



- (51) **International Patent Classification:**  
A63F 13/10 (2006.01) A63F 13/06 (2006.01)
- (21) **International Application Number:**  
PCT/US2012/034695
- (22) **International Filing Date:**  
23 April 2012 (23.04.2012)
- (25) **Filing Language:** English
- (26) **Publication Language:** English
- (30) **Priority Data:**  
61/480,062 28 April 2011 (28.04.2011) US  
13/237,584 20 September 2011 (20.09.2011) US
- (71) **Applicant (for all designated States except US): MICROSOFT CORPORATION [US/US];** One Microsoft Way, Redmond, WA 98052-6399 (US).
- (72) **Inventors: LANSDALE, Thomas William;** c/o Microsoft Corporation, LCA - International Patents, One Microsoft Way, Redmond, WA 98052-6399 (US). **GRIFFITHS, Charles Robert;** c/o Microsoft Corporation, LCA - International Patents, One Microsoft Way, Redmond, WA 98052-6399 (US).
- (81) **Designated States (unless otherwise indicated, for every kind of national protection available):** AE, AG, AL, AM, AO, AT, AU, AZ, BA, BB, BG, BH, BR, BW, BY, BZ, CA, CH, CL, CN, CO, CR, CU, CZ, DE, DK, DM, DO,

DZ, EC, EE, EG, ES, FI, GB, GD, GE, GH, GM, GT, HN, HR, HU, ID, IL, IN, IS, JP, KE, KG, KM, KN, KP, KR, KZ, LA, LC, LK, LR, LS, LT, LU, LY, MA, MD, ME, MG, MK, MN, MW, MX, MY, MZ, NA, NG, NI, NO, NZ, OM, PE, PG, PH, PL, PT, QA, RO, RS, RU, RW, SC, SD, SE, SG, SK, SL, SM, ST, SV, SY, TH, TJ, TM, TN, TR, TT, TZ, UA, UG, US, UZ, VC, VN, ZA, ZM, ZW.

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

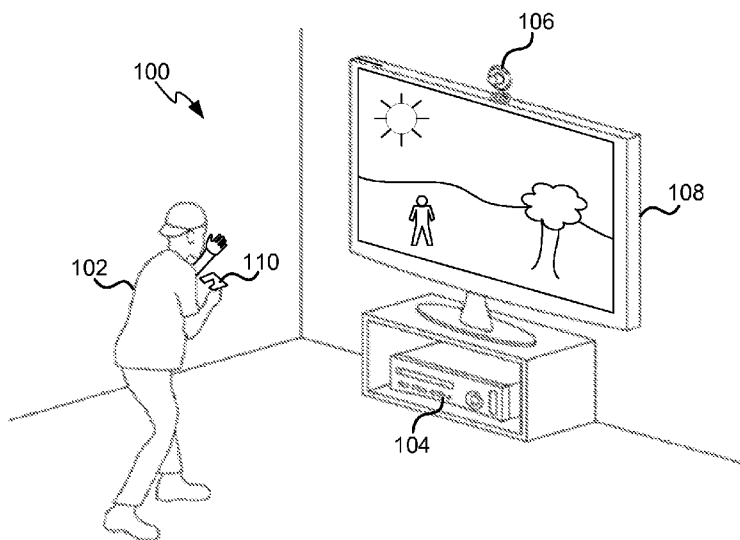
**Declarations under Rule 4.17:**

- as to applicant's entitlement to apply for and be granted a patent (Rule 4.17(ii))
- as to the applicant's entitlement to claim the priority of the earlier application (Rule 4.17(iii))

**Published:**

- without international search report and to be republished upon receipt of that report (Rule 48.2(g))

(54) **Title:** CONTROL OF SEPARATE COMPUTER GAME ELEMENTS



**FIG. 1**

(57) **Abstract:** Control of separate computer game elements is described. In one example, a computer game system displays a computer game to a user, the computer game comprising at least one avatar and an environment. Manual input commands are received from a hand-operated controller operated by the user, and are used to control the display of the avatar in the computer game. Substantially concurrently with the manual input, a stream of images from an image capture device depicting the user is also received. A portal, separate to the avatar, is displayed in a region of the computer game display, which shows the received stream of images to the user. The stream of images of the user is analyzed and a gesture performed by the user is recognized. The recognized gesture is used to control displayed interaction between the portal and the environment.



## CONTROL OF SEPARATE COMPUTER GAME ELEMENTS

### BACKGROUND

[0001] Existing video and computer game control systems use hand held  
controllers which incorporate buttons and joysticks to enable a player to control an  
5 avatar or other objects depicted at a game display. Design of these types of hand  
held controllers seeks to enable fine grained control of game play in robust, easy to  
use and intuitive manners.

[0002] More recently, some computer game control systems use voice  
recognition technology and gesture recognition to enable a player to control a game  
10 interface. In this situation gamers have no hand held controller and are able to  
interact with the game in a straightforward manner without being restricted by  
physical user input devices such as hand held controllers.

[0003] The embodiments described below are not limited to implementations  
which solve any or all of the disadvantages of known game control systems

### 15 SUMMARY

[0004] The following presents a simplified summary of the disclosure in  
order to provide a basic understanding to the reader. This summary is not an  
extensive overview of the disclosure and it does not identify key/critical elements  
or delineate the scope of the specification. Its sole purpose is to present a selection  
20 of concepts disclosed herein in a simplified form as a prelude to the more detailed  
description that is presented later.

[0005] Control of separate computer game elements is described. In one  
example, a computer game system displays a computer game to a user, the  
computer game comprising at least one avatar and an environment. Manual input  
25 commands are received from a hand-operated controller operated by the user, and  
are used to control the display of the avatar in the computer game. Substantially  
concurrently with the manual input, a stream of images from an image capture  
device depicting the user is also received. A portal, separate to the avatar, is  
displayed in a region of the computer game display, which shows the received  
30 stream of images to the user. The stream of images of the user is analyzed and a

gesture performed by the user is recognized. The recognized gesture is used to control displayed interaction between the portal and the environment.

[0006] Many of the attendant features will be more readily appreciated as the same becomes better understood by reference to the following detailed description considered in connection with the accompanying drawings.

### **DESCRIPTION OF THE DRAWINGS**

[0007] The present description will be better understood from the following detailed description read in light of the accompanying drawings, wherein:

FIG. 1 is a schematic diagram of a player holding a game controller and gesturing before a game apparatus having a depth camera;

FIG. 2 is a schematic diagram of a game system incorporating an image capture device, a hand held controller, a computing device and a display;

FIG. 3 is a plan view of a hand held controller;

FIG. 4 is a perspective view of the hand held controller of FIG. 3;

FIG. 5 is a schematic diagram of a display during a first game play;

FIG. 6 is a schematic diagram of a display during a second game play;

FIG. 7 is a flow diagram of a method of operation of a game system;

FIG. 8 is a flow diagram of a gesture recognition operation; and

FIG. 9 illustrates an exemplary computing-based device in which embodiments of a game system may be implemented.

Like reference numerals are used to designate like parts in the accompanying drawings.

### **DETAILED DESCRIPTION**

[0008] The detailed description provided below in connection with the appended drawings is intended as a description of the present examples and is not intended to represent the only forms in which the present example may be constructed or utilized. The description sets forth the functions of the example and the sequence of steps for constructing and operating the example. However, the same or equivalent functions and sequences may be accomplished by different examples.

[0009] Although the present examples are described and illustrated herein as being implemented in a game system for two dimensional side-scrolling platformer games, the system described is provided as an example and not a limitation. As those skilled in the art will appreciate, the present examples are suitable for application in a variety of different types of game systems.

