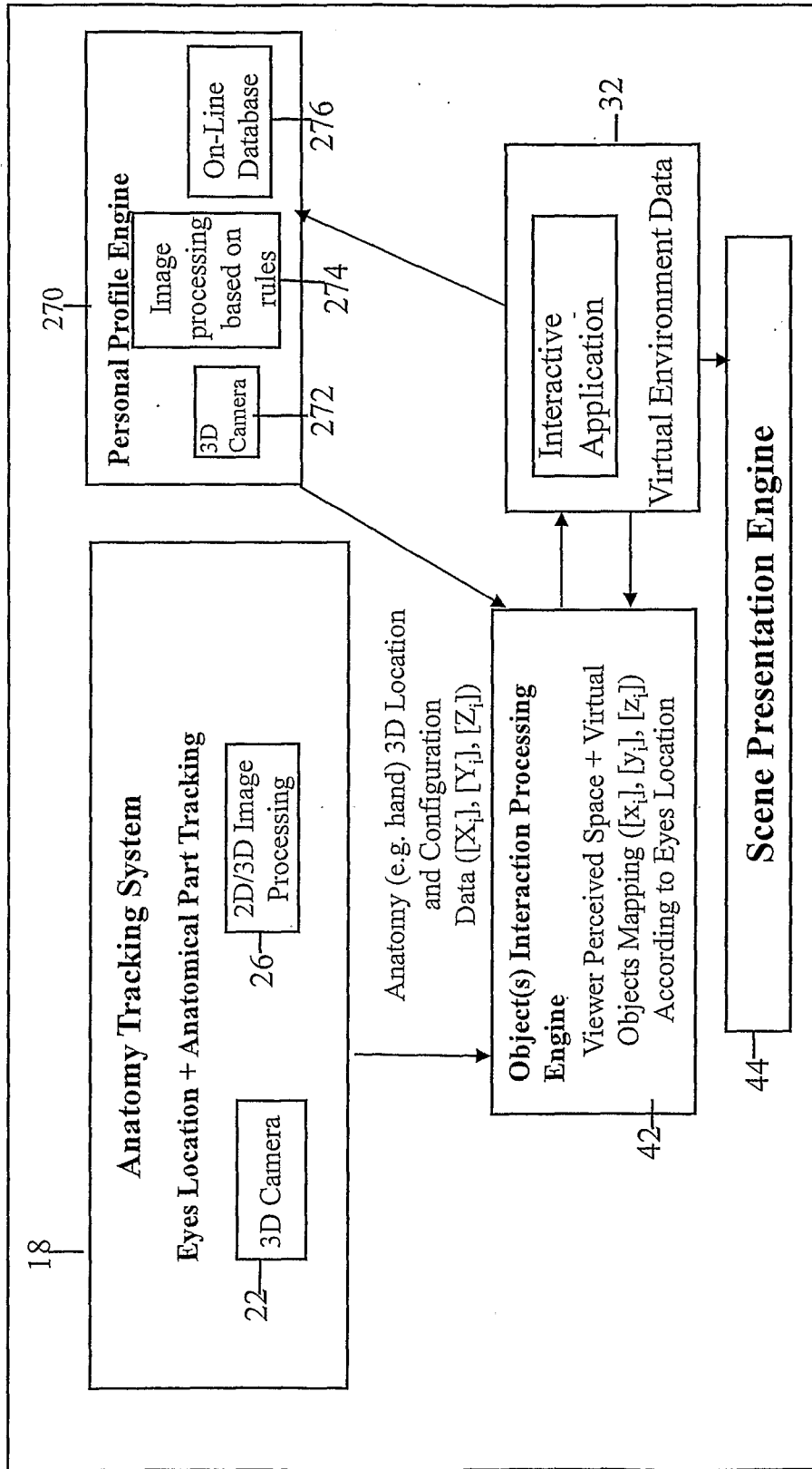


Fig 9

