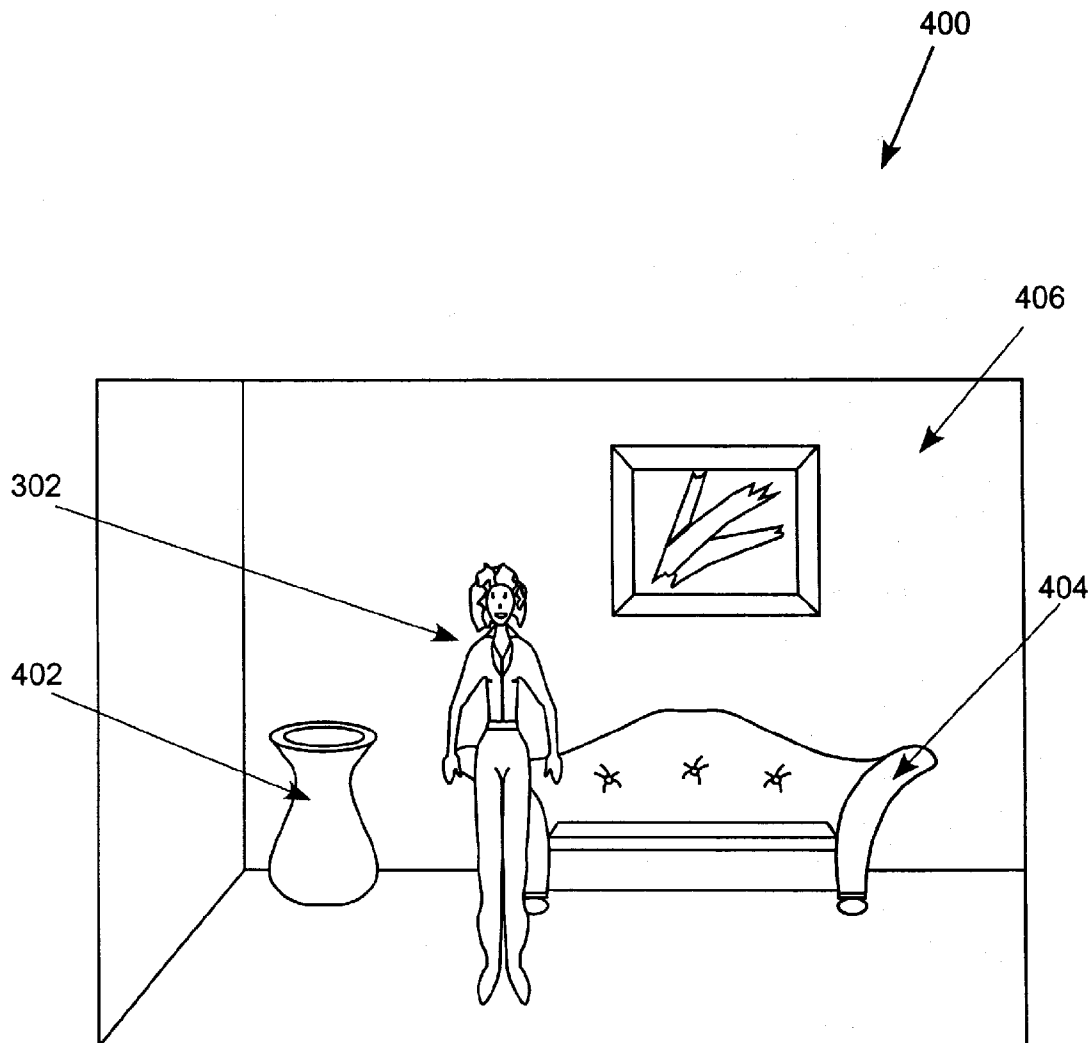




US 20120038637A1

(19) **United States**(12) **Patent Application Publication**
Marks(10) **Pub. No.: US 2012/0038637 A1**(43) **Pub. Date: Feb. 16, 2012**(54) **USER-DRIVEN THREE-DIMENSIONAL
INTERACTIVE GAMING ENVIRONMENT**(52) **U.S. Cl. 345/419**(75) **Inventor: Richard Marks, Foster City, CA
(US)**(73) **Assignee: Sony Computer Entertainment
Inc., Tokyo (JP)**(21) **Appl. No.: 13/282,369**(22) **Filed: Oct. 26, 2011****Related U.S. Application Data**(63) Continuation of application No. 10/448,614, filed on
May 29, 2003, now Pat. No. 8,072,470.**Publication Classification**(51) **Int. Cl. G06T 15/00 (2011.01)**(57) **ABSTRACT**

An invention is provided for affording a real-time three-dimensional interactive environment using a depth sensing device. The invention includes obtaining depth values indicating distances from one or more physical objects in a physical scene to a depth sensing device. The depth sensing device is configurable to be maintained at a particular depth range defined by a plane so that objects between the particular depth range and the depth sensing device are processed by the depth sensing device, wherein the particular depth range establishes active detection by the depth sensing device, as depth values of objects placed through the particular depth range and toward the depth sensing device are detected and depth values of objects placed beyond the particular depth range are not detected. The objects placed through the particular depth range are rendered and displayed in a virtual scene based on geometric characteristics of the object itself.



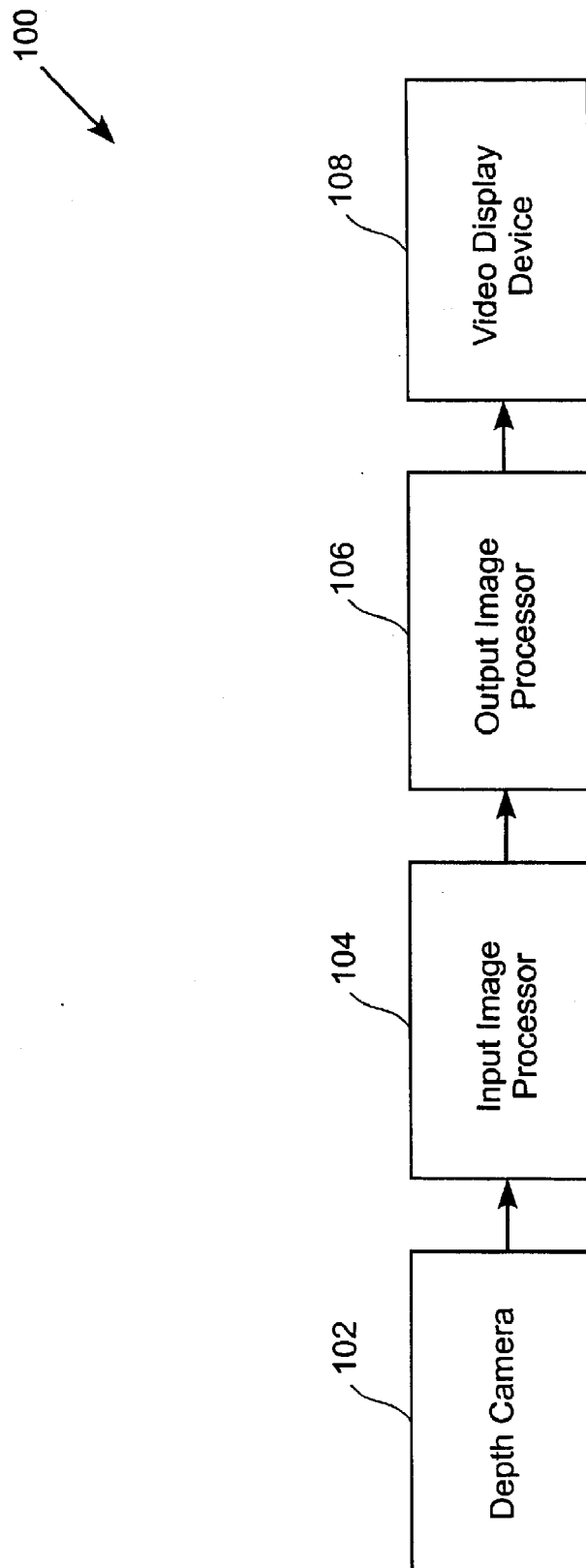
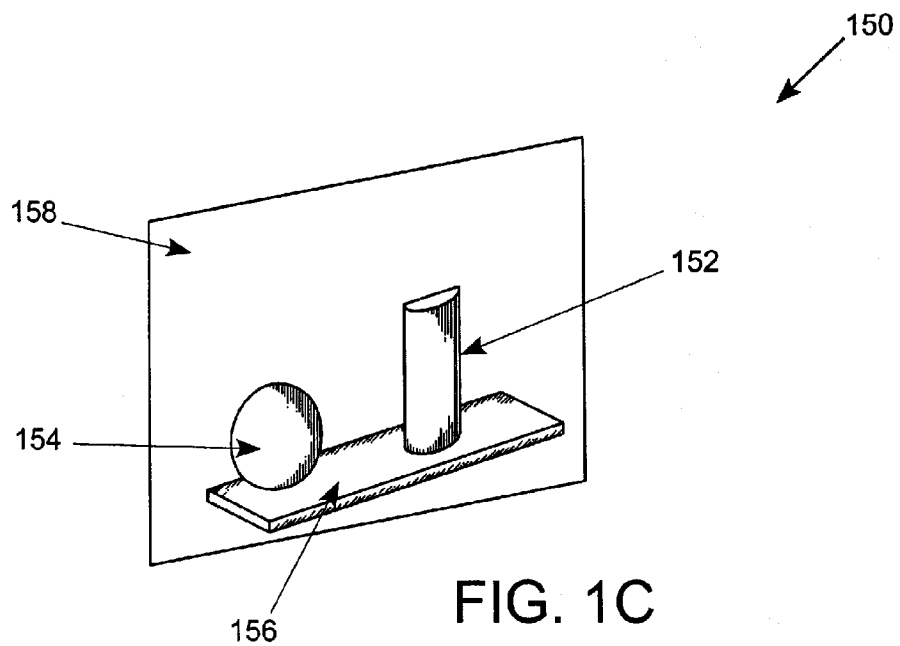
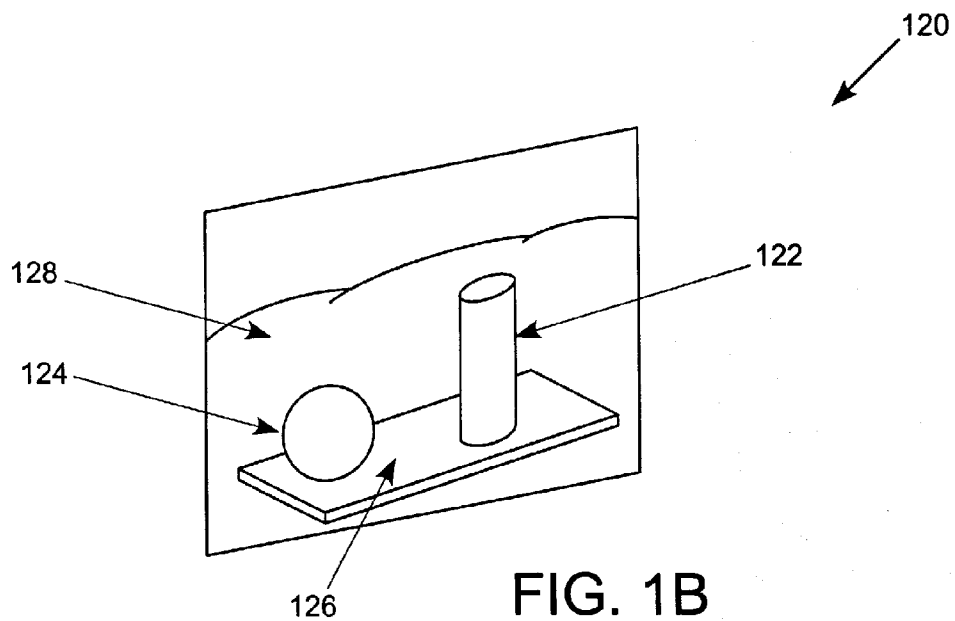