[0010] Reference is first made to FIG. 1, which illustrates an example control system 100 for controlling a computer game. In this example, the control system comprises both a hand-held controller and a camera-based control system. By integrating both types of control a game player experiences the benefits of both types of control system. Integration is achieved as described herein to enable fine grained control of game systems in a robust, easy to use manner which enhances the player experience. FIG. 1 shows a user 102 playing, in this illustrative example, a two dimensional side-scrolling platformer game. This type of game is may be clearly depicted in two dimensional drawings; however, the methods described herein are also applicable to three dimensional games, augmented reality applications and games of other types. In some examples, camera-based control system 100 can be used to, among other things, determine body pose, bind, recognize, analyze, track, associate to a human target, provide feedback, and/or adapt to aspects of a human target such as user 102 (also referred to herein as a player). In this example one player is depicted for clarity. However, two or more players may also use the control system at the same time.

[0011] The camera-based control system 100 comprises a computing device 104. The computing device 104 can be a general purpose computer, gaming system or console, or dedicated image processing device. The computing device 104 can include hardware components and/or software components such that the computing device 104 can be used to execute applications such as gaming applications and/or non-gaming applications. The structure of the computing device 104 is discussed hereinafter with reference to FIG. 9.

[0012] The camera-based control system 100 further comprises a capture device 106. The capture device 106 can be, for example, an image sensor or detector that can be used to visually monitor one or more users (such as user 102)

such that gestures performed by the one or more users can be captured, analyzed, processed, and tracked to perform one or more controls or actions within a game or application, as described in more detail below.

[0013] The camera-based control system 100 can further comprise a display  
5 device 108 connected to the computing device 104. The display device can be a television, a monitor, a high-definition television (HDTV), or the like that can provide game or application visuals (and optionally audio) to the user 102.

[0014] In operation, the user 102 can be tracked using the capture device 106  
10 such that the position, movements and size of user 102 can be interpreted by the computing device 104 (and/or the capture device 106) as controls that can be used to affect the application being executed by computing device 104. As a result, the user 102 can move his or her body (or parts of his or her body) to control an executed game or application.

[0015] In the illustrative example of FIG. 1, the application executing on the  
15 computing device 104 is a two dimensional side-scrolling platformer game that the user 102 is playing. In this example, the computing device 104 controls the display device 108 to provide a visual representation of a terrain comprising a landscape, tree, and the sun to the user 102. The computing device 104 also controls the display device 108 to provide a visual representation of a user avatar that the user  
20 102 can control with his or her movements and/or by using a hand held controller 110. For example, the computing device 104 can comprise a body pose estimator that is arranged to recognize and track different body parts of the user, and map these onto the avatar. In this way, the avatar copies the movements of the user 102 such that if the user 102, for example walks in physical space, this causes the user  
25 avatar to walk in game space.

[0016] However, only copying user movements in game space limits the type  
and complexity of the interaction between the user and the game. For example, many in-game controls are momentary actions or commands, which may be triggered using button presses in traditional gaming systems. Examples of these  
30 include actions such as punch, shoot, change weapon, throw, kick, jump, and/or crouch. Such actions or commands may be controlled by recognizing that the user

is performing one of these actions and triggering a corresponding in-game action, rather than merely copying the user's movements. In addition, combinations of user inputs at the hand held controller and user inputs via the camera-based control system may be used to control the game apparatus.

5 [0017] Reference is now made to FIG. 2, which illustrates a schematic diagram of the capture device 106 that can be used in the camera-based control system 100 of FIG. 1. In the example of FIG. 2 the capture device 106 is configured to capture video images with depth information. Such a capture device can be referred to as a depth camera. The depth information can be in the form of a  
10 depth image that includes depth values, i.e. a value associated with each image element of the depth image that is related to the distance between the depth camera and an item or object located at that image element. Note that the term "image element" is used to refer to a pixel, group of pixels, voxel, group of voxels or other higher level component of an image.

15 [0018] The depth information can be obtained using any suitable technique including, for example, time-of-flight, structured light, stereo image, or the like. In some examples, the capture device 106 can organize the depth information into "Z layers," or layers that may be perpendicular to a Z-axis extending from the depth camera along its line of sight.

20 [0019] As shown in FIG. 2, the capture device 106 comprises at least one imaging sensor 200. In the example shown in FIG. 2, the imaging sensor 200 comprises a depth camera 202 arranged to capture a depth image of a scene. The captured depth image can include a two-dimensional (2-D) area of the captured scene where each image element in the 2-D area represents a depth value such as a  
25 length or distance of an object in the captured scene from the depth camera 202.

[0020] The capture device can also include an emitter 204 arranged to illuminate the scene in such a manner that depth information can be ascertained by the depth camera 202. For example, in the case that the depth camera 202 is an infra-red (IR) time-of-flight camera, the emitter 204 emits IR light onto the scene,  
30 and the depth camera 202 is arranged to detect backscattered light from the surface of one or more targets and objects in the scene. In some examples, pulsed infrared

light can be emitted from the emitter 204 such that the time between an outgoing light pulse and a corresponding incoming light pulse can be detected by the depth camera and measured and used to determine a physical distance from the capture device 106 to a location on the targets or objects in the scene. Additionally, in some examples, the phase of the outgoing light wave from the emitter 204 can be compared to the phase of the incoming light wave at the depth camera 202 to determine a phase shift. The phase shift can then be used to determine a physical distance from the capture device 106 to a location on the targets or objects. In a further example, time-of-flight analysis can be used to indirectly determine a physical distance from the capture device 106 to a location on the targets or objects by analyzing the intensity of the reflected beam of light over time via various techniques including, for example, shuttered light pulse imaging.

[0021] In another example, the capture device 106 can use structured light to capture depth information. In such a technique, patterned light (e.g., light displayed as a known pattern such as spot, grid, or stripe pattern, which may also be time-varying) can be projected onto the scene using the emitter 204. Upon striking the surface of one or more targets or objects in the scene, the pattern becomes deformed. Such a deformation of the pattern can be captured by the depth camera 202 and then be analyzed to determine a physical distance from the capture device 106 to a location on the targets or objects in the scene.

[0022] In another example, the depth camera 202 can be in the form of two or more physically separated cameras that view a scene from different angles, such that visual stereo data is obtained that can be resolved to generate depth information. In this case the emitter 204 can be used to illuminate the scene or can be omitted.

[0023] In some examples, in addition or alternative to the depth camera 202, the capture device 106 can comprise a video camera, which is referred to as an RGB camera 206. The RGB camera 206 is arranged to capture sequences of images of the scene at visible light frequencies, and can hence provide images that can be used to augment the depth images. In some examples, the RGB camera 206 can be used instead of the depth camera 202. The capture device 106 can also

optionally comprise a microphone 207 or microphone array (which can be directional and/or steerable), which is arranged to capture sound information such as voice input from the user and can be used for speech recognition.

[0024] The capture device 106 shown in FIG. 2 further comprises at least one processor 208, which is in communication with the imaging sensor 200 (i.e. depth camera 202 and RGB camera 206 in the example of FIG. 2), the emitter 204, and the microphone 207. The processor 208 can be a general purpose microprocessor, or a specialized signal/image processor. The processor 208 is arranged to execute instructions to control the imaging sensor 200, emitter 204 and microphone 207 to capture depth images, RGB images, and/or voice signals. The processor 208 can also optionally be arranged to perform processing on these images and signals, as outlined in more detail hereinafter.

[0025] The capture device 106 shown in FIG. 2 further includes a memory 210 arranged to store the instructions that for execution by the processor 208, images or frames of images captured by the depth camera 202 or RGB camera 206, or any other suitable information, images, or the like. In some examples, the memory 210 can include random access memory (RAM), read only memory (ROM), cache, Flash memory, a hard disk, or any other suitable storage component. The memory 210 can be a separate component in communication with the processor 208 or integrated into the processor 208.

[0026] The capture device 106 also comprises an output interface 212 in communication with the processor 208 and is arranged to provide data to the computing device 104 via a communication link. The communication link can be, for example, a wired connection (such as USB, Firewire, Ethernet or similar) and/or a wireless connection (such as WiFi, Bluetooth® or similar). In other examples, the output interface 212 can interface with one or more communication networks (such as the internet) and provide data to the computing device 104 via these networks.