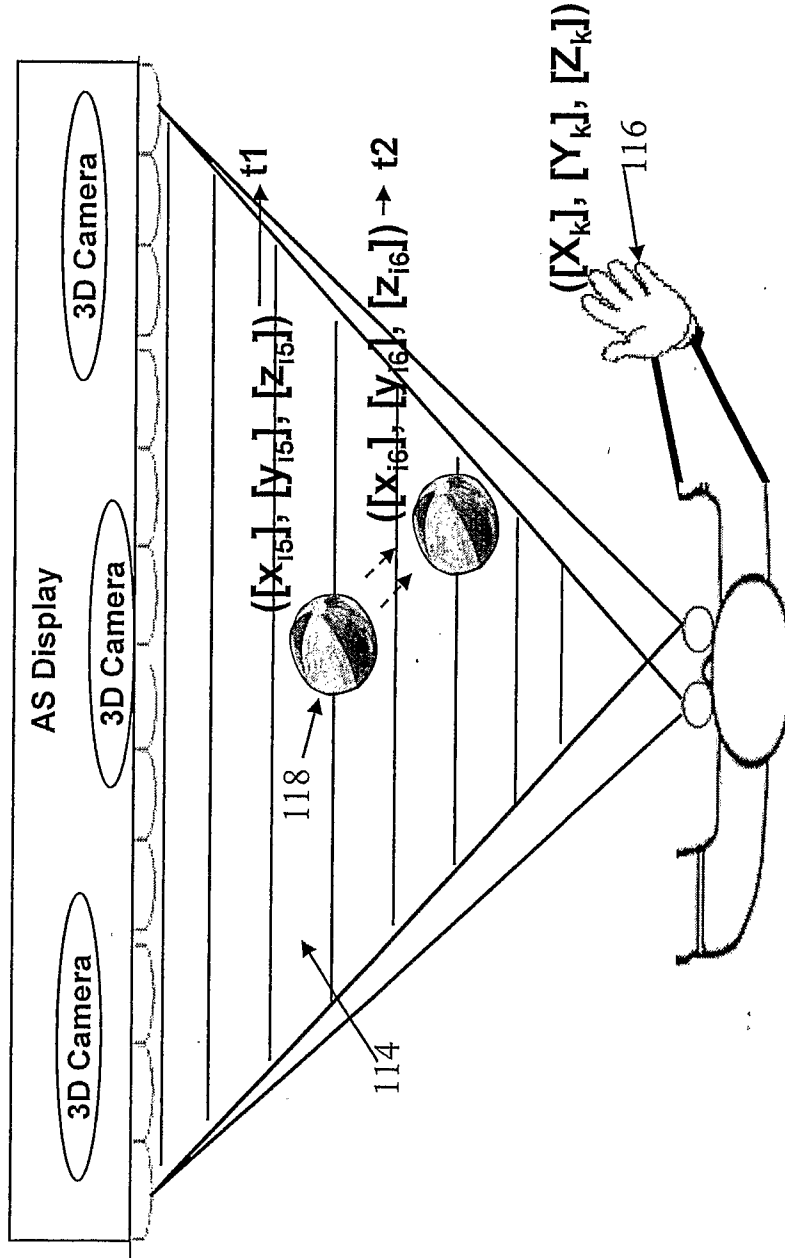
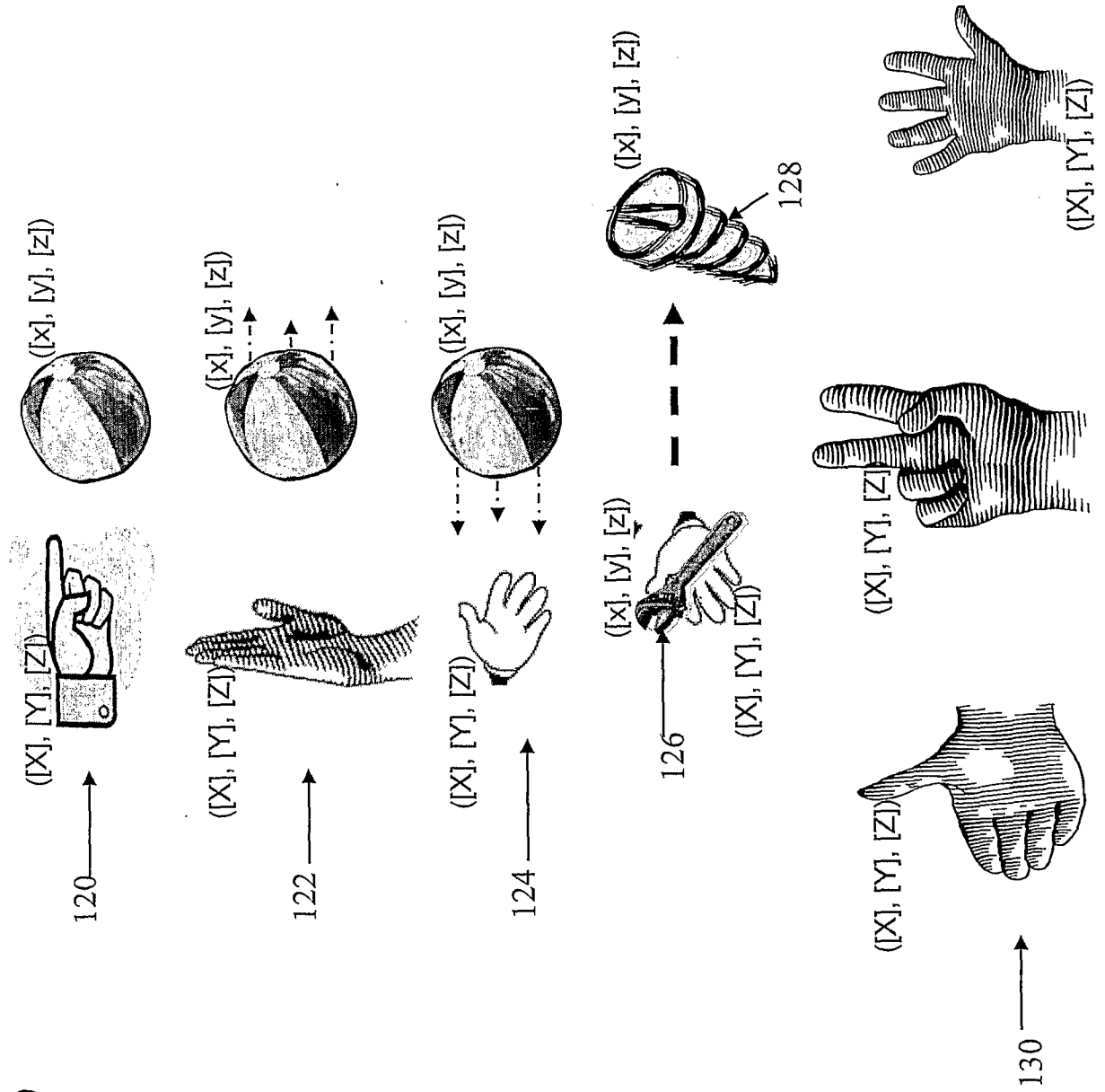