FIG. 1A



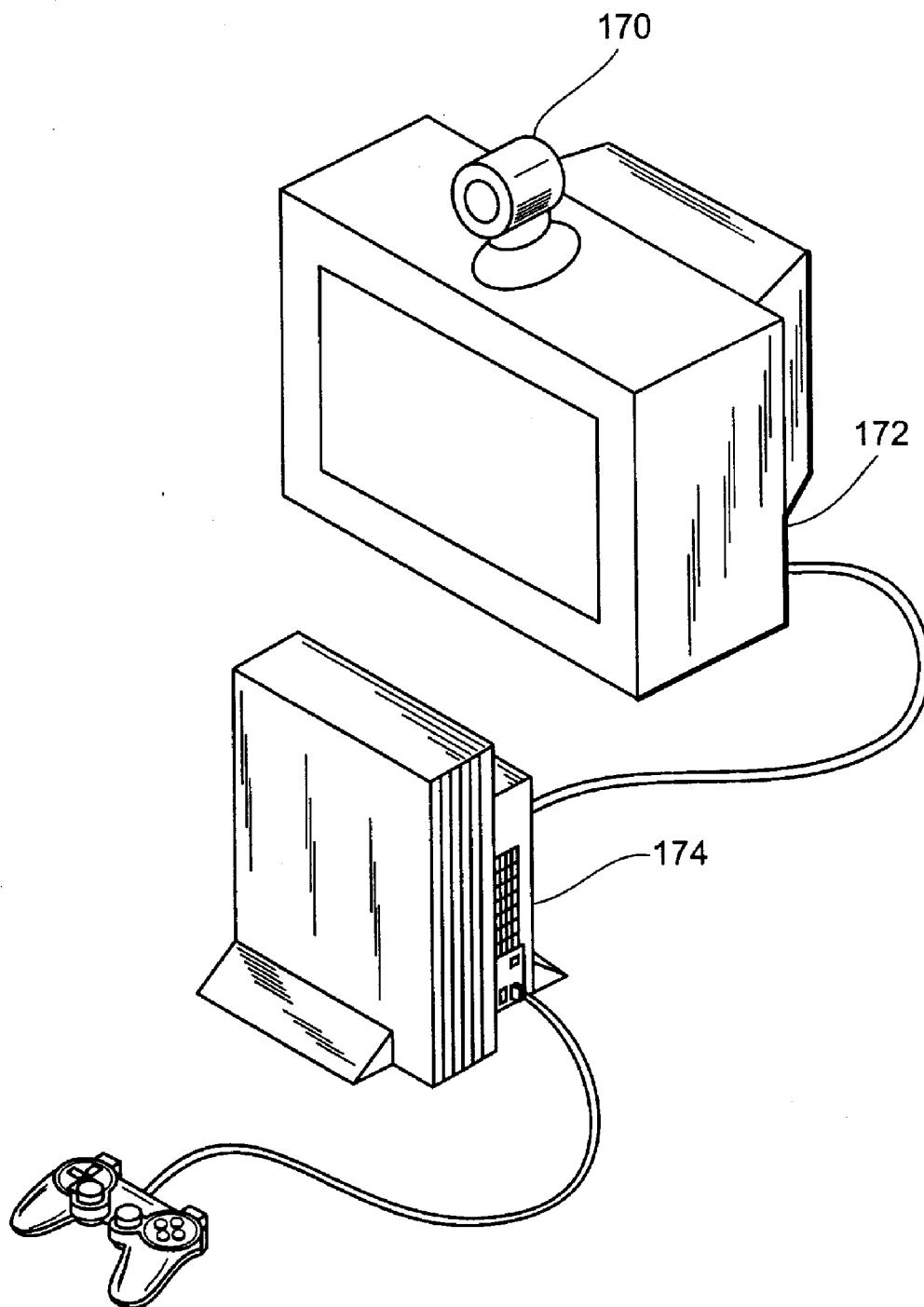


FIG. 1D

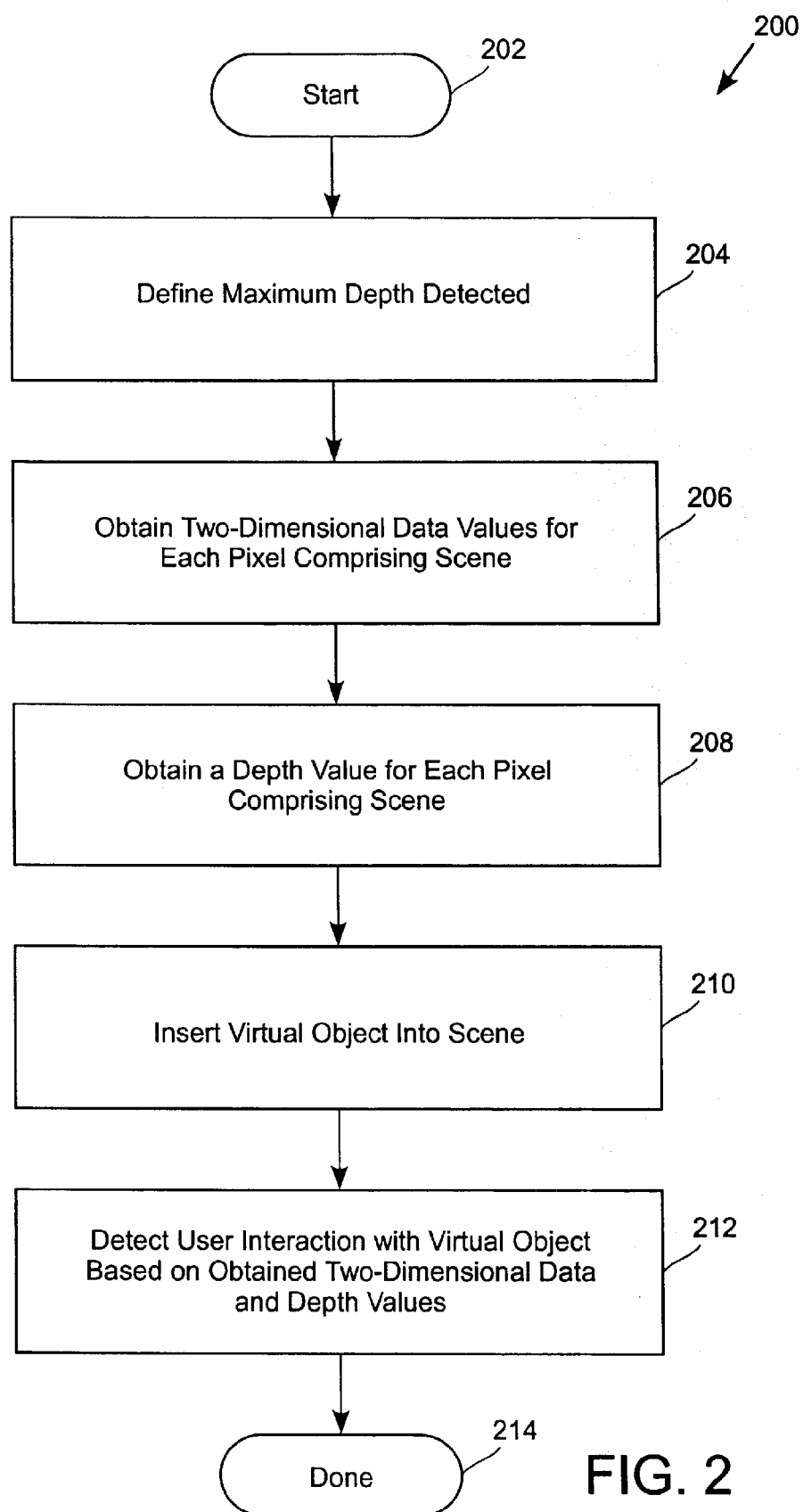


FIG. 2

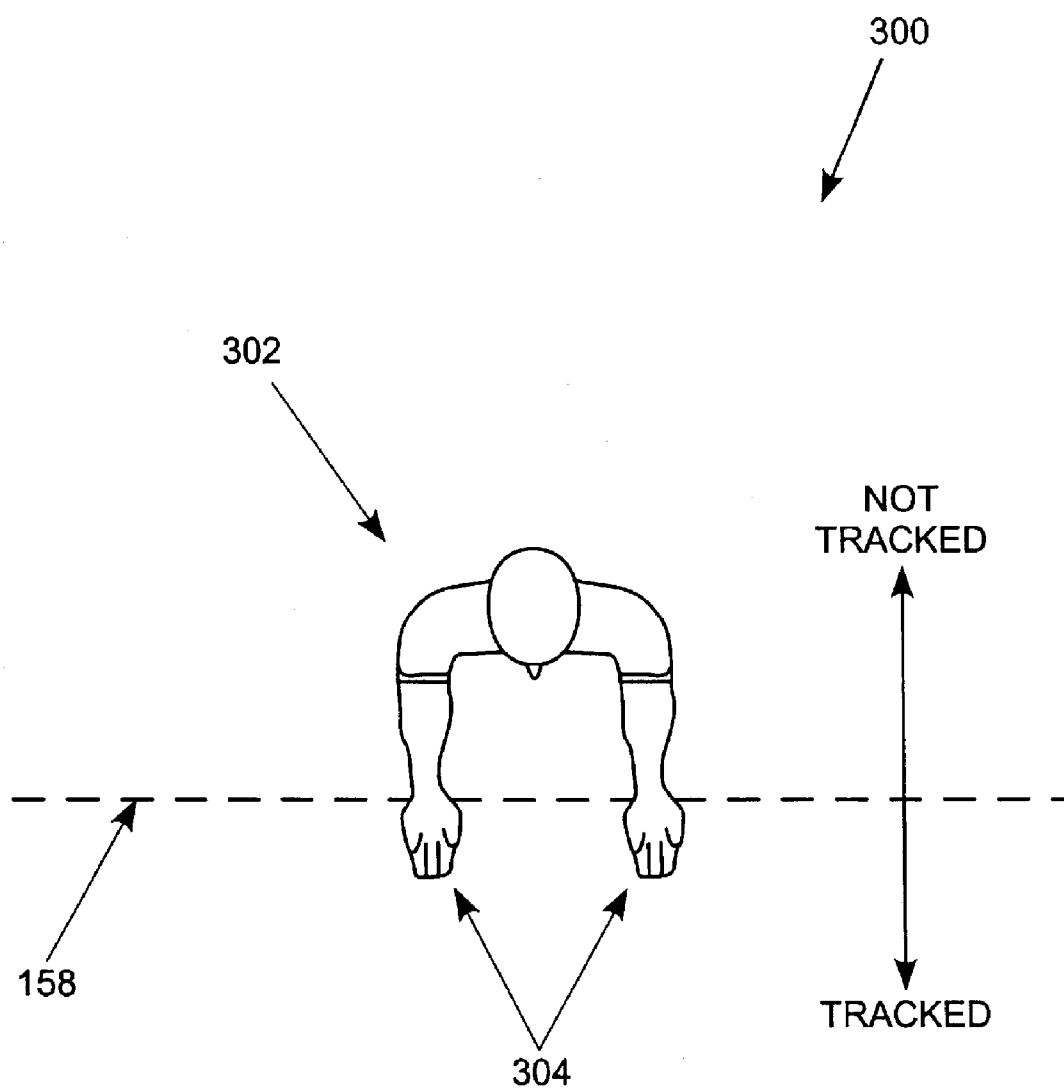


FIG. 3

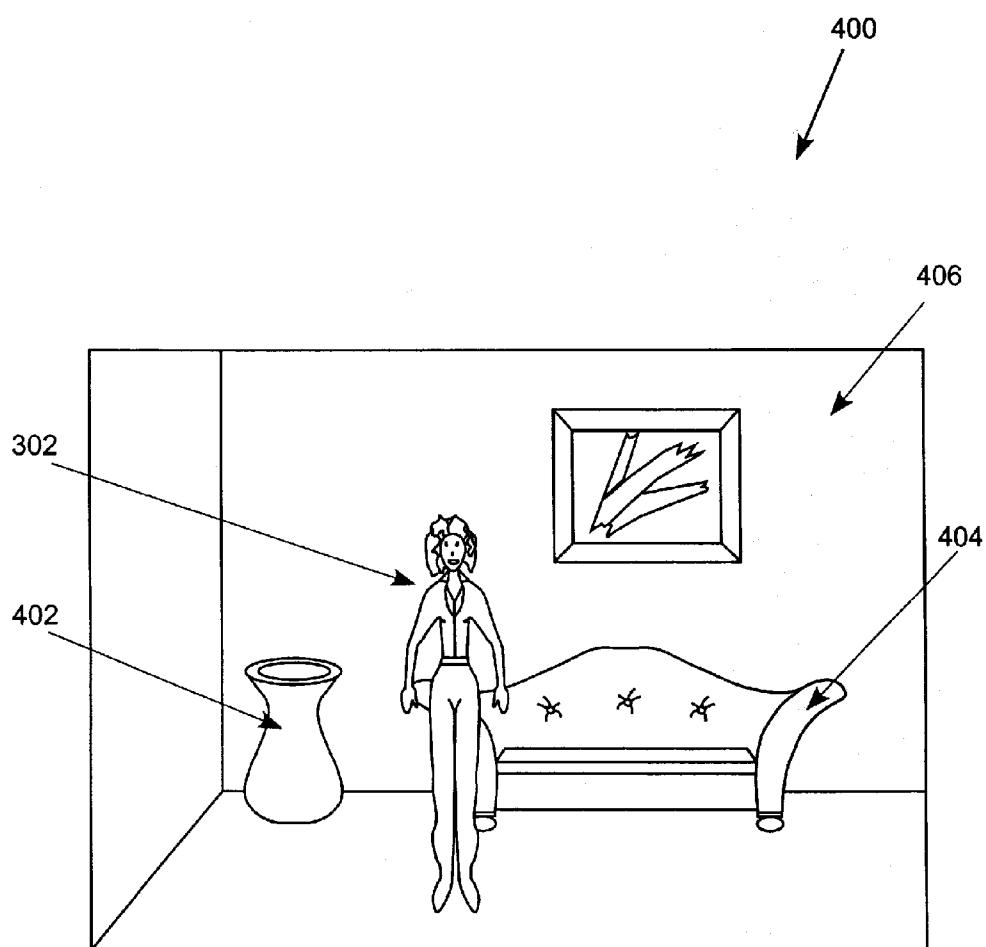


FIG. 4

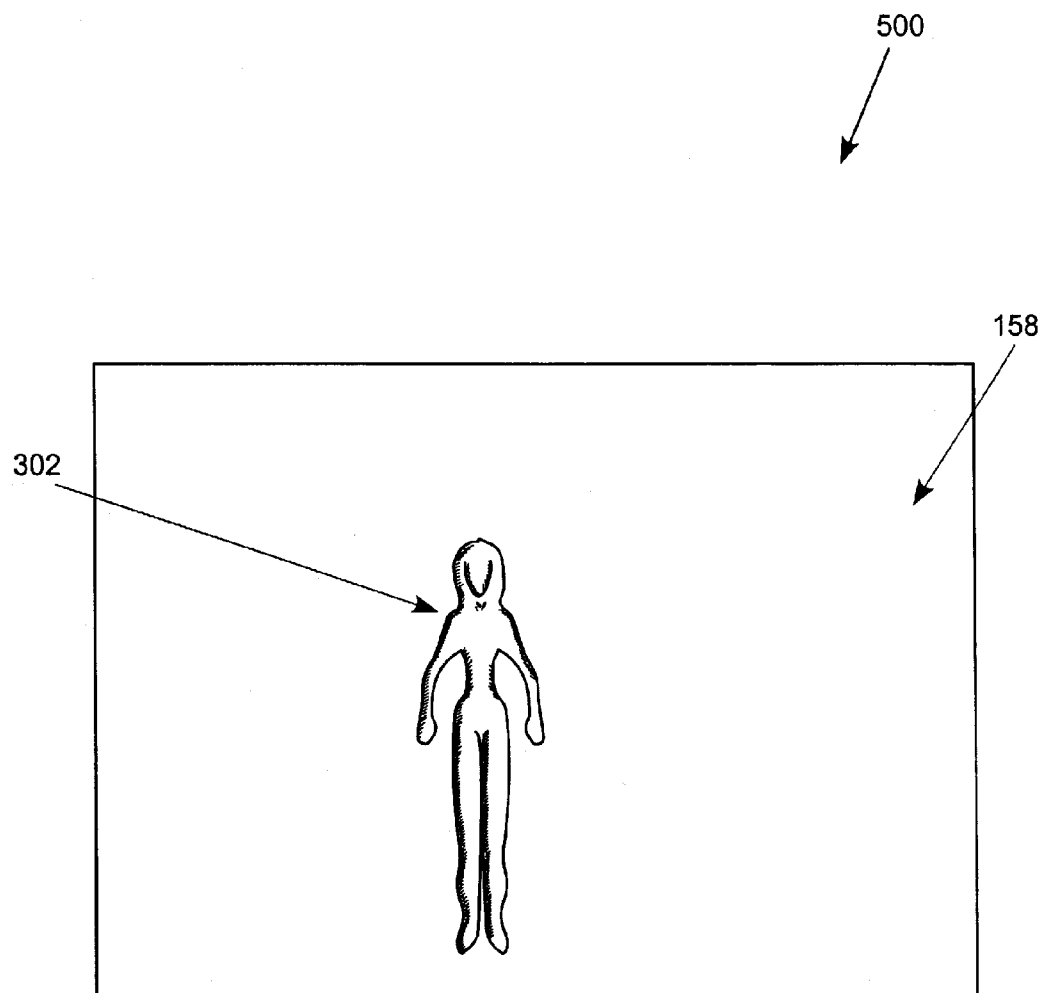


FIG. 5

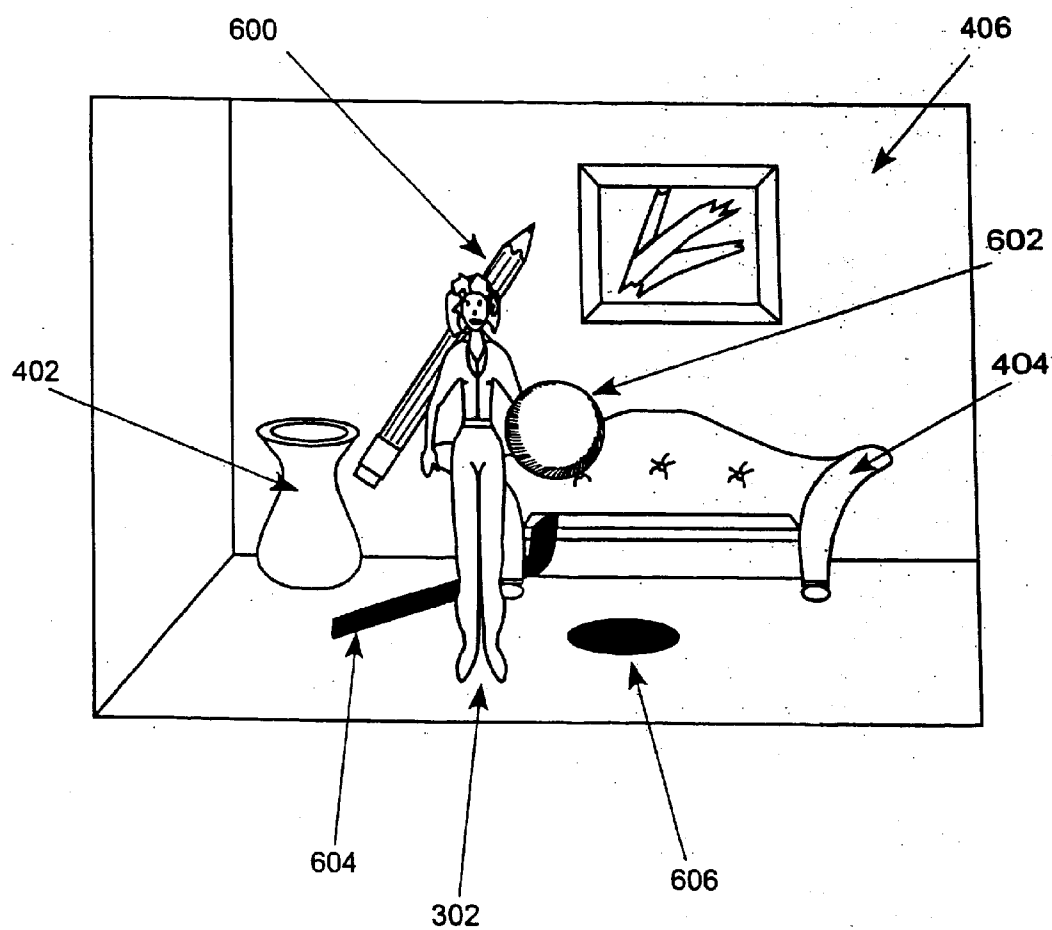


FIG. 6

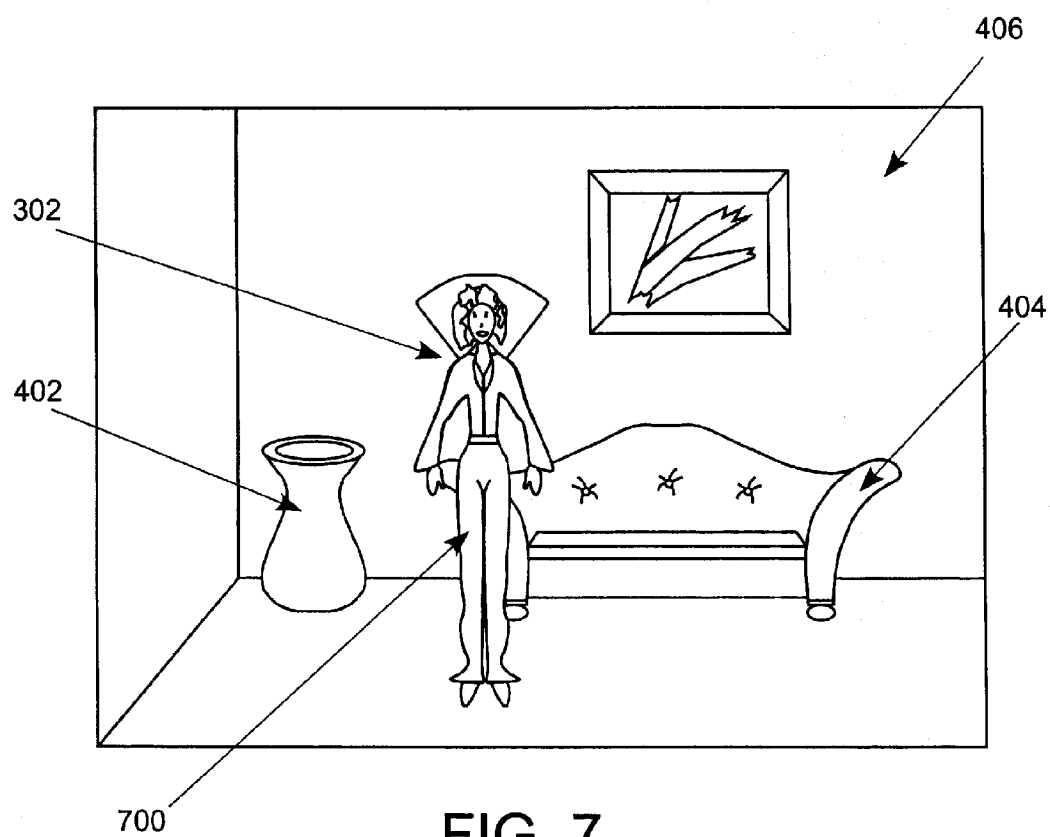
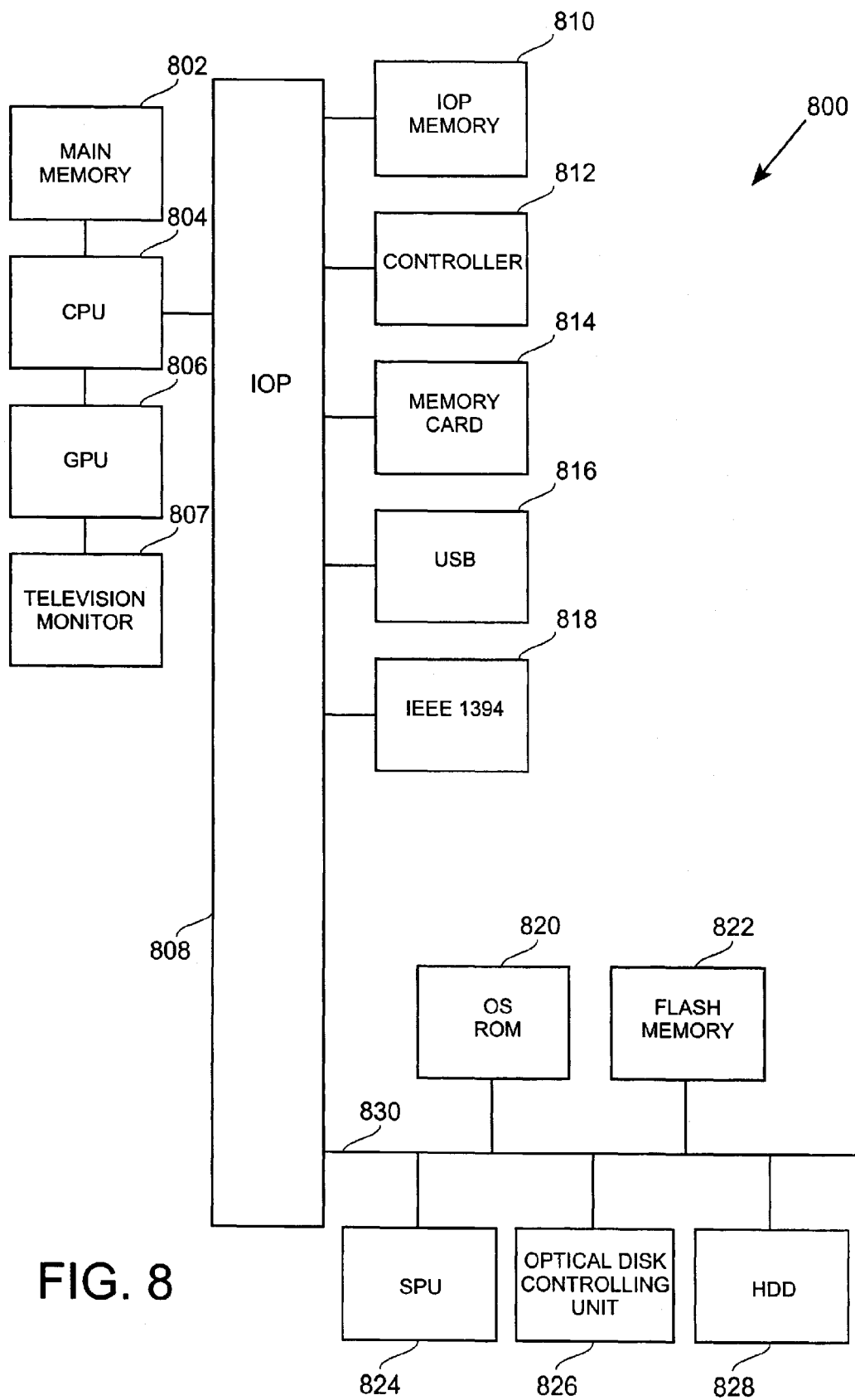


FIG. 7



USER-DRIVEN THREE-DIMENSIONAL INTERACTIVE GAMING ENVIRONMENT

CLAIM OF PRIORITY

[0001] This application claims priority as a continuation of U.S. patent application Ser. No. 10/448,614, filed May 29, 2003, and entitled "SYSTEM AND METHOD FOR PROVIDING A REAL-TIME THREE-DIMENSIONAL INTERACTIVE ENVIRONMENT," which is incorporated herein by reference.

CROSS REFERENCE TO RELATED APPLICATIONS

[0002] This application is related to U.S. patent application Ser. No. 10/365,120, filed Feb. 11, 2003, and entitled "Method and Apparatus for Real-Time Motion Capture," which is incorporated herein by reference.

BACKGROUND OF THE INVENTION

[0003] 1. Field of the Invention

[0004] This invention relates generally to video image processing, and more particularly to providing a real-time interactive computer environment using a three-dimensional camera.

[0005] 2. Description of the Related Art

[0006] With the increased processing capabilities of today's computer technology, new creative methods for interacting with computer systems have become available. For example, new on-line keyboards allow individuals to enter information without the need for a physical keyboard, and new game controllers with a plurality of joysticks and directional keys enhance the user's ability to interact with the computer system. In addition to hand held input devices, input devices employing video images are currently available that allow user control of objects on a graphical display such as a video monitor.