[0027] A controller 110 is also provided as part of the camera-based control system. The controller may be a hand held controller as depicted schematically in FIG. 1 or may be integral with another larger device that is not hand held. The

controller comprises a plurality of user input devices such as buttons, joysticks, touch pads, switches and enables a player to make input to a game system. User input data is sent from the controller to the computing device 104 by a wired connection and/or a wireless connection.

5 [0028] The computing device 104 executes a number of functions relating to the camera-based gesture recognition, such as an optional body pose estimator 214 and a gesture recognition engine 216. The body pose estimator 214 is arranged to use computer vision techniques to detect and track different body parts of the user. An example of a body pose estimator is given in US patent publication US-2010-10 0278384-A1 "Human body pose estimation" filed 20 May 2009. The body pose estimator 214 can provide an output to the gesture recognition engine in the form of a time-series of data relating to the user's body pose. This can be in the form of a fully tracked skeletal model of the user, or a more coarse identification of the visible body parts of the user. For example, these time-series sequences can 15 comprise data relating to a time-varying angle between at least two body parts of the user, a rate of change of angle between at least two body parts of the user, a motion velocity for at least one body part of the user, or a combination thereof. The different types of data (angles between certain body parts, velocities, etc.) are known as "features". In other examples, the body pose estimator 214 can derive 20 other data sequences (i.e. other features) from the changing pose of the user over time. In further examples, the gesture recognition engine 216 can utilize input (i.e. features) derived from different sources other than the body pose estimator. Application software 218 can also be executed on the computing device 104 and controlled using the gestures. The application software is arranged to control 25 display of the game at a display 108.

[0029] FIG. 3 is a plan view of an example hand held controller 110. It has a generally winged shape with each wing or shoulder 122 being 316 being sized and shaped to be clasped in one hand. The controller comprises a housing supporting a plurality of buttons, switches and joysticks as now described in more detail. 30 However, this is an example only and other types of controller 110 may be used.

[0030] Four digital action buttons 302 are provided on the right face of the controller comprising a green A button, red B button, blue X button and amber Y button. Two analog joysticks 310 and 312 are provided. These joysticks may also be depressed or clicked in to active a digital button beneath each joystick. Digital start 306, back 308 and guide 304 buttons are centrally positioned on the housing.  
5 For example, the guide button is used to turn on the controller and access a menu.

[0031] FIG. 4 is a perspective view of the controller and shows a left bumper 406 and a right bumper 404 each of which are buttons that may be pressed by the user. A left trigger 400 and a right trigger 402 which are both analog are given on the underside of the controller (visible in FIG. 4). A connector 408 may be  
10 provided to enable wired connection to the computing device 104.

[0032] Reference is now made to FIG. 5, which illustrates a schematic diagram of an example computer game display during game play. The example computer game illustrated in FIG. 5 is able to utilize input from both a manual,  
15 hand-held controller (such as that in FIG. 3 and 4) and motion-based input from a camera (such as the capture device 106).

[0033] The example computer game of FIG. 5 comprises an avatar 508 which is the primary game-play element that the user controls using the controller 110. The avatar 508 is located within a game environment comprising active  
20 game-play elements, such as rotatable platform 504 and background elements such as tree 512 and sun 500. The example computer game of FIG. 5 also comprises a portal 510, which is a representation of the user based on the stream of images captured by the capture device. This representation can be in the form of a depth image, an RGB image, or a stylized, processed, or colored version of the same.  
25 The user can therefore see himself or herself represented in the portal 510. In some examples, the portal 510 can be represented as a TV screen to reflect the display device on which the user is playing the game.

[0034] The portal 510 acts as a secondary user-controlled game-play element separate from the avatar 508, and is controlled by the capture device 106. The  
30 computer game is configured such that the user can "reach outside" the portal by extending one or more body parts outside of the boundaries shown in the portal

510. When this is performed, a computer-generated representation of the extended body part is displayed, such as computer-generated hand 506 reaching out from the right-hand side of the portal 510 in FIG. 5. The computer-generated representation of the extended body part can be used to interact with game-play elements such as the rotatable platform 504.

[0035] When playing the example game in FIG. 5, the user is aiming to enable the avatar 508 to proceed from the left to the right side of the environment. The user can control the avatar 508 directly using the controller 110, for example to move left or right, jump, duck, etc. Substantially concurrently with this, the user can control the portal 510 using movement captured by the capture device 106 in order to facilitate the avatar's progress. For example, in the case of the game shown in FIG. 5, the user is able to reach out from the portal 510 and punch the rotatable platform 504 in order to make it spin such that the avatar 508 can pass it and proceed further with the game.

[0036] Further game-play elements can also be present in some examples. For example, chain 502 is shown in FIG. 5, and this can interact with the portal 510. For example, the user can reach out from the portal 510 and grab the chain and pull their hand back towards the portal 510 in order to move the portal 510 across the environment, so that it can be positioned as desired to interact with game-play elements. In another example, the user can move the portal 510 using the chain 502 by reaching out to "grab" the chain then while continuing to reach out, sliding their hand along the chain in the direction opposite the one they want the portal to travel, and then finally (and optionally, as the interaction is now complete) bringing their hand back to a rest position.

[0037] FIG. 6 shows a schematic diagram of a further example computer game display. In this example, the avatar 508 is in a more complex environment comprising a plurality of platforms and other game-play elements, such as rotatable platform 612 and pushable block 610. In this example, the user can again control the movement of the avatar 508 using the controller 110. In addition, in this example, a plurality of portals are present, which are initially concealed. In this example the portals are shown concealed within presents (such as present 602),

although they can be concealed in any suitable manner. The user can reveal a portal by performing a gesture. For example, the user can perform a punch forwards gesture to tear through the present and reveal the portal. This is illustrated in the example of present 604, which has been partially torn to reveal the portal inside. Once a portal has been revealed, such as portal 606, the user can reach out from it to interact with game play elements in the environment, e.g by rotating rotatable platform 612 or sliding pushable block 610, thereby facilitating the avatar's progress through the game.

[0038] In further examples, many other different types of interaction are possible between the portal 510, 606 and the game environment. For example, the user can reach out from the portal to punch, pull, push, grab, activate, twist, or pick up in-game elements, or perform any other suitable action.

[0039] Reference is now made to FIG. 7 and 8, which illustrate flowcharts of processes for controlling the operation of the game with combined controller and motion-based input. Firstly, FIG. 7 shows a flowchart for the overall control of the game operation. The process of FIG. 7 can be performed by the computing device 104 described above. The game display, comprising the avatar and game environment is displayed 702 on the display device 108.

[0040] Manual input from a hand-operated controller operated by the user is received 704 at the computing device. Substantially concurrently with this, a stream of images of the user is received 706 from the capture device 106. The stream of images are processed by the computing device and used to display 708 the portal 510, 606 in a portion of the game display. The displayed portal 510, 606 shows at least a portion of the image stream in (at least near) real-time, such that movements of the user can be seen reflected in the portal representation. The computing device can process the image stream to color, stylize, crop, transform, or remove the background from the image stream before displaying it in the portal.

[0041] The image stream from the capture device is also processed to recognize 710 movements of the player. For example, the image stream is processed to determine whether the user is reaching out from the portal, or punching forwards to reveal the portal, as mentioned above. This operation is

discussed in more detail with reference to FIG. 8 below. The result of the movement recognition is a determination of whether interaction is occurring between the portal and the environment, and, if so, what type of interaction (punch, pull, etc.).

5 [0042] The computing device then controls both the display of the avatar on the basis of the manual input and the displayed interaction between the portal and the environment on the basis of the recognized movement of the player. In other words, two separate game-play entities (the avatar and portal) are separately controlled; both by the same user but using different input methodologies  
10 (controller and movement). These two game-playing elements are combined together within a single game to provide an overall game-play experience.