Fig 10



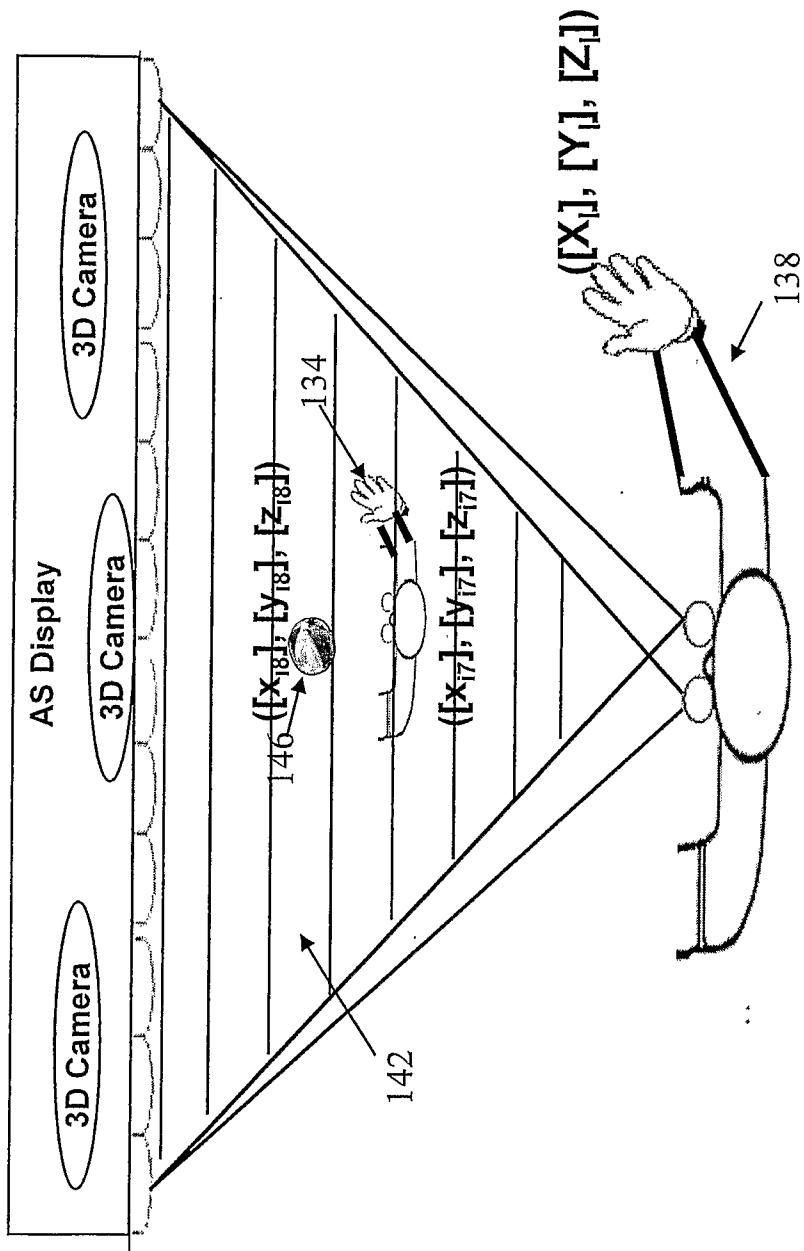


Fig 11

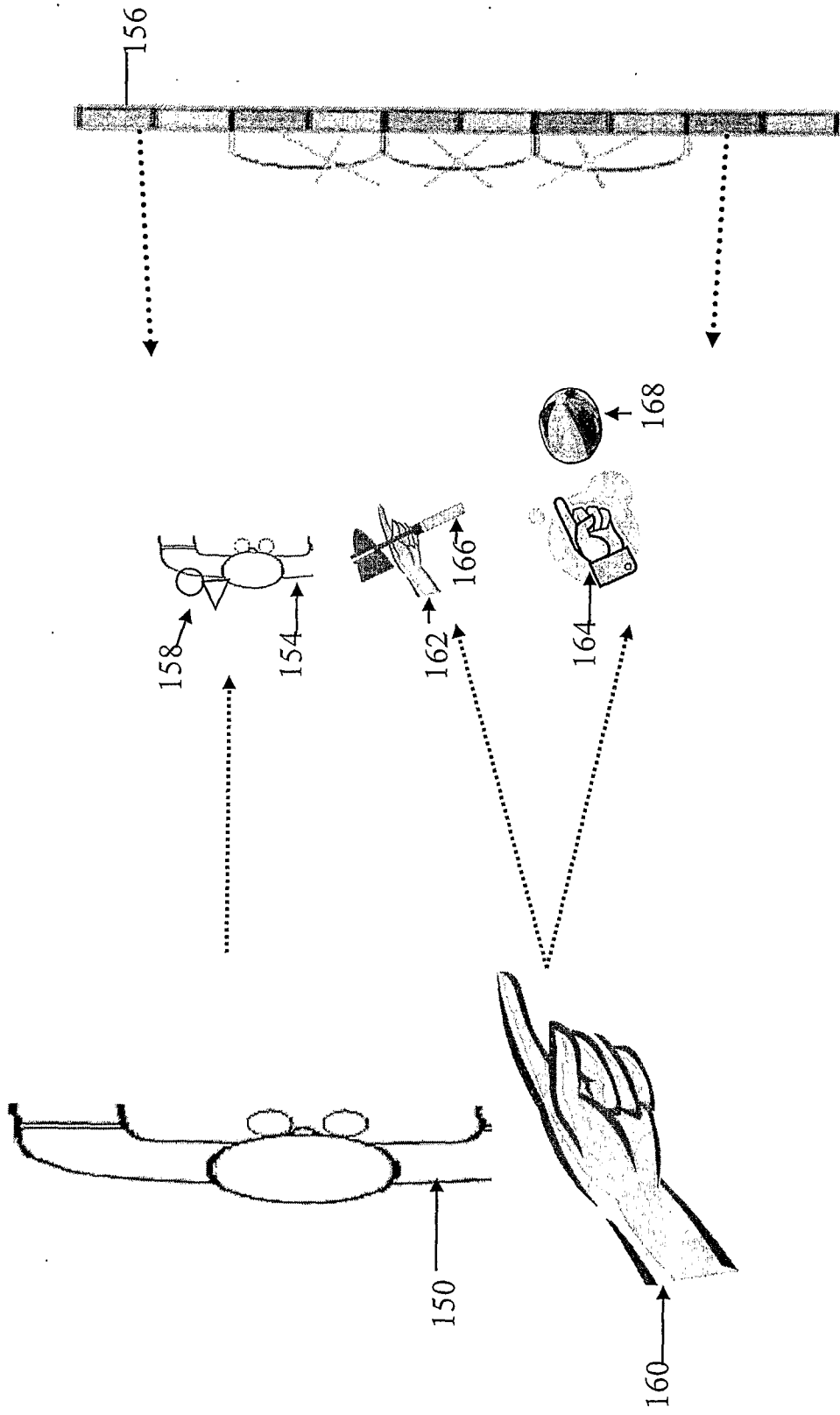