[0007] Such video input devices often are responsive to the movement or position of a user in the field of view of a video capture device. More recently, video image processing has been used to translate the movement of the user that has been captured as a sequence of video images into signals for game control. Prior art input systems include a video capture device that scans a field of view in which a system user stands. The captured video image is applied to a video digitizer that provides digital output to a processor that analyzes and processes the digital information received from the digitizer.

[0008] Based upon the position or movement of the participant in the field of view, the processor produces signals that are used by the graphics generating system to move objects on the display. Although the operation or output of the devices or graphical displays can thereby be affected by the position or movement of the participant, the computer processing time required is frequently very extensive and complex, tending to require substantial computer and/or time resources.

[0009] In addition, known devices and methods employing user video image data that are used to affect the movement of an object on a graphical display are typically characterized by significant encumbrances upon the participant within the video camera field of view. Such systems may include additional equipment that the participant is required to wear, such as arm coverings or gloves with integral, more easily detectable portions or colors, and/or visible light sources such as light emitting diodes. Unfortunately, such systems do not

allow for the ease-of-use, quick response, and simplicity needed to provide a user input device capable of meeting marketability requirements for consumer items such as might be required of video game controllers.

[0010] In view of the foregoing, there is a need for enhanced systems and methods that allow interaction in a three-dimensional environment. The methods should allow user interaction without requiring additional equipment, such as arm coverings or gloves. In addition, the method should not require overly burdensome processing ability and should have the ability to function in real-time, thus providing the user with a natural computer interaction experience.

SUMMARY OF THE INVENTION

[0011] Broadly speaking, embodiments of the present invention fill these needs by providing a real-time three-dimensional interactive environment using a three-dimensional camera. Generally, embodiments of the present invention allow the user to interact with, and affect, computer-generated objects and environments that are combined visually with the user's actual physical environment. In one embodiment, a method is disclosed for providing a real-time three-dimensional interactive environment. The method includes obtaining two-dimensional data values for a plurality of pixels representing a physical scene, and obtaining a depth value for each pixel of the plurality of pixels using a depth sensing device. Each depth value indicates a distance from a physical object in the physical scene to the depth sensing device. At least one computer-generated virtual object is inserted into the scene, and an interaction between a physical object in the scene and the virtual object is detected based on coordinates of the virtual object and the obtained depth values. For example, the two-dimensional values for the plurality of pixels can be color values, and each depth value can indicate a distance from a physical object in the physical scene represented by the corresponding pixel to the sensing device. In one aspect, the interaction can be a collision between a physical object in the scene and the virtual object. In this aspect, the collision is detected when the virtual object and a physical object occupy a same three-dimensional space based on three-dimensional coordinates of the virtual