[0043] Reference is now made to FIG. 8, which illustrates a flowchart of a process for recognizing a gesture for portal interaction. The process of FIG. 8 can be implemented at the gesture recognition engine 216 of FIG. 2. The body pose  
15 estimator 214 discussed above with reference to FIG. 2 is used to track 802 one or more body parts of the user. The body pose estimator 214 comprises a skeletal tracker that can identify body parts and joints and track their location and/or velocity. For example, as described below one or more hands of the user can be tracked. In other examples, different body parts can be tracked, such as the feet,  
20 arms, legs or head.

[0044] The tracking of the body part or parts is then used to determine 804  
whether a gesture is being performed. This can be performed in a number of ways, depending on the gestures that can be used in the game. For example, in the case of  
25 the "reaching out" gesture, where the user reaches out from the portal with one or more of their hands, the location of the user's hands can be tracked relative to the boundary of the images from the capture device. If a hand of the user is moved to within a predefined distance of one edge of the image stream, then this triggers a reaching out gesture from the corresponding edge of the portal. In one example, it can be determined whether the user's hand is within a boundary of 10% of the  
30 width or height from the edge of the image stream, although other percentage or absolute values can also be used. In a further example, it can be determined

whether the user's hand has touched or passed the edge of the image stream. In yet further examples, the velocity of the user's hands can be tracked using the body pose estimator, and this can be used to detect a change in the velocity of a hand (e.g. as the user punches) in the direction of one of the edges of the image stream.

5 [0045] In the case of the punch forwards gesture (e.g. as used to reveal the portal as described with reference to FIG. 6), the body pose estimator can be used to track the depth values of the user's hands using the depth camera 202. The punch forwards gesture can then be detected by determining that the distance between the capture device 106 and the user's hand is less than a predetermined  
10 value. In alternative examples, a change of depth for the hand within a predefined time period can be detected, or the velocity of the hand can be tracked and a rapid change in velocity detected.

[0046] In further examples, many different types of gestures can be detected by tracking the location, depth, and/or velocity of one or more body parts in the  
15 image stream. In further examples, higher-level processing can also be applied to the tracked body parts to detect gestures, such as the use of machine learning classifiers (e.g. decision trees).

[0047] If it was determined 804 that no gesture was detected at the current time, then no interaction between the portal and the game environment is displayed  
20 806. The portal still displays the user's movements from the image stream, but no wider interaction with game elements is performed.

[0048] Conversely, if it was determined 804 that a gesture was performed by the user, then this indicates that interaction between the portal and the game environment can be controlled. However, prior to displaying the interaction  
25 between the portal and the game environment, a further optional check can be performed to ensure that accidental or unintended interactions are not triggered.

[0049] To perform the optional additional check, it is determined whether or not the user has both hands on the hand-held controller 110. If the user does have both hands on the controller, then this indicates that the user is not intending to  
30 perform a gesture, but may have accidentally triggered it e.g. by moving their hands

around when using the controller such that they approached the edge of the image stream, or walked towards the capture device.

[0050] In one example, this can be determined by comparing 808 the image of the user to pre-recorded stance examples 810 showing the user holding the controller 110 in both hands. If the image of the user is similar to one or more of these pre-recorded stance examples 810, then this indicates that the user is still holding the controller in both hands. In other examples, this can be determined using the body pose estimator. For example, it can be determined whether both hands are within a predetermined distance of each other (indicating they are both on the controller). The location of the hands can also be compared to other body parts, such as the user's torso, to indicate that both hands are still on the controller. The motion of the user's hands can also be tracked, such that if the user's hands move substantially in sync with each other, then this indicates that they are both holding a common object, such as the controller.

15 [0051] If this optional additional check determines 812 that both hands are holding onto the controller, then this indicates that the gesture triggering was unintended, and no interaction between the portal and the game environment is displayed 806. Conversely, if the optional check determines 812 that both hands are not on the controller, then this indicates that the gesture was intentional. When a gesture is determined to be intentional, the portal is controlled to interact with the game environment, as outlined above (e.g. by punching out from the sides or top of the portal or revealing a portal).

[0052] FIG. 9 illustrates various components of an exemplary computing device 104 which may be implemented as any form of a computing and/or electronic device, and in which embodiments of the above-described game control techniques may be implemented.

[0053] Computing device 104 comprises one or more processors 902 which may be microprocessors, controllers or any other suitable type of processors for processing computing executable instructions to control a game system. In some examples, for example where a system on a chip architecture is used, the processors 30 902 may include one or more fixed function blocks (also referred to as accelerators)

which implement a part of the game control methods in hardware (rather than software or firmware).

[0054] The computing-based device 104 also comprises an input interface 904 arranged to receive input from one or more devices, such as the capture device 5 106 of FIG. 2 and/or the controller of FIG. 3 and FIG. 4. An output interface 906 is also provided and arranged to provide output to, for example, a display system integral with or in communication with the computing-based device (such as display device 108). The display system may provide a graphical user interface, or other user interface of any suitable type although this is not essential. A 10 communication interface 908 may optionally be provided, which can be arranged to communicate with one or more communication networks (e.g. the internet).

[0055] The computer executable instructions may be provided using any computer-readable media that is accessible by computing based device 104. Computer-readable media may include, for example, computer storage media such 15 as memory 910 and communications media. Computer storage media, such as memory 910, includes volatile and non-volatile, 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, EPROM, 20 EEPROM, flash memory or other memory technology, CD-ROM, digital versatile disks (DVD) or other optical storage, magnetic cassettes, magnetic tape, magnetic disk storage or other magnetic storage devices, or any other non-transmission medium that can be used to store information for access by a computing device. In contrast, communication media may embody computer readable instructions, data 25 structures, program modules, or other data in a modulated data signal, such as a carrier wave, or other transport mechanism. As defined herein, computer storage media does not include communication media. Therefore, a computer storage medium should not be interpreted to be a propagating signal per se. Although the computer storage media (memory 910) is shown within the computing-based 30 device 104 it will be appreciated that the storage may be distributed or located

remotely and accessed via a network or other communication link (e.g. using communication interface 908).

[0056] Platform software comprising an operating system 912 or any other suitable platform software may be provided at the computing-based device to  
5 enable application software 218 to be executed on the device. The memory 910 can store executable instructions to implement the functionality of the body pose estimator 214 and the gesture recognition engine 216. The memory 910 can also provide a data store 914, which can be used to provide storage for data used by the processors 902 when performing the game control techniques, such as for any  
10 stance templates, thresholds, parameters, screen space mapping functions, or other data.

[0057] The term 'computer' is used herein to refer to any device with processing capability such that it can execute instructions. Those skilled in the art will realize that such processing capabilities are incorporated into many different  
15 devices and therefore the term 'computer' includes PCs, servers, mobile telephones, personal digital assistants and many other devices.

[0058] The methods described herein may be performed by software in machine readable form on a tangible storage medium e.g. in the form of a computer program comprising computer program code means adapted to perform all the steps  
20 of any of the methods described herein when the program is run on a computer and where the computer program may be embodied on a computer readable medium. Examples of tangible (or non-transitory) storage media include disks, thumb drives, memory etc and do not include propagated signals. The software can be suitable for execution on a parallel processor or a serial processor such that the method steps  
25 may be carried out in any suitable order, or simultaneously.

[0059] This acknowledges that software can be a valuable, separately tradable commodity. It is intended to encompass software, which runs on or controls “dumb” or standard hardware, to carry out the desired functions. It is also intended to encompass software which “describes” or defines the configuration of  
30 hardware, such as HDL (hardware description language) software, as is used for

designing silicon chips, or for configuring universal programmable chips, to carry out desired functions.

[0060] Those skilled in the art will realize that storage devices utilized to store program instructions can be distributed across a network. For example, a remote computer may store an example of the process described as software. A local or terminal computer may access the remote computer and download a part or all of the software to run the program. Alternatively, the local computer may download pieces of the software as needed, or execute some software instructions at the local terminal and some at the remote computer (or computer network).

10 Those skilled in the art will also realize that by utilizing conventional techniques known to those skilled in the art that all, or a portion of the software instructions may be carried out by a dedicated circuit, such as a DSP, programmable logic array, or the like.