Fig 12

17/24

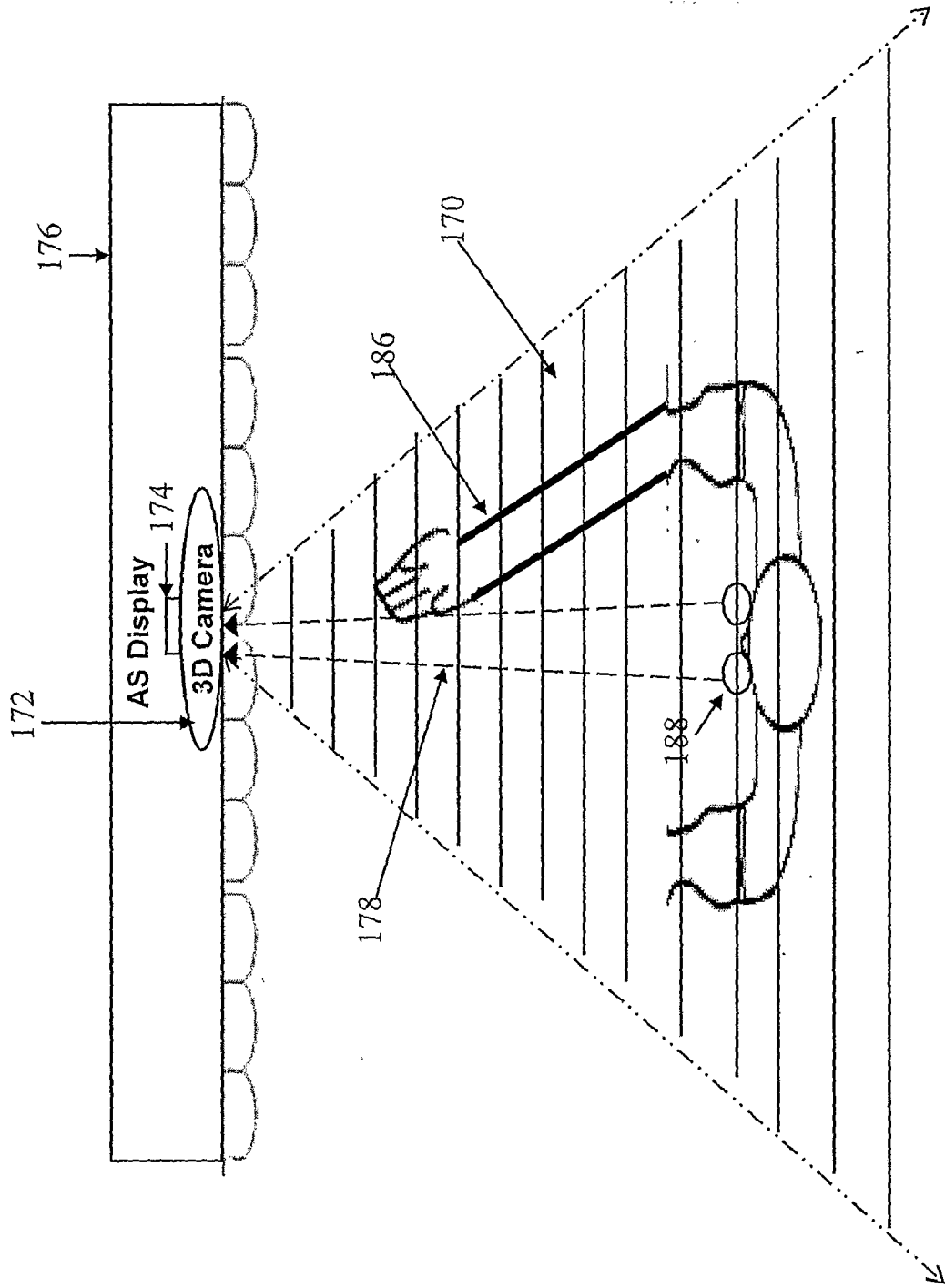
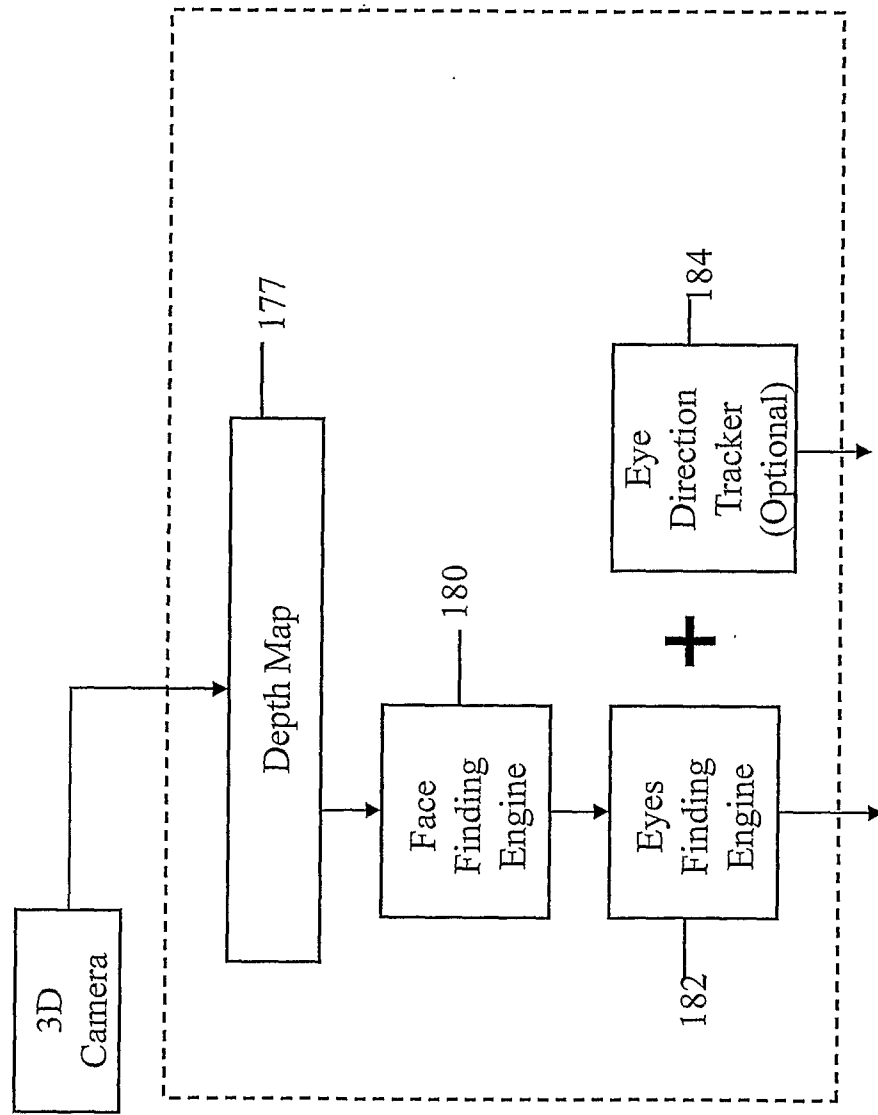


