US008568230B2

(12) **United States Patent**
Marks et al.

(10) **Patent No.: US 8,568,230 B2**
(45) **Date of Patent: *Oct. 29, 2013**

(54) **METHODS FOR DIRECTING POINTING DETECTION CONVEYED BY USER WHEN INTERFACING WITH A COMPUTER PROGRAM**

(75) Inventors: **Richard L. Marks**, Foster City, CA (US); **Hrishikesh R. Deshpande**, Foster City, CA (US)

(73) Assignee: **Sony Entertainment Computer Inc.**, Tokyo (JP)

(*) Notice: Subject to any disclaimer, the term of this patent is extended or adjusted under 35 U.S.C. 154(b) by 132 days.

This patent is subject to a terminal disclaimer.

(21) Appl. No.: **12/616,117**

(22) Filed: **Nov. 10, 2009**

(65) **Prior Publication Data**

US 2010/0056277 A1 Mar. 4, 2010

Related U.S. Application Data

(63) Continuation of application No. 11/301,673, filed on Dec. 12, 2005, now Pat. No. 7,646,372, which is a continuation-in-part of application No. 10/663,236, filed on Sep. 15, 2003, now Pat. No. 7,883,415, and a continuation-in-part of application No. 10/759,782, filed on Jan. 16, 2004, now Pat. No. 7,623,115.

(51) **Int. Cl.**
A63F 13/06 (2006.01)

(52) **U.S. Cl.**
USPC **463/36**; 463/39; 382/103; 382/312; 345/156; 348/169

(58) **Field of Classification Search**
USPC 463/36, 39; 345/156; 382/103, 312; 348/135, 136, 137, 139, 140, 141, 142, 348/143, 169
See application file for complete search history.

(56) **References Cited**

U.S. PATENT DOCUMENTS

5,394,168 A 2/1995 Smith, III et al. 345/156
5,473,701 A 12/1995 Cezanne et al. 381/92

(Continued)

FOREIGN PATENT DOCUMENTS

EP 0835676 4/1998
EP 1098686 5/2003

(Continued)

OTHER PUBLICATIONS

Ephraim et al. "Speech Enhancement Using a Minimum Mean-Square Error Log-Spectral Amplitude Estimator", 1985, IEEE.

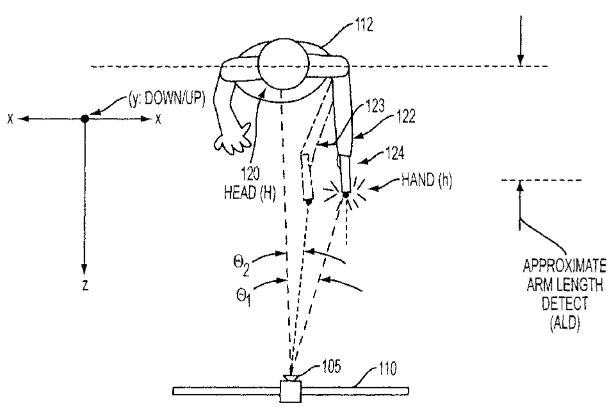
(Continued)

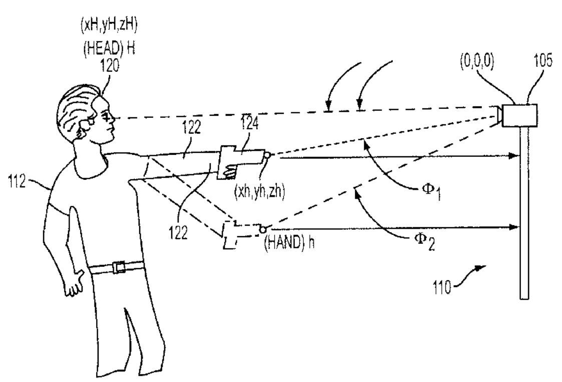
Primary Examiner — Damon Pierce
(74) *Attorney, Agent, or Firm* — Martine Penilla Group, LLP

(57) **ABSTRACT**

A method for detecting direction conveyed by a user when interfacing with a computer program executed on a computing device. The computing device being interfaced with an image capture device. The method includes detecting a human head of a person in the image taken with the image capture device and assigning the human head a head location. And, detecting an object held by the person in the image and assigning the object an object location. The method determines a relative position in space between the head location and the object location from the capture location, such that the relative position is used to identify a pointing direction of the object.