[0061] Any range or device value given herein may be extended or altered without losing the effect sought, as will be apparent to the skilled person.

[0062] 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.

[0063] It will be understood that the benefits and advantages described above may relate to one embodiment or may relate to several embodiments. The embodiments are not limited to those that solve any or all of the stated problems or those that have any or all of the stated benefits and advantages. It will further be understood that reference to 'an' item refers to one or more of those items.

[0064] The steps of the methods described herein may be carried out in any suitable order, or simultaneously where appropriate. Additionally, individual blocks may be deleted from any of the methods without departing from the spirit and scope of the subject matter described herein. Aspects of any of the examples described above may be combined with aspects of any of the other examples described to form further examples without losing the effect sought.

[0065] The term 'comprising' is used herein to mean including the method blocks or elements identified, but that such blocks or elements do not comprise an exclusive list and a method or apparatus may contain additional blocks or elements.

[0066] It will be understood that the above description is given by way of  
5 example only and that various modifications may be made by those skilled in the art. The above specification, examples and data provide a complete description of the structure and use of exemplary embodiments. Although various embodiments have been described above with a certain degree of particularity, or with reference to one or more individual embodiments, those skilled in the art could make  
10 numerous alterations to the disclosed embodiments without departing from the spirit or scope of this specification.

**CLAIMS**

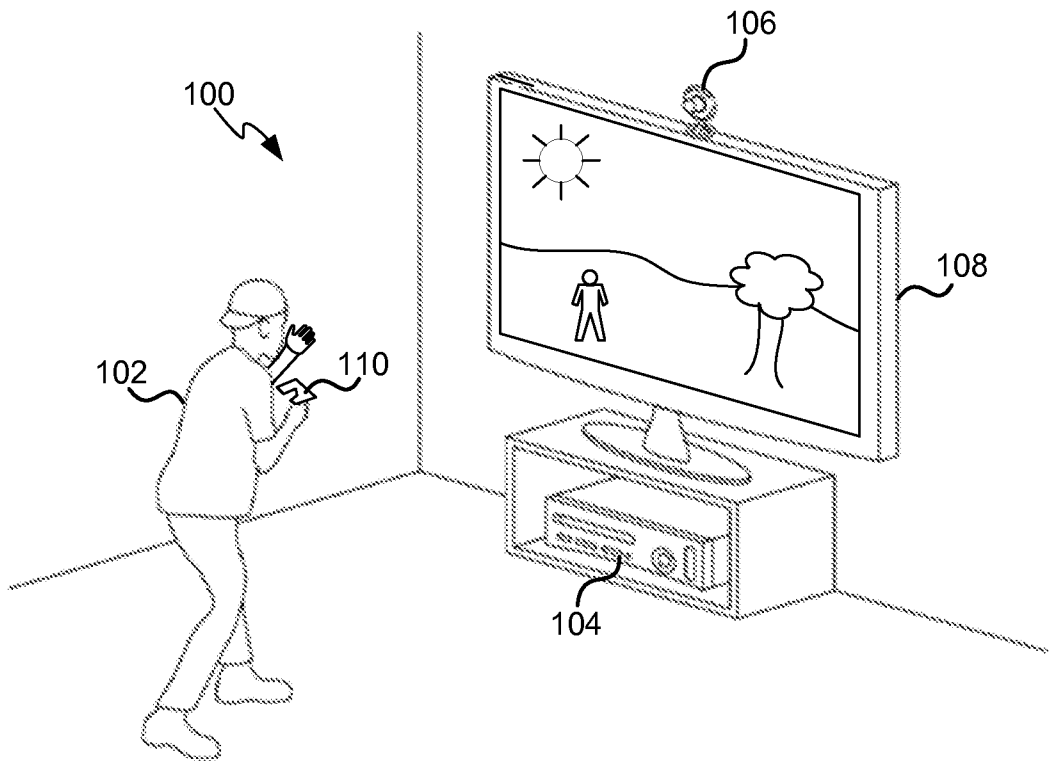
1. A method of controlling a computer game system comprising:  
displaying a computer game at a display viewable by a user of the game, the display comprising at least one avatar and an environment;  
receiving manual input from a hand-operated controller operated by the user;  
receiving a stream of images from an image capture device depicting the user;  
displaying a portal as part of the displayed computer game, the portal being a region of the display representing the stream of images depicting the user;  
recognizing a gesture of the user by analyzing the stream of images; and  
controlling the display of the avatar on the basis of the manual input and controlling displayed interaction between the portal and the environment on the basis of the recognized gesture of the user.
2. A method as claimed in claim 1, wherein the steps of recognizing the gesture and receiving manual input occur substantially concurrently.
3. A method as claimed in any of the preceding claims, wherein the step of recognizing a gesture of the user by analyzing the stream of images comprises tracking at least one body part of the user in the stream of images.
4. A method as claimed in claim 3, wherein the tracking comprises tracking a location for the at least one body part of the user in the stream of images.
5. A method as claimed in claim 4, wherein the location tracked for the at least one body part is a distance between the at least one body part and the image capture device.
6. A method as claimed in claim 3, wherein the tracking comprises tracking a motion velocity for the at least one body part of the user in the stream of images.

7. A method as claimed in claim 3, wherein the capture device comprises a depth camera arranged to capture a sequence of depth images of the user, and wherein the step of tracking is performed using a body-pose estimator taking an input from the depth camera.

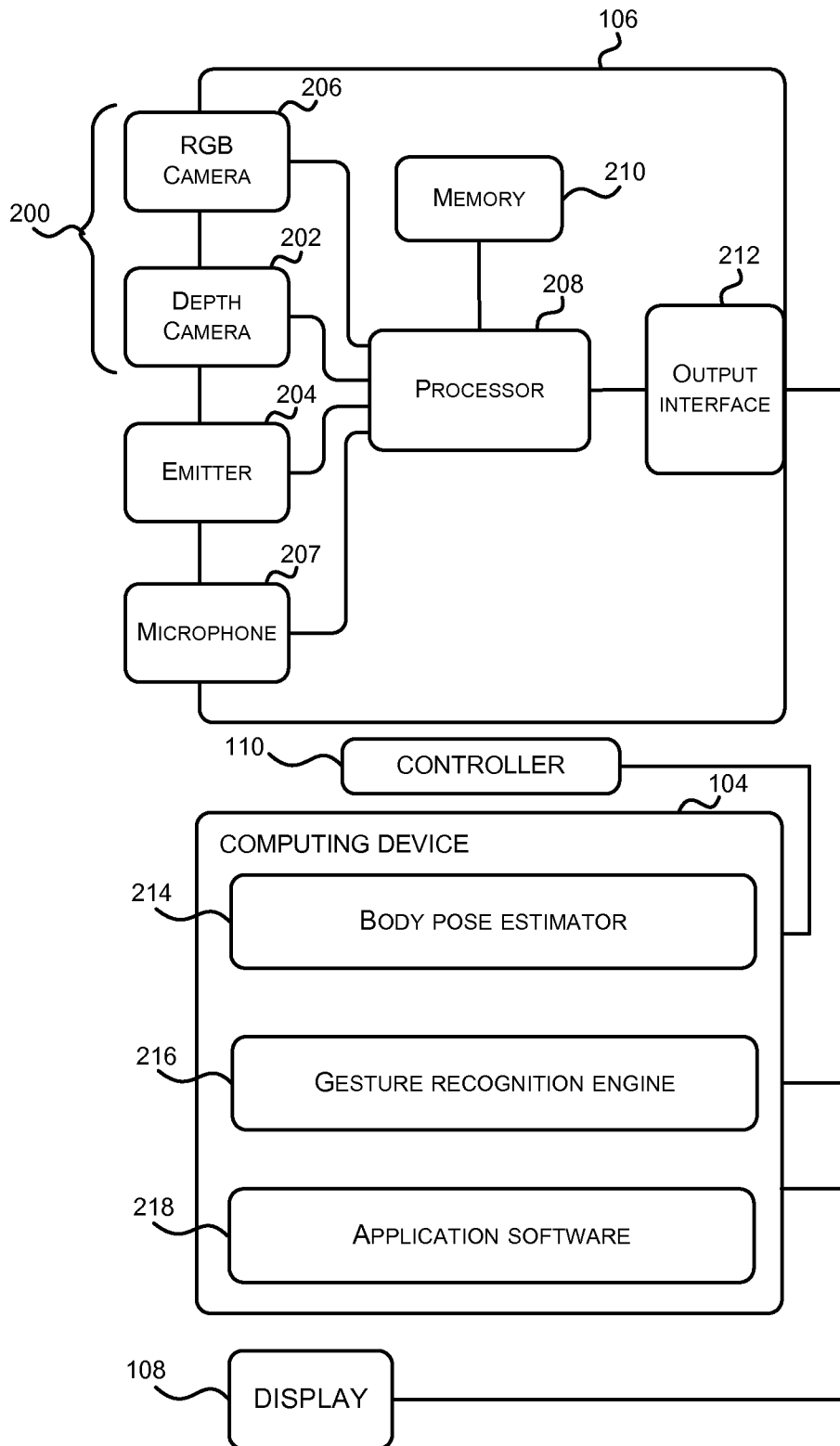
8. A method as claimed in any of the preceding claims, wherein the step of recognizing a gesture comprises determining whether the user is holding the controller with both hands.