Fig 13A

Fig 13B



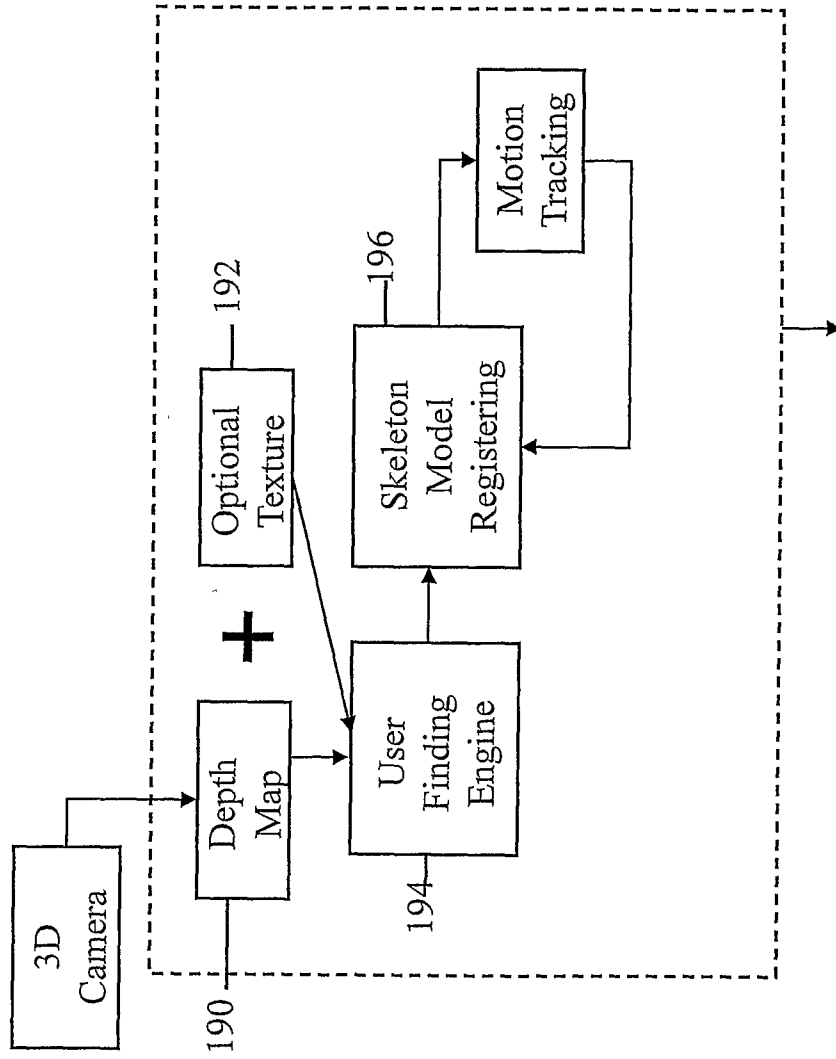


Fig 13C

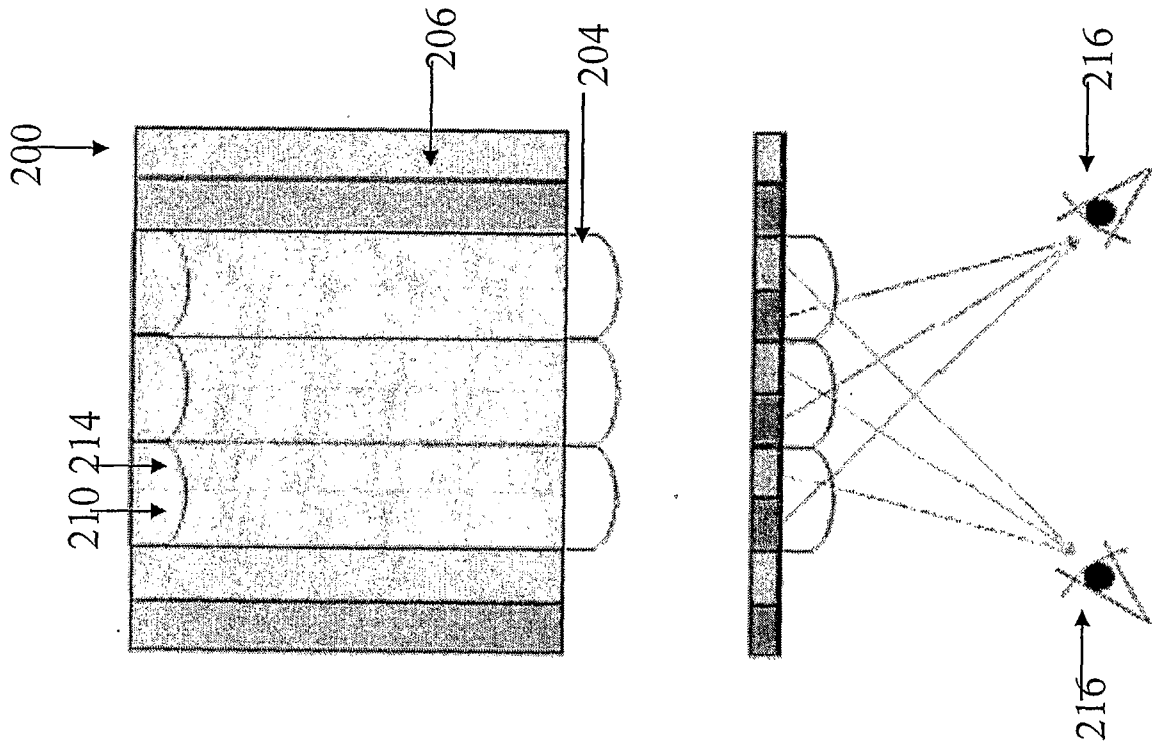