22 Claims, 11 Drawing Sheets





(56) **References Cited**

U.S. PATENT DOCUMENTS

5,485,273	A	1/1996	Mark et al.	356/350
5,611,731	A	3/1997	Bouton et al.	463/37
5,675,825	A	10/1997	Dreyer et al.	395/800
5,846,086	A	12/1998	Bizzi et al.	434/247
5,861,910	A	1/1999	McGarry et al.	348/87
5,930,741	A	7/1999	Kramer	702/153
5,937,081	A	8/1999	O'Brill et al.	382/111
5,959,596	A	9/1999	McCarten et al.	345/2
6,144,367	A	11/2000	Berstis	345/158
6,147,678	A *	11/2000	Kumar et al.	345/158
6,498,860	B1	12/2002	Sasaki et al.	382/103
6,533,420	B1	3/2003	Eichenlaub	353/7
6,587,573	B1	7/2003	Stam et al.	382/104
6,674,415	B2	1/2004	Nakamura et al.	345/32
6,727,988	B2	4/2004	Kim et al.	356/319
6,741,741	B2	5/2004	Farrell	382/199
6,746,124	B2	6/2004	Fischer et al.	353/43
6,772,057	B2	8/2004	Breed et al.	701/45
6,791,531	B1	9/2004	Johnston et al.	345/157
6,928,180	B2	8/2005	Stam et al.	382/104
6,965,684	B2 *	11/2005	Chen et al.	382/103
6,990,639	B2	1/2006	Wilson	715/863
7,043,056	B2	5/2006	Edwards et al.	382/103
7,223,173	B2	5/2007	Masuyama et al.	463/36
7,259,747	B2 *	8/2007	Bell	345/156
7,263,462	B2	8/2007	Funge et al.	702/179
7,296,007	B1	11/2007	Funge et al.	706/47
7,352,359	B2	4/2008	Zalewski et al.	345/156
7,391,409	B2	6/2008	Zalewski et al.	345/156
7,436,887	B2	10/2008	Yeredor et al.	375/240
7,545,926	B2	6/2009	Mao	379/406.08
7,558,698	B2	7/2009	Funge et al.	702/179
7,576,727	B2 *	8/2009	Bell	345/158
7,613,610	B1	11/2009	Zimmerman et al.	704/235
7,623,115	B2	11/2009	Marks	345/156
7,627,139	B2	12/2009	Marks et al.	382/103
7,636,645	B1	12/2009	Yen et al.	702/152
7,697,700	B2	4/2010	Mao	381/94.3
7,721,231	B2	5/2010	Wilson	715/863
2003/0022716	A1	1/2003	Park et al.	463/36
2004/0046736	A1	3/2004	Pryor et al.	345/156
2004/0178576	A1	9/2004	Hillis et al.	273/148
2004/0227725	A1	11/2004	Calarco et al.	345/156
2005/0059488	A1	3/2005	Larsen et al.	463/36
2005/0226431	A1	10/2005	Mao	381/61
2006/0038819	A1	2/2006	Festejo et al.	345/530
2006/0204012	A1	9/2006	Marks et al.	381/26
2006/0233389	A1	10/2006	Mao et al.	381/92
2006/0252541	A1	11/2006	Zalewski et al.	463/156
2006/0256081	A1	11/2006	Zalewski et al.	
2006/0264258	A1	11/2006	Zalewski et al.	463/36
2006/0264259	A1	11/2006	Zalewski et al.	463/36
2006/0264260	A1	11/2006	Zalewski et al.	463/36
2006/0269072	A1	11/2006	Mao	381/56
2006/0269073	A1	11/2006	Mao	381/56
2006/0274032	A1	12/2006	Mao et al.	345/156
2006/0274911	A1	12/2006	Mao et al.	381/334
2006/0280312	A1	12/2006	Mao	381/56
2006/0282873	A1	12/2006	Zalewski et al.	725/133
2006/0287084	A1	12/2006	Mao et al.	463/37
2006/0287085	A1	12/2006	Mao et al.	463/37
2006/0287086	A1	12/2006	Zalewski et al.	436/37
2006/0287087	A1	12/2006	Zalewski et al.	463/37
2007/0015559	A1	1/2007	Zalewski et al.	463/1
2007/0021208	A1	1/2007	Mao et al.	463/36
2007/0025562	A1	2/2007	Zalewski et al.	381/92
2007/0061413	A1	3/2007	Larsen et al.	709/217
2007/0072675	A1	3/2007	Hammano et al.	463/42
2007/0260340	A1	11/2007	Mao	700/94
2007/0260517	A1	11/2007	Zalewski et al.	705/14
2007/0261077	A1	11/2007	Zalewski et al.	725/35
2008/0091421	A1	4/2008	Gustavsson	704/233
2009/0010494	A1	1/2009	Bechtel et al.	382/104
2009/0016642	A1	1/2009	Hart	382/278
2009/0221368	A1	9/2009	Yen et al.	463/32
2009/0221374	A1	9/2009	Yen et al.	463/42
2009/0288064	A1	11/2009	Yen et al.	717/106
2010/0004896	A1	1/2010	Yen et al.	702/153
2010/0137064	A1	6/2010	Shum et al.	463/36

FOREIGN PATENT DOCUMENTS

GB	2 398 691 A	8/2004
WO	WO 88/05942	8/1988
WO	WO 98/48571	10/1998
WO	WO 2004/041379 A2	5/2004
WO	WO 2008/056180	5/2008

OTHER PUBLICATIONS

Ephraim et al. "Speech Enhancement Using a Minimum Mean-Square Error Short-Time Spectral Amplitude Estimator", 1984, IEEE.

Richardson et al. "Virtual Network Computing", 1998, IEEE Internet Computing vol. 2.

XP-002453974, "CFS and FS95/98/2000: How to Use the Trim Controls to Keep Your Aircraft Level", Aug. 10, 2007, http://support.microsoft.com/?scid=kb%3Ben-us%3B175195&x=13&y=15.

* cited by examiner

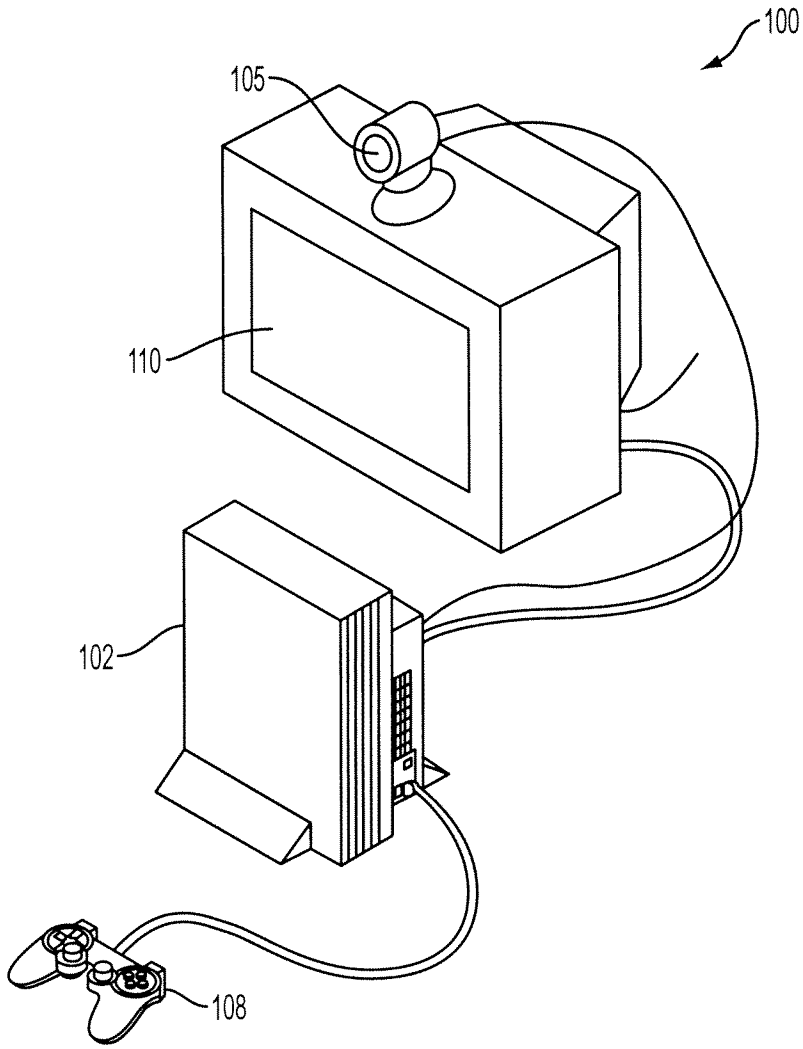


FIG. 1A

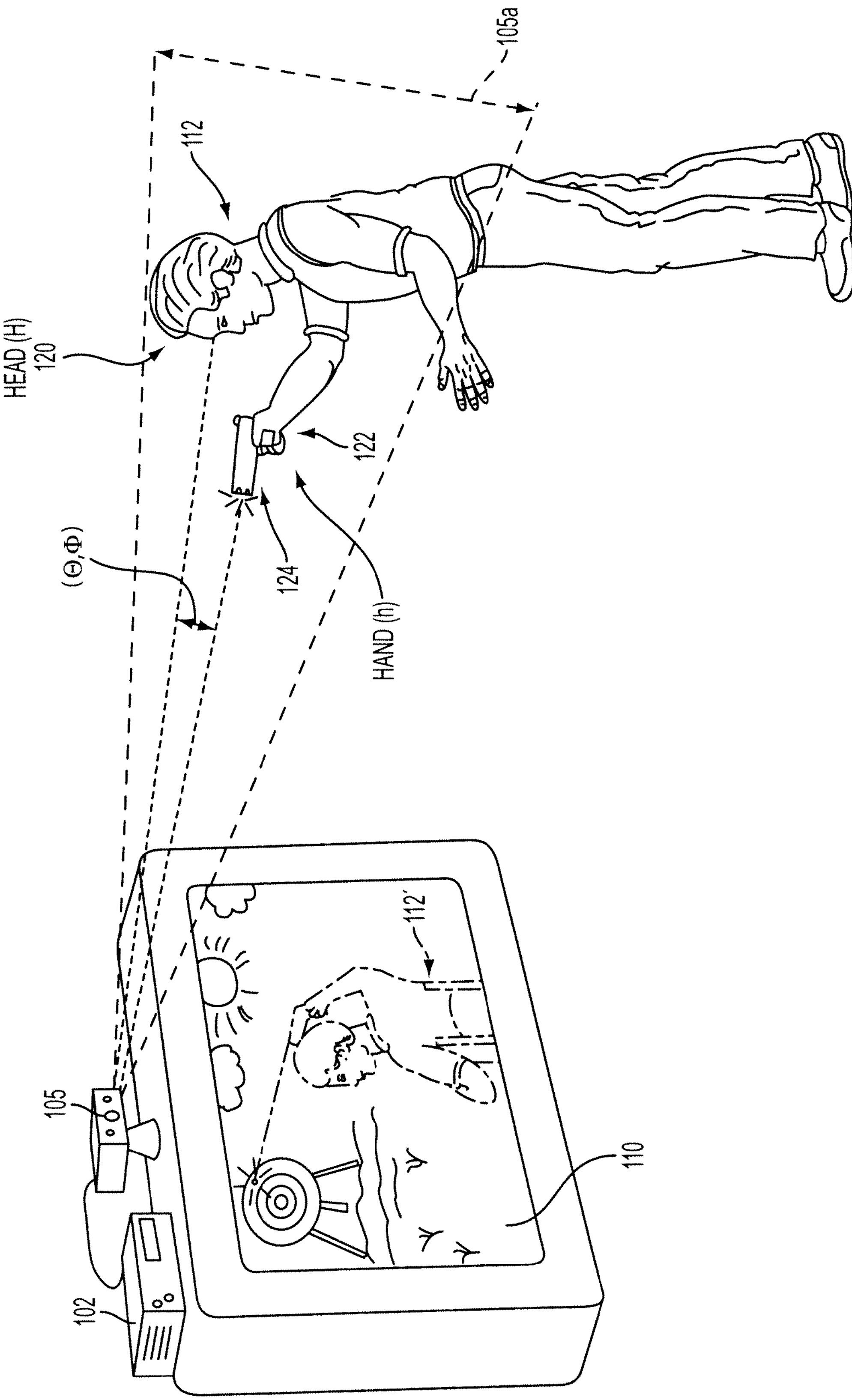


FIG. 1B