9. A computer game system, comprising:  
an input interface arranged to receive a stream of images depicting a user from a capture device, and receive manual input commands from a hand-operated controller operable by the user;  
an output interface arranged to provide a computer game display to a display device viewable by a user of the computer game system, the computer game display comprising at least one avatar and an environment;  
a processor arranged to: control the output interface to display a portal as part of the computer game display, the portal being a region of the display representing the stream of images depicting the user, recognizing a gesture of the user by analyzing the stream of images, control the display of the avatar on the basis of the manual input and control displayed interaction between the portal and the environment on the basis of the recognized gesture of the user.

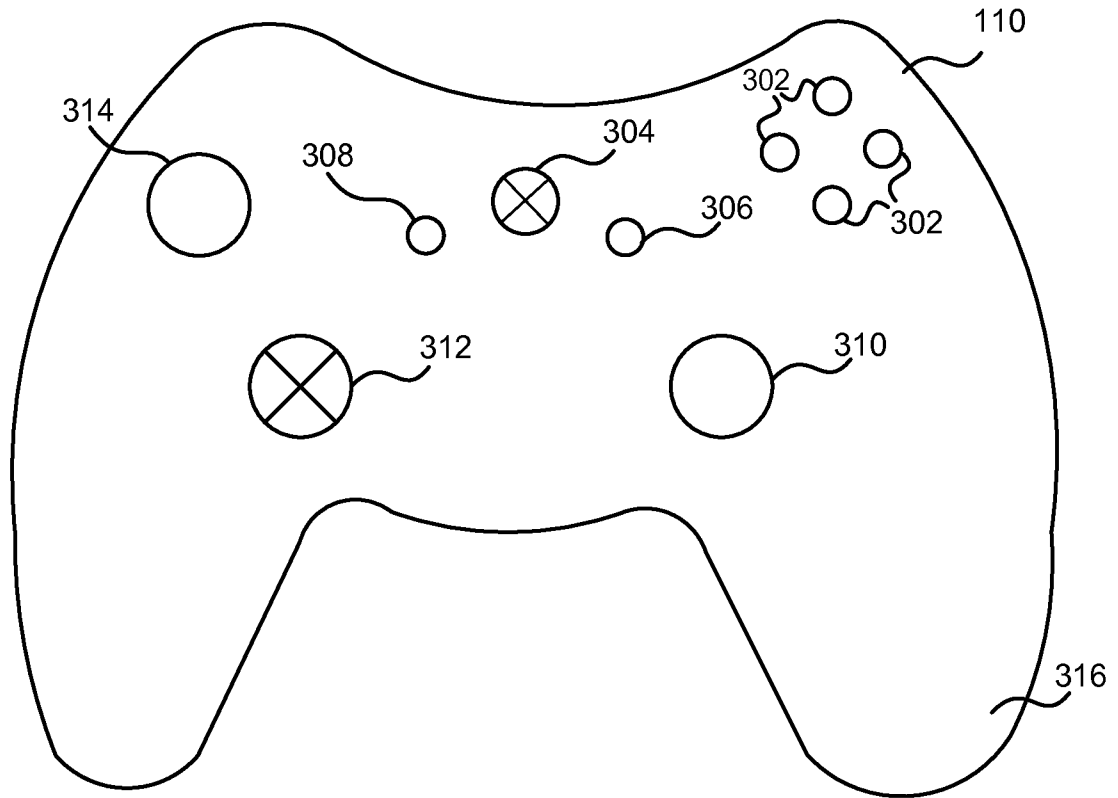
10. A system according to claim 9, wherein the capture device comprises a depth camera.