Fig 14

Fig 15

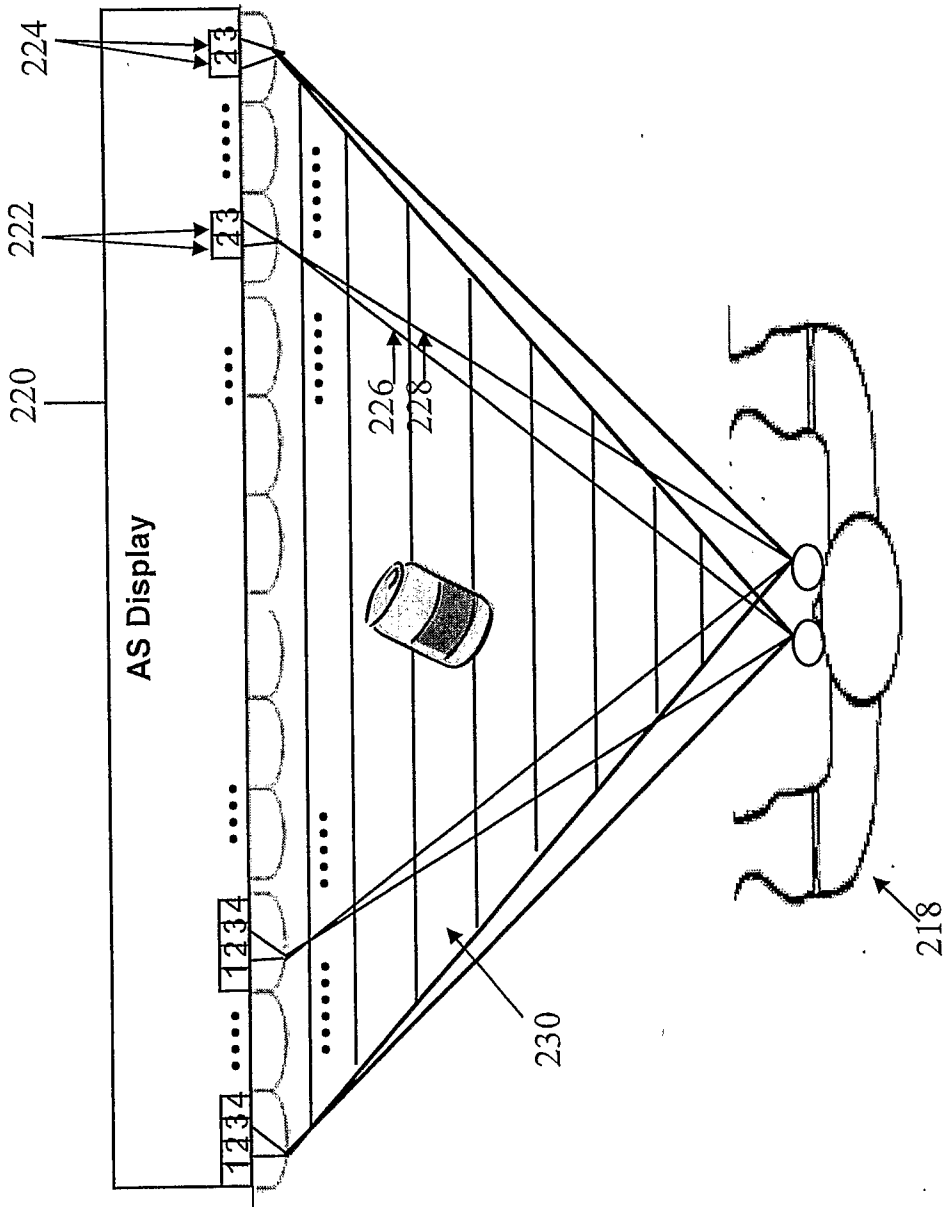
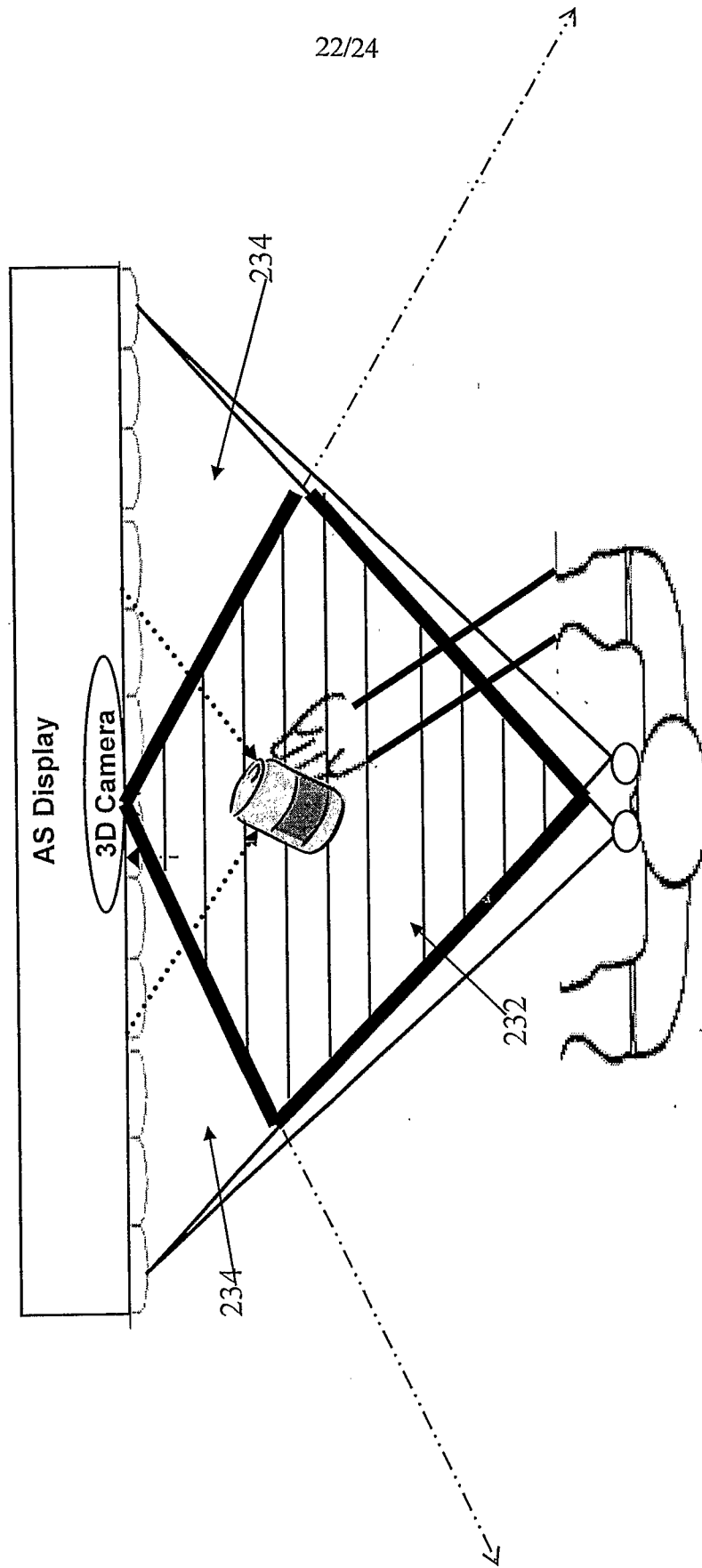