object and three-dimensional coordinates of the physical object. Optionally, an appearance of a physical object in the scene can be visually altered. For example, the physical object can be a user, and computer-generated clothing can be mapped to the user based on the depth values for pixels representing the user. In addition, a maximum depth range can be defined that indicates the farthest distance for which depth values will be obtained. In this aspect, depth values for the user may be detected only when the user is within a distance less than the maximum depth range to the sensing device.

[0012] A computer program embodied on a computer readable medium for providing a real-time three-dimensional interactive environment is disclosed in an additional embodiment of the present invention. The computer program includes program instructions that obtain two-dimensional data values for a plurality of pixels representing a physical scene. Also, program instructions are included that obtain a depth value for each pixel of the plurality of pixels using a depth sensing device. As above, each depth value indicates a distance from a physical object in the physical scene to the depth sensing device. Program instructions also are included that insert at least one virtual object into the scene, the virtual

object being computer-generated. Further, program instructions are included that detect an interaction between a physical object in the scene and the virtual object based on coordinates of the virtual object and the obtained depth values. As above, the two-dimensional values for the plurality of pixels are color values, and each depth value can indicate a distance from a physical object in the physical scene represented by the corresponding pixel to the sensing device. Optionally, program instructions can be included that define a maximum depth range that indicates the farthest distance for which depth values will be obtained.