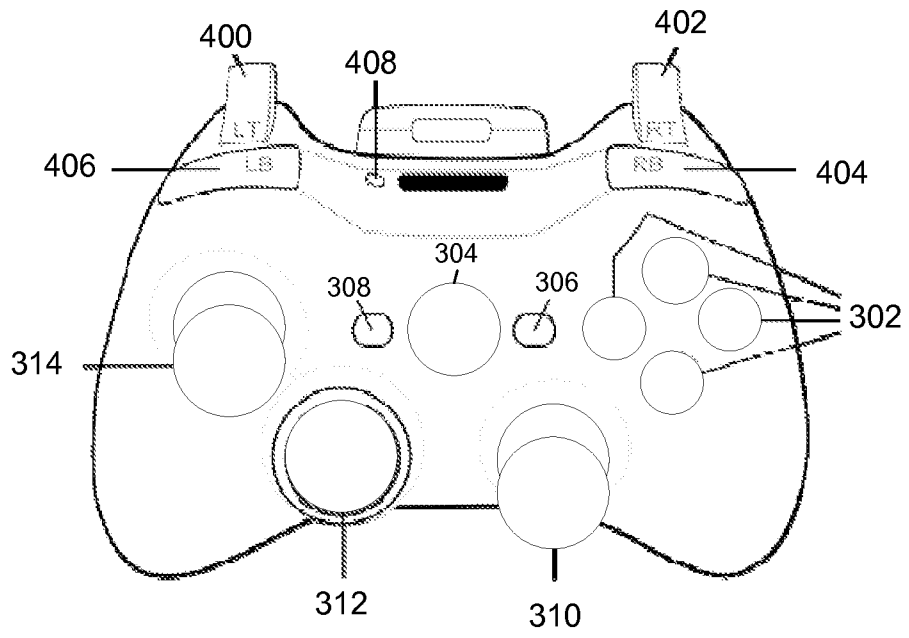
**FIG. 1**



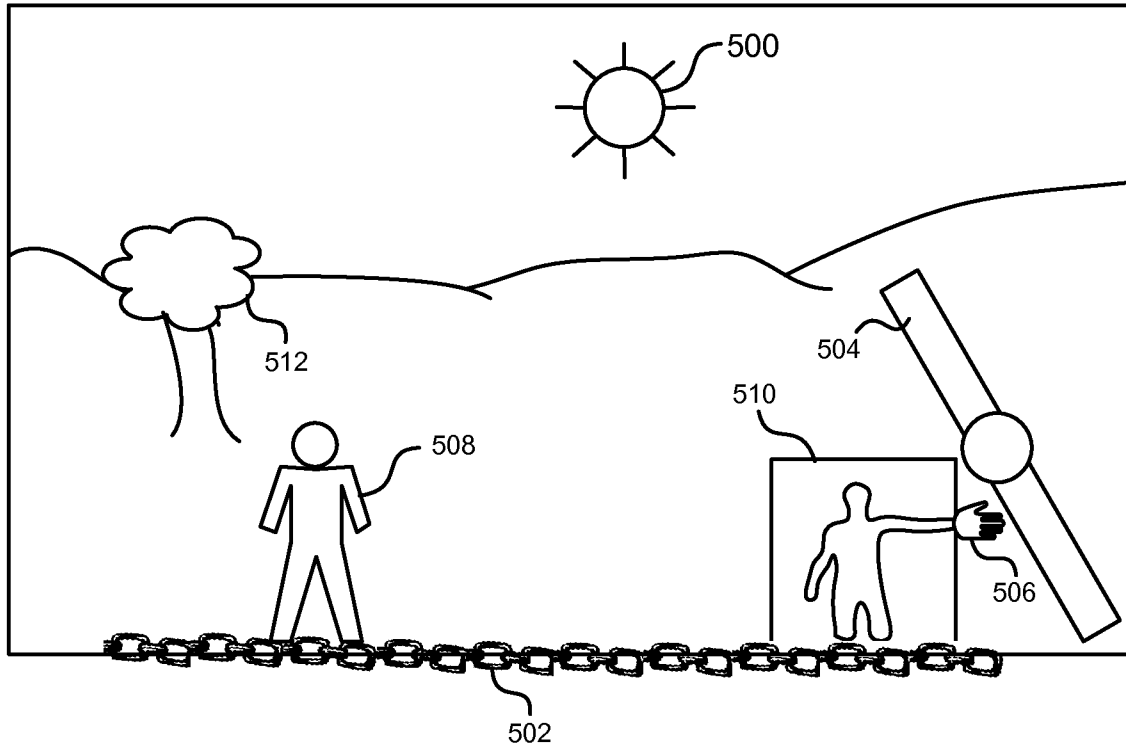
**FIG. 2**



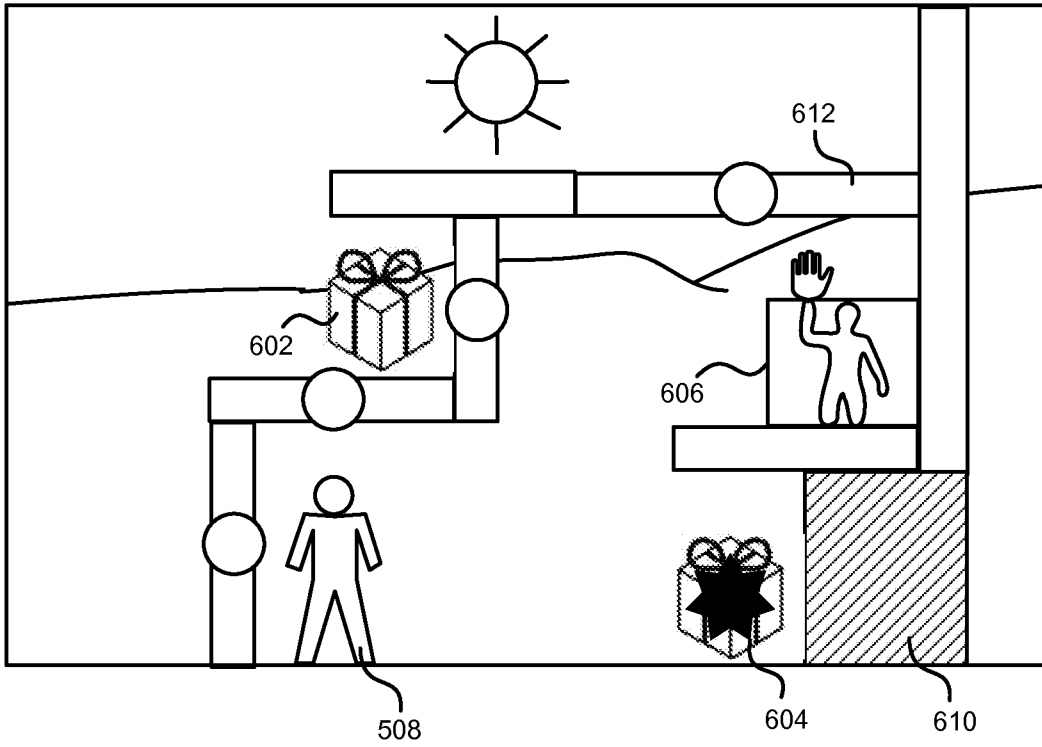
**FIG. 3**



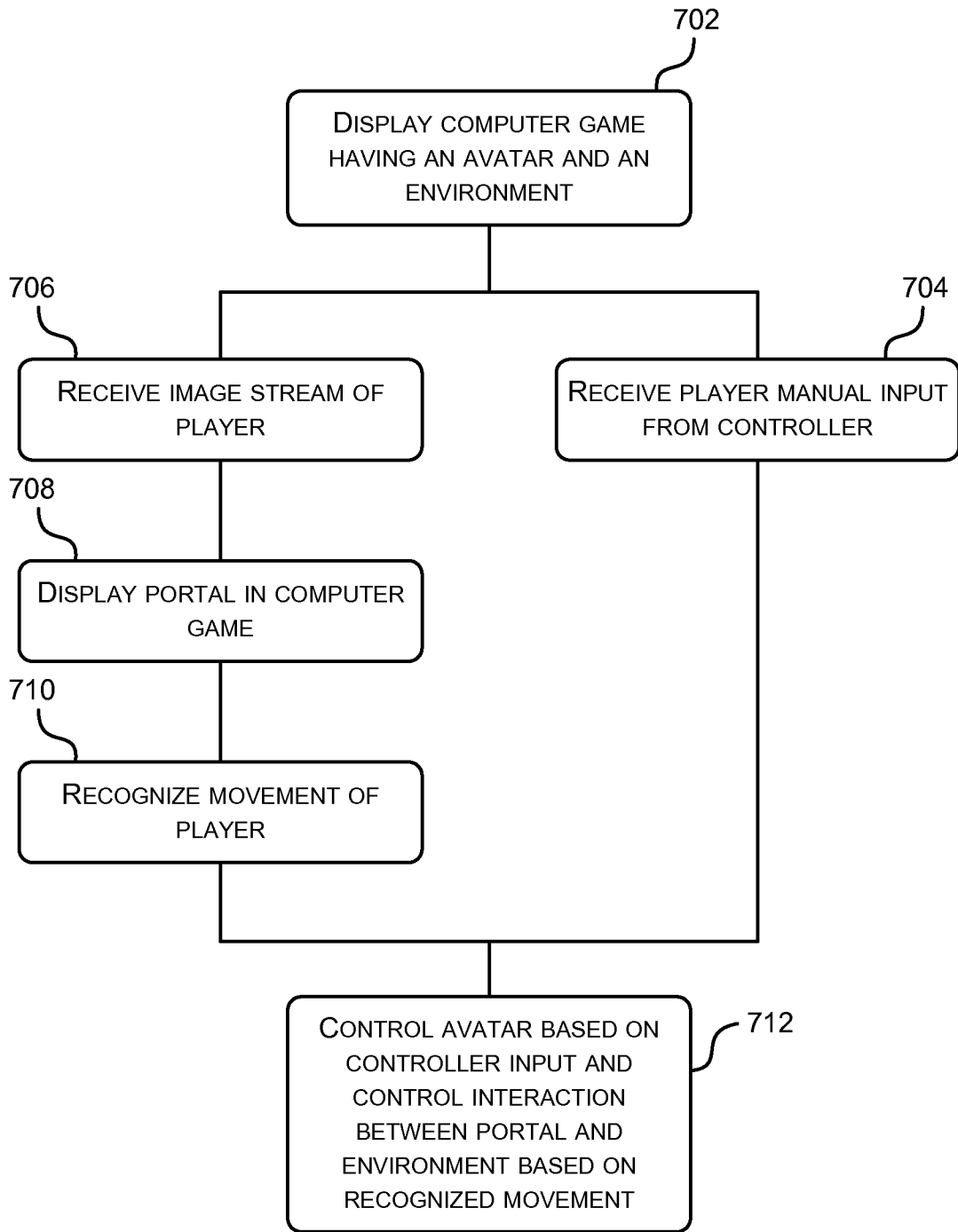
**FIG. 4**



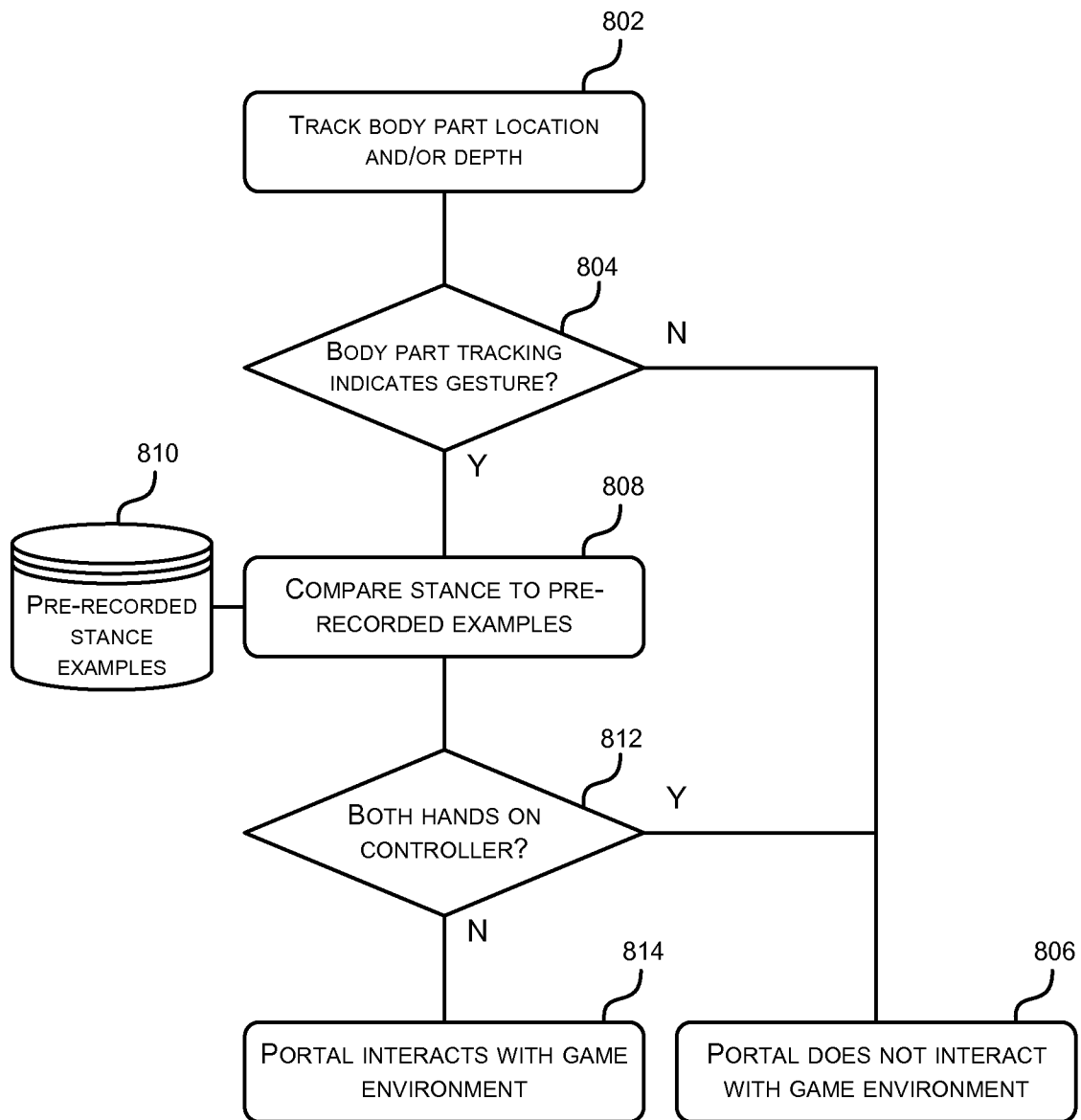
**FIG. 5**



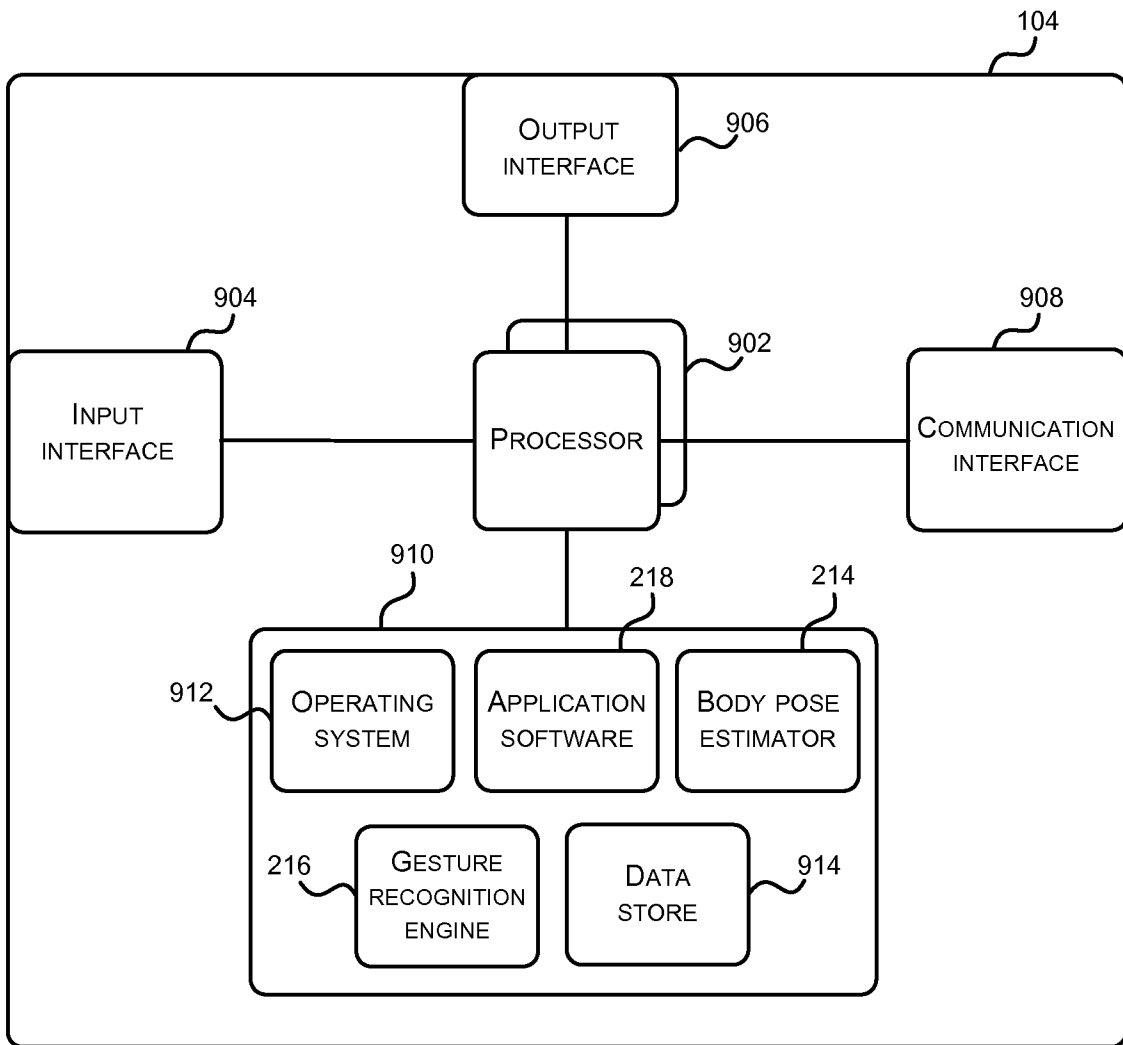
**FIG. 6**



**FIG. 7**



**FIG. 8**



**FIG. 9**