Fig 16



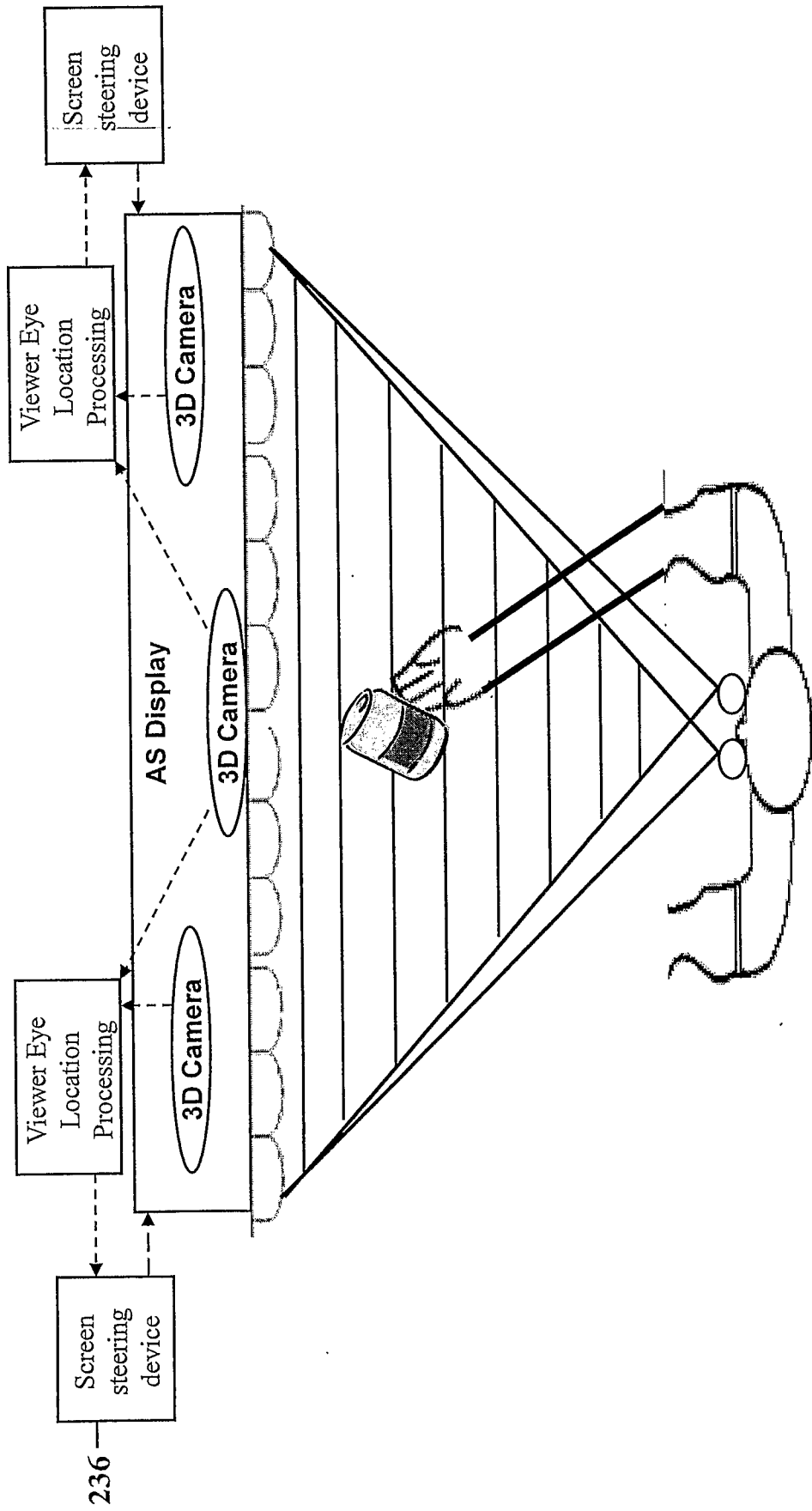


Fig 17

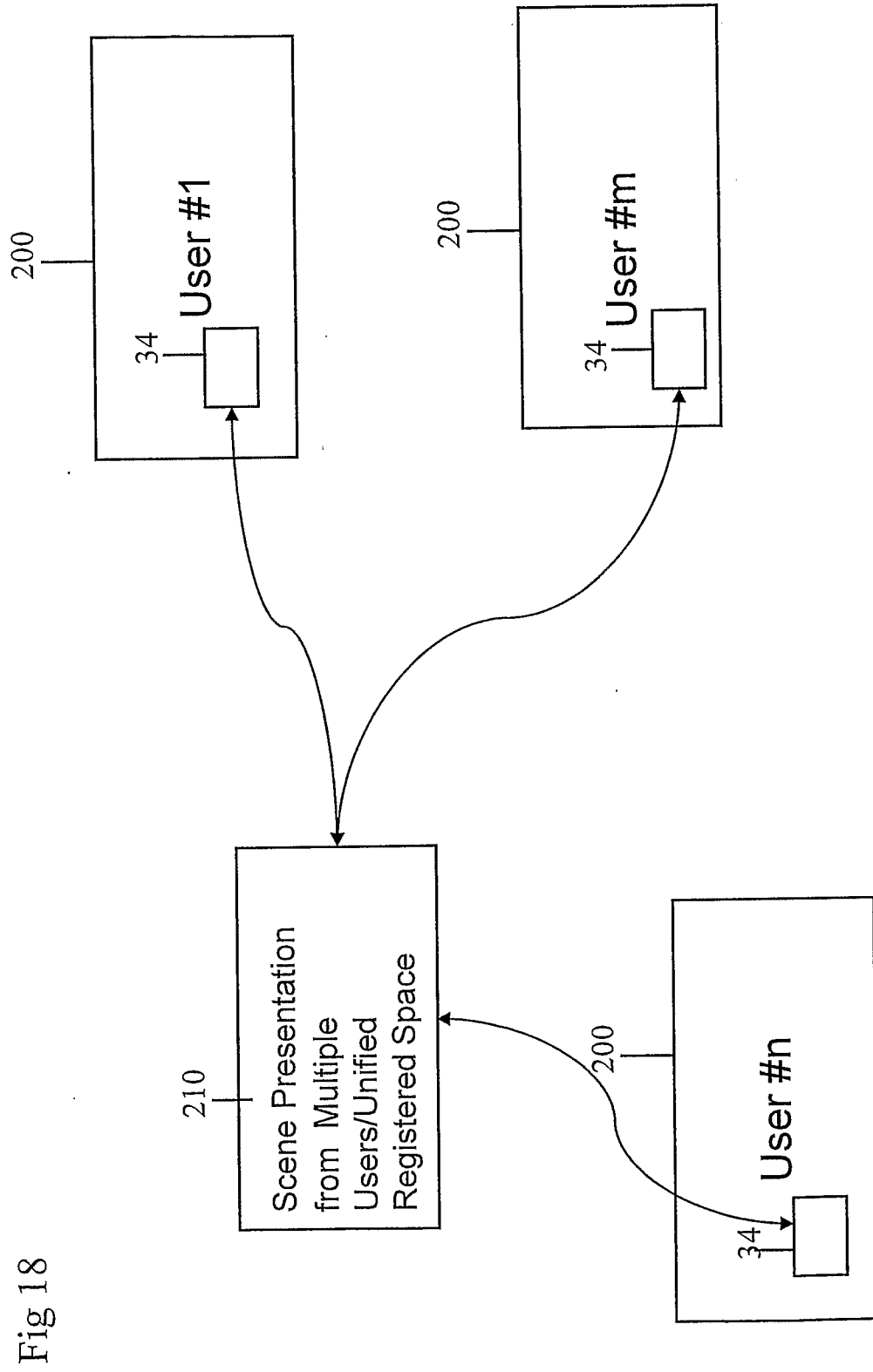


Fig 18

INTERNATIONAL SEARCH REPORT

International application No
PCT/IL2008/000530

A. CLASSIFICATION OF SUBJECT MATTER

INV. G06F3/01 G06F3/03 H04N13/00

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

B. FIELDS SEARCHED

Minimum documentation searched (classification system followed by classification symbols)
G06F H04N

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

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

EPO-Internal, WPI Data

C. DOCUMENTS CONSIDERED TO BE RELEVANT

Category*	Citation of document, with indication, where appropriate, of the relevant passages	Relevant to claim No.
X	US 2004/036763 A1 (SWIFT DAVID C [US] ET AL) 26 February 2004 (2004-02-26)	1-4, 6-17, 19-21
Y	abstract	5
A	paragraphs [0012] - [0022] - paragraphs [0042] - [0167]; claims 1-8	18
X	US 2006/012675 A1 (ALPASLAN ZAHIR Y [US] ET AL) 19 January 2006 (2006-01-19)	1-4, 6-17, 19-21
Y	abstract	5
A	paragraphs [0006] - [0009]; figures 1a-7b paragraphs [0018] - [0057]; claims 1-34	18
Y	US 2004/223218 A1 (PUTILIN ANDREY N [RU] ET AL) 11 November 2004 (2004-11-11)	5
A	abstract the whole document	1-4, 6-21
	----- -/--	

Further documents are listed in the continuation of Box C.

See patent family annex.

* Special categories of cited documents :

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

- *T* later document published after the international filing date or priority date and not in conflict with the application but cited to understand the principle or theory underlying the invention
- *X* document of particular relevance; the claimed invention cannot be considered novel or cannot be considered to involve an inventive step when the document is taken alone
- *Y* document of particular relevance; the claimed invention cannot be considered to involve an inventive step when the document is combined with one or more other such documents, such combination being obvious to a person skilled in the art.