[0013] In a further embodiment, a system is disclosed for providing a real-time three-dimensional interactive environment. The system includes a depth sensing device capable of obtaining two-dimensional data values for a plurality of pixels representing a physical scene. The depth sensing device is further capable of obtaining a depth value for each pixel of the plurality of pixels. As above, each depth value indicates a distance from a physical object in the physical scene to the depth sensing device. Also included in the system is logic that inserts at least one computer-generated virtual object into the scene. Further, the system includes logic that detects an interaction between a physical object in the scene and the virtual object based on coordinates of the virtual object and the obtained depth values. As above, the two-dimensional values for the plurality of pixels can be color values, and each depth value can indicate a distance from a physical object in the physical scene represented by the corresponding pixel to the sensing device. Optionally, the system can include logic that defines a maximum depth range, the maximum depth range indicating the farthest distance for which depth values will be obtained. In this aspect, logic can also be included that that detects depth values for a user only when the user is within a distance less than the maximum depth range to the sensing device.

[0014] A further method for providing a real-time three-dimensional interactive environment is disclosed in an additional embodiment of the present invention. As above, the method includes obtaining two-dimensional data values for a plurality of pixels representing a physical scene. Also as above, a depth value for each pixel of the plurality of pixels is obtained using a depth sensing device. Each depth value indicates a distance from a physical object in the physical scene to the depth sensing device. Based on the obtained two-dimensional data values and the obtained depth values, three-dimensional volume information is estimated for each physical object in the physical scene. In addition, computer-generated virtual objects having three-dimensional volume information for the virtual object can be inserted into the scene. In this manner, interactions between physical and virtual objects in the scene can be detected based on the coordinates of the three-dimensional volume information for the virtual object and the physical object. Other aspects and advantages of the invention will become apparent from the following detailed description, taken in conjunction with the accompanying drawings, illustrating by way of example the principles of the invention.

BRIEF DESCRIPTION OF THE DRAWINGS

[0015] The invention, together with further advantages thereof, may best be understood by reference to the following description taken in conjunction with the accompanying drawings in which:

[0016] FIG. 1A is a block diagram of an exemplary system for providing a real-time three-dimensional interactive environment, in accordance with an embodiment of the present invention;

[0017] FIG. 1B is an illustration showing a two-dimensional data captured using a typical depth camera;

[0018] FIG. 1C is an illustration showing depth data captured using a typical depth camera;

[0019] FIG. 1D illustrates an exemplary system environment for providing a real-time three-dimensional interactive environment, in accordance with an embodiment of the present invention;

[0020] FIG. 2 is a flowchart showing a method for providing a real-time three-dimensional interactive environment, in accordance with an embodiment of the present invention;

[0021] FIG. 3 is an illustration showing a top view of a user interacting with a maximum range plane, in accordance with an embodiment of the present invention;

[0022] FIG. 4 is an illustration showing two-dimensional data for an exemplary scene, in accordance with an embodiment of the present invention;

[0023] FIG. 5 illustrates z-values for the exemplary scene of FIG. 4, in accordance with an embodiment of the present invention;

[0024] FIG. 6 is an illustration showing computer generated virtual objects inserted into a scene, in accordance with an embodiment of the present invention;

[0025] FIG. 7 is an illustration showing computer-generated changes to the physical objects within the room, in accordance with an embodiment of the present invention; and

[0026] FIG. 8 is a block diagram of a computer processing system for providing a three-dimensional interactive environment, in accordance with an embodiment of the present invention.

DETAILED DESCRIPTION OF THE PREFERRED EMBODIMENTS

[0027] An invention is disclosed for providing a real-time three-dimensional interactive environment using a three-dimensional camera. Generally, embodiments of the present invention allow the user to interactive with, and affect, computer objects and environments that are combined visually with the user's actual physical environment. Through the use of a three-dimensional camera, three-dimensional images can be obtained in real-time. These three-dimensional images are utilized to place digital objects within the user's environment, track the user's movement, and accurately detect when the user interacts with the digital objects. In the following description, numerous specific details are set forth in order to provide a thorough understanding of the present invention. It will be apparent, however, to one skilled in the art that the present invention may be practiced without some or all of these specific details. In other instances, well known process steps have not been described in detail in order not to unnecessarily obscure the present invention.

[0028] In the following description, the terms "depth camera" and "three-dimensional camera" will refer to any camera that is capable of obtaining distance or depth information as well as two-dimensional pixel information. For example, a depth camera can utilize controlled infrared lighting to obtain distance information. Another exemplary depth camera can be a stereo camera pair, which triangulates distance information using two standard cameras. Similarly, the term "depth

sensing device” will refer to any type of device that is capable of obtaining distance information as well as two-dimensional pixel information.

[0029] Recent advances in three-dimensional imagery have opened the door for increased possibilities in real-time interactive computer animation. In particular, new “depth cameras” provide the ability to capture and map the third-dimension in addition to normal two-dimensional video imagery. With the new depth data, embodiments of the present invention allow the placement of computer-generated objects in various positions within a video scene in real-time, including behind other objects.

[0030] Moreover, embodiments of the present invention provide real-time interactive gaming experiences for users. For example, users can interact with various computer-generated objects in real-time. Furthermore, video scenes can be altered in real-time to enhance the user’s game experience. For example, computer generated costumes can be inserted over the user’s clothing, and computer generated light sources can be utilized to project virtual shadows within a video scene. Hence, using the embodiments of the present invention and a depth camera, users can experience an interactive game environment within their own living room.

[0031] FIG. 1A is a block diagram of an exemplary system 100 for providing a real-time three-dimensional interactive environment, in accordance with an embodiment of the present invention. As shown in FIG. 1A, the system 100 includes a depth camera 102, an input image processor 104, an output image processor 106, and a video display device 108.

[0032] As mentioned above, the depth camera 102 provides the ability to capture and map the third-dimension in addition to normal two-dimensional video imagery. FIGS. 1B and 1C illustrated the images generated by a typical depth camera 102. In particular, FIG. 1B is an illustration showing two-dimensional data 120 captured using a typical depth camera. Similar to normal cameras, a depth camera captures two-dimensional data for a plurality of pixels that comprise the video image. These values are color values for the pixels, generally red, green, and blue (RGB) values for each pixel. In this manner, objects captured by the camera appear as two-dimension objects on a monitor. For example, in FIG. 1B, the exemplary scene includes a cylinder object 122 and a sphere object 124 disposed on a table 126, which is situated among hills 128, illustrated in the background.

[0033] However, unlike a conventional camera, a depth camera also captures depth values for the scene. FIG. 1C is an illustration showing depth data 150 captured using a typical depth camera. As illustrated in FIG. 1B, the depth camera captures the x and y components of a scene using RGB values for each pixel in the scene. However, as shown in FIG. 1C, the depth camera also captures the z-components of the scene, which represent the depth values for the scene. Since the depth values correspond to the z-axis, the depth values are often referred to as z-values.

[0034] In operation, a z-value is captured for each pixel of the scene. Each z-value represents a distance from the camera to a particular object in the scene corresponding to the related pixel. For example, in FIG. 1C, z-values are illustrated for the cylinder object 152, the sphere object 154, and part of the table 156. In addition, a maximum detection range is defined beyond which depth values will not be detected. For example, in FIG. 1C the maximum depth range 158 appears as vertical plane wherein all pixels are given the same depth value. As

will be described in greater detail below, this maximum range plane can be utilized by the embodiments of the present invention to provide user defined object tracking. Thus, using a depth camera, each object can be tracked in three dimensions. As a result, a computer system of the embodiments of the present invention can utilize the z-values, along with the two-dimensional pixel data, to create an enhanced three-dimensional interactive environment for the user.

[0035] Referring back to FIG. 1A, the input image processor 104 translates the captured video images and depth data into signals that are delivered to an output image processor. The output image processor 106 is programmed to effect movement and status of virtual objects on the video display device 108 in response to signals received from the input image processor 104.

[0036] These and additional aspects of the present invention may be implemented by one or more processors which execute software instructions. According to one embodiment of the present invention, a single processor executes both input image processing and output image processing. However, as shown in the figures and for ease of description, the processing operations are shown as being divided between an input image processor 104 and an output image processor 106. It should be noted that the invention is in no way to be interpreted as limited to any special processor configuration, such as more than one processor. The multiple processing blocks shown in FIG. 1A and the other Figures are shown only for convenience of description.

[0037] FIG. 1D illustrates an exemplary system environment for providing a real-time three-dimensional interactive environment, in accordance with an embodiment of the present invention. The system environment includes, depth camera 170, video display device 172, and console 174 having processor functionality, such as a video game machine. Generally, a user and their environment, such as a living room, are located within the field of view of the depth camera 170. The processing system 174 can be implemented by an entertainment system, such as a Sony® Playstation™ II or Sony® Playstation™ I type of processing and computer entertainment system. It should be noted, however, that processing system 174 can be implemented in other types of computer systems, such as personal computers, workstations, laptop computers, wireless computing devices, or any other type of computing device that is capable of receiving and processing graphical image data.

[0038] FIG. 2 is a flowchart showing a method 200 for providing a real-time three-dimensional interactive environment, in accordance with an embodiment of the present invention. In an initial operation 202, preprocess operations are performed. Preprocess operations can include defining three-dimensional objects, adjusting a depth camera for optimum performance, and other preprocess operations that will be apparent to those skilled in the art after a careful reading of the present disclosure.

[0039] In operation 204, a maximum depth range is defined. As described above, a maximum depth range is defined beyond which depth values will not be detected. Typically, the maximum depth range appears as vertical plane wherein all pixels are given the same depth value. This maximum range plane can be utilized by the embodiments of the present invention to provide user defined object tracking, as illustrated in FIG. 3.

[0040] FIG. 3 is an illustration showing a top view 300 of a user 302 interacting with a maximum depth range plane 158,

in accordance with an embodiment of the present invention. As shown in FIG. 3, a maximum depth range plane 158, which defines tracking distance, is defined. Objects in front of the maximum depth range plane 158 are tracked, while objects behind the maximum depth range plane 158 are not tracked. In this manner, the user 302 can determine when to interact with the system by allowing part of the user's body, or an object, to cross the maximum depth range plane 158.

[0041] For example, when the user 302 of FIG. 3 places their hands 304 in front of the maximum depth range plane 158, the system detects and tracks their hands 304. In this manner, the user 302 controls when to interact with the system, and the system can avoid any confusing information caused, for example, by unexpected body movement. In addition, motion confusion caused by other people moving behind the user, or for example, a family pet, can be avoided.

[0042] For example, in one embodiment of the present invention, the user 302 is allowed to drag and drop objects on the screen by gesturing with their hands across the maximum depth range plane 158. In this embodiment, a user can extend their hand 304 or other object across the maximum depth range plane 158 to initiate interaction with objects on a screen. The movement of the user's hand is then tracked using the depth data provided by the depth camera. Tracking is then terminated when the user retracts their hand behind the maximum depth range plane 158. During tracking, objects encountered by the user's hand movement can be moved and manipulated, as described in greater detail subsequently.

[0043] Referring back to FIG. 2, two-dimensional data values are obtained for each pixel comprising the scene, in operation 206. As mentioned above, a depth camera can capture two-dimensional data for a plurality of pixels that comprise a video image. These values are color values for the pixels, and generally red, green, and blue (RGB) values for each pixel. In this manner, objects captured by the camera appear as two-dimension objects on a monitor.

[0044] For example, FIG. 4 is an illustration showing two-dimensional data 400 for an exemplary scene, in accordance with an embodiment of the present invention. The exemplary scene on FIG. 4 illustrates a user 302 in the living room of their home. However, it should be noted that embodiments of the present invention can be utilized in any location, as desired by the user. As can be appreciated, various physical objects are located in this environment. For example, in FIG. 4 there is a vase 402 and sofa 404, as well as a picture on the back wall 406 of the room. As will be discussed in greater detail subsequently these exemplary objects will illustrate properties of the embodiments of the present invention.

[0045] Generally, the user 302 positions the depth camera in a suitable position in front of them. In addition, various adjustments can be made to the camera angle, aperture setting, and other settings that will be apparent to those skilled in the art after a careful reading of the present disclosure. The camera then captures video data for the scene, generally comprising color values for the pixels comprising the scene.

[0046] Referring back to FIG. 2, depth values are obtained for each pixel comprising the scene, in operation 208. In addition to two-dimensional data capture, a depth camera also captures depth values for the scene. As discussed above, the depth camera captures the x and y components of a scene using RGB values for each pixel in the scene. However, the depth camera also captures the z-components of the scene, which represent the depth values for the scene.

[0047] Thus, in operation 208, a z-value is captured for each pixel of the scene. Each z-value represents a distance from the camera to a particular object in the scene corresponding to the related pixel. For example, FIG. 5 illustrates z-values for the exemplary scene of FIG. 4, in accordance with an embodiment of the present invention. The z-values are included for the user 302, however, in the example of FIG. 5 the maximum depth range plane 158 has been defined just behind the user 302. Thus, excluding depth values for the other objects in the scene, including the vase, sofa, and back wall. However, it should be noted that the maximum depth range plane 158 can be defined at any distance. Thus, the maximum depth range plane 158 can be defined farther back in the exemplary scene to include the vase, sofa, and back wall.

[0048] In this manner, the position and movement of the user 302 can be tracked. Moreover, using the depth information, the user 302 can be tracked in three dimensions, thus allowing for realistic placement of objects within the scene. Furthermore, using the three-dimensional data allows users to interact with a virtual environment in a realistic manner thus enhancing the user's 302 experience.

[0049] In addition, one embodiment of the present invention can construct complete 3D volume information for objects in the scene using the z-values. In general, a depth camera does not itself provide full volume information. That is, the depth camera provides z-values for pixels of object surfaces that are visible to the camera. Hence, the z-values for the surfaces, such as the user's 302 back are not provided by the depth camera. Thus, one embodiment of the present invention estimates the complete volume information for objects in the scene to create complete 3D volumes, which can later be intersected with other 3D objects to determine collisions or for measuring distances between the objects.

[0050] For example, in FIG. 5, one embodiment of the present invention estimates the "back" z-values of the user 302, which are not visible to the depth camera. In one embodiment, a pre-generated model is utilized to estimate the volume of a particular object. Although the pre-generated model may not be absolutely accurate, and good estimation of volume can be achieved. For example, when estimating a volume of a particular person, the depth of the person can be estimated to be equal to the width of the person. When the model is accurate, embodiments of the present invention orient the model to match the orientation the actual object, and then utilize the model to estimate the volume of the object. For example, when the object is a couch, embodiments of the present invention orient a model couch to match the couch object, then determine the volume of the couch object based on the couch size and the model data.

[0051] In this manner, a complete 3D volume of the user 302 can be constructed, which can later be utilized to interact with computer generated virtual objects. In this manner, embodiments of the present invention can process both real and virtual objects in a single consistent manner.

[0052] Referring back to FIG. 2, in operation 210, virtual objects are inserted into the scene. With the new depth data obtained in operation 208, embodiments of the present invention allow the placement of computer-generated objects in various positions within a video scene in real-time, including behind other objects in operation 210.

[0053] In this manner, embodiments of the present invention provide real-time interactive gaming experiences for users. For example, users can interact with various computer-

generated objects in real-time. Furthermore, video scenes can be altered in real-time to enhance the user's game experience. For example, computer generated costumes can be inserted over the user's clothing, and computer generated light sources can be utilized to project virtual shadows within a video scene. Hence, using the embodiments of the present invention and a depth camera, user's can experience an interactive game environment within their own living room.

[0054] For example, FIG. 6 is an illustration showing computer-generated virtual objects inserted into a scene, in accordance with an embodiment of the present invention. As in FIG. 4, the scene includes a vase 402 and sofa 404, as well as a picture on the back wall 406 of the room. The depth camera captures these physical objects using two-dimensional pixel data, as described previously. In addition, also described above, the depth camera captures depth data, in this example for the user 302. Using the depth data, embodiments of the present invention insert virtual objects into the scene. For example, in FIG. 6 two virtual objects 600 and 602 were added to the scene. As illustrated, the virtual objects 600 and 602 can be inserted into the scene in a realistic manner because of the added depth information available.

[0055] That is, the depth data obtained in operation 208 can be utilized to determine the exact position of the user 302 in three-dimensional space. As a result, the virtual "pencil" object 600 can be positioned, altered, and animated to appear to be "behind" the user 302. Similarly, the virtual sphere 602 can be positioned, altered, and animated to appear, for example, in "front" of the user 302. Moreover, by extending the maximum depth range to approximately the position of the back wall 406, the inserted virtual objects can appear to interact with other objects in the user's room. In addition, one embodiment of the present invention inserts a virtual light source in the scene to cast "shadows" 604 and 606 from the virtual objects, which further increase the realism of the virtual scene. Since, the exact three-dimensional position of the floor and sofa 404 can be determined from the depth data, the computer-generated shadow 606 of the virtual sphere 602 can appear to be cast on the floor and the computer-generated shadow 604 of the virtual pencil 602 can appear to be cast on the sofa 404 and on the floor. Virtual objects can also include computer-generated changes to the physical objects within the room, as illustrated in FIG. 7.

[0056] FIG. 7 is an illustration showing computer-generated changes to the physical objects within the room, in accordance with an embodiment of the present invention. As in FIG. 4, the scene includes a vase 402 and sofa 404, as well as a picture on the back wall 406 of the room. The depth camera captures these physical objects using two-dimensional pixel data, as described previously. In addition, also described above, the depth camera captures depth data, in this example for the user 302. Using the depth data, embodiments of the present invention can visually alter physical objects in the scene. For example, in FIG. 7, a computer-generated costume 700 has been inserted into the scene over the user's clothing. Since the z-values obtained from the depth camera allow the system to track the user's movement, the computer-generated costume 700 can be animated to move with the user, creating the appearance that the user 302 is "wearing" the computer-generated costume.

[0057] Referring back to FIG. 2, the user's interactions with the virtual objects are detected based on the obtained two-dimensional data and the depth values. As is well known in the art, computer-generated three-dimensional objects are

located in a virtual three-dimensional space and processed, often using matrixes, to generate a two-dimensional projection of the three-dimensional scene, typically viewed using a monitor or television. In one embodiment of the present invention, the virtual three-dimensional space is configured to coincide with the physical space of the user. For example, referring to FIG. 6, the virtual three-dimensional space can be configured to coincide with the living room of the user 302. In this manner, embodiments of the present invention can detect when objects, both virtual and physical, occupy the same three-dimensional space.

[0058] Thus, embodiments of the present invention can, utilizing the z-values from the depth camera, allow the user 302 to interact with the virtual objects. For example, a user can swing at the virtual sphere 602 and the system can detect when the user's 302 hand, for example, occupies the same space as the virtual sphere 602, indicating a collision. Thereafter, an appropriate response to the collision can be generated, for example, the virtual sphere 602 can be made the "virtually fly" across the room.

[0059] Post process operations are performed in operation 214. Post process operations can include saving locations of virtual objects on a computer storage medium, loading of saved virtual objects from the computer storage medium, and other post process operation that will be apparent to those skilled in the art after a careful reading of the present disclosure.

[0060] In one embodiment, the three-dimensional interactive system and methods of the embodiments of the present invention are implemented using a computer processing system illustrated by the block diagram of FIG. 8. The processing system may represent a computer-based entertainment system embodiment that includes a central processing unit ("CPU") 804 coupled to a main memory 802 and graphical processing unit ("GPU") 806. The CPU 804 is also coupled to an Input/Output Processor ("IOP") Bus 808. In one embodiment, the GPU 806 includes an internal buffer for fast processing of pixel based graphical data. Additionally, the GPU can include an output processing portion or functionality to convert the image data processed into standard television signals, for example NTSC or PAL, for transmission to a television monitor 807 connected external to the entertainment system 800 or elements thereof. Alternatively, data output signals can be provided to a display device other than a television monitor, such as a computer monitor, LCD (Liquid Crystal Display) device, or other type of display device.

[0061] The IOP bus 808 couples the CPU 804 to various input/output devices and other busses or device. IOP bus 808 is connected to input/output processor memory 810, a controller 812, a memory card 814, a Universal Serial Bus (USB) port 816, an IEEE1394 (also known as a Firewire interface) port, and bus 830. Bus 830 couples several other system components to CPU 804, including operating system ("OS") ROM 820, flash memory 822, a sound processing unit ("SPU") 824, an optical disc controlling unit 826, and a hard disk drive ("HDD") 828. In one aspect of this embodiment, the video capture device can be directly connected to the IOP bus 808 for transmission therethrough to the CPU 804; there, data from the video capture device can be used to change or update the values used to generate the graphics images in the GPU 806.

[0062] Programs or computer instructions embodying aspects of the present invention can be provided by several different methods. For example, the user input method for

interaction with graphical images can be provided in the form of a program stored in HDD **828**, flash memory **822**, OS ROM **820**, or on a memory card **812**. Alternatively, the program can be downloaded to the processing unit **800** through one or more input ports coupled to the CPU **804**. The program modules defining the input method can be provided with the game or application program that is executed by the CPU **804** and displayed on display device **807** or they may be provided separately from the application program, such as for execution from local main memory **802**.

[0063] Embodiments of the present invention also contemplate distributed image processing configurations. For example, the invention is not limited to the captured image and display image processing taking place in one or even two locations, such as in the CPU or in the CPU and one other element. For example, the input image processing can just as readily take place in an associated CPU, processor or device that can perform processing; essentially all of image processing can be distributed throughout the interconnected system. Thus, the present invention is not limited to any specific image processing hardware circuitry and/or software; it is also not limited to any specific combination of general hardware circuitry and/or software, nor to any particular source for the instructions executed by processing components.

[0064] Although the foregoing invention has been described in some detail for purposes of clarity of understanding, it will be apparent that certain changes and modifications may be practiced within the scope of the appended claims. Accordingly, the present embodiments are to be considered as illustrative and not restrictive, and the invention is not to be limited to the details given herein, but may be modified within the scope and equivalents of the appended claims.

What is claimed is:

1. A computer implemented method having access to memory, the method providing a real-time three-dimensional interactive environment, comprising the operations of:

obtaining depth values indicating distances from one or more physical objects in a physical scene to a depth sensing device, the depth sensing device configurable to be maintained at a particular depth range defined by a plane so that objects between the particular depth range and the depth sensing device are processed by the depth sensing device, wherein the particular depth range establishes active detection by the depth sensing device, as depth values of objects placed through the particular depth range and toward the depth sensing device are detected and depth values of objects placed beyond the particular depth range are not detected, and the objects placed through the particular depth range are rendered and displayed in a virtual scene based on geometric characteristics of the object itself.

2. A method as recited in claim 1, further comprising, initiating tracking of the objects when the objects are placed through the particular depth range and toward the depth sensing device, and terminating tracking of the objects when the objects are placed beyond the particular depth range.

3. A method as recited in claim 1, further comprising, inserting at least one virtual object into the virtual scene after obtaining the depth values, the virtual object being computer-generated and configured to be inserted within and beyond the particular depth range.

4. A method as recited in claim 3, further comprising, detecting an interaction between only objects placed through the particular depth range and the virtual object based on coordinates of the virtual object and the obtained depth values of the objects placed through the particular depth range.

5. A method as recited in claim 1, wherein the depth sensing device is a depth camera using controlled infrared lighting.

6. A method as recited in claim 1, further comprising, estimating three-dimensional volume information for each physical object within the particular depth range based on the obtained depth values.

7. A computer program embodied on a computer readable medium for providing a real-time three-dimensional interactive environment, comprising:

program instructions that obtain depth values indicating distances from one or more physical objects in a physical scene to a depth sensing device, the depth sensing device configurable to be maintained at a particular depth range defined by a plane so that objects between the particular depth range and the depth sensing device are processed by the depth sensing device, wherein the particular depth range establishes active detection by the depth sensing device, as depth values of objects placed through the particular depth range and toward the depth sensing device are detected and depth values of objects placed beyond the particular depth range are not detected, and the objects placed through the particular depth range are rendered and displayed in a virtual scene based on geometric characteristics of the object itself.

8. A computer program as recited in claim 7, further comprising,

program instructions that initiate tracking of the objects when the objects are placed through the particular depth range and toward the depth sensing device, and terminate tracking of the objects when the objects are placed beyond the particular depth range.

9. A computer program as recited in claim 7, further comprising,

program instructions that insert at least one virtual object into the virtual scene after obtaining the depth values, the virtual object being computer-generated and configured to be inserted within and beyond the particular depth range.

10. A computer program as recited in claim 9, further comprising,

program instructions that detect an interaction between only objects placed through the particular depth range and the virtual object based on coordinates of the virtual object and the obtained depth values of the objects placed through the particular depth range.

11. A computer program as recited in claim 7, wherein the depth sensing device is a depth camera using controlled infrared lighting.

12. A computer program as recited in claim 7, further comprising, program instructions that estimate three-dimensional volume information for each physical object within the particular depth range based on the obtained depth values.

13. A system for providing a real-time three-dimensional interactive environment, comprising:

a depth sensing device capable of obtaining depth values indicating distances from one or more physical objects in a physical scene to a depth sensing device, the depth

sensing device configurable to be maintained at a particular depth range defined by a plane so that objects between the particular depth range and the depth sensing device are processed by the depth sensing device, wherein the particular depth range establishes active detection by the depth sensing device, as depth values of objects placed through the particular depth range and toward the depth sensing device are detected and depth values of objects placed beyond the particular depth range are not detected; and

a console having logic configured to render and display the objects placed through the particular depth range in a virtual scene based on geometric characteristics of the object itself.

14. A system as recited in claim **13**, further comprising, logic that initiates tracking of the objects when the objects are placed through the particular depth range and toward the depth sensing device, and terminates tracking of the objects when the objects are placed beyond the particular depth range.

15. A system as recited in claim **13**, further comprising, logic that inserts at least one virtual object into the virtual scene after obtaining the depth values, the virtual object being computer-generated and configured to be inserted within and beyond the particular depth range.

16. A system as recited in claim **15**, further comprising, logic that detects an interaction between only objects placed through the particular depth range and the virtual object based on coordinates of the virtual object and the obtained depth values of the objects placed through the particular depth range.

17. A system as recited in claim **13**, wherein the depth sensing device is a depth camera using controlled infrared lighting.

18. A system as recited in claim **13**, further comprising, logic that estimates three-dimensional volume information for each physical object within the particular depth range based on the obtained depth values.

* * * * *