- *&* document member of the same patent family

Date of the actual completion of the international search

14 August 2008

Date of mailing of the international search report

03/09/2008

Name and mailing address of the ISA/

European Patent Office, P.B. 5818 Patentlaan 2
NL - 2280 HV Rijswijk
Tel. (+31-70) 340-2040, Tx. 31 651 epo nl,
Fax: (+31-70) 340-3016

Authorized officer

Wierzejewski, Piotr

INTERNATIONAL SEARCH REPORT

International application No
PCT/IL2008/000530

C(Continuation). DOCUMENTS CONSIDERED TO BE RELEVANT		
Category*	Citation of document, with indication, where appropriate, of the relevant passages	Relevant to claim No.
X	US 2004/160389 A1 (SUYAMA SHIRO [JP] ET AL) 19 August 2004 (2004-08-19)	5
A	abstract paragraphs [0016] - [0063]	1-4,6-21
A	PASTOOR S ET AL: "17.D: INVITED PAPER: AUTOSTEREOSCOPIC USER-COMPUTER INTERFACE WITH VISUALLY CONTROLLED INTERACTIONS" 1997 SID INTERNATIONAL SYMPOSIUM DIGEST OF TECHNICAL PAPERS. BOSTON, MAY 13 - 15, 1997; [SID INTERNATIONAL SYMPOSIUM DIGEST OF TECHNICAL PAPERS], SANTA ANA, SID, US, vol. VOL. 28, 13 May 1997 (1997-05-13), pages 277-280, XP000722705 the whole document	1-21
A	ALPASLAN ZAHIR Y ET AL: "Three-dimensional interaction with autostereoscopic displays" PROCEEDINGS OF THE SPIE, SPIE, BELLINGHAM, VA, vol. 5291, 19 January 2004 (2004-01-19), pages 227-236, XP007905371 ISSN: 0277-786X the whole document	1-21
A	US 2005/264527 A1 (LIN JULIUS J [US]) 1 December 2005 (2005-12-01) abstract paragraphs [0017] - [0027]; figures 1-45 paragraphs [0081] - [0228]	1-21

INTERNATIONAL SEARCH REPORT

Information on patent family members

International application No

PCT/IL2008/000530

Patent document cited in search report	Publication date	Patent family member(s)	Publication date
US 2004036763	A1	26-02-2004	NONE
US 2006012675	A1	19-01-2006	NONE
US 2004223218	A1	11-11-2004	NONE
US 2004160389	A1	19-08-2004	US 2004164927 A1 26-08-2004
US 2005264527	A1	01-12-2005	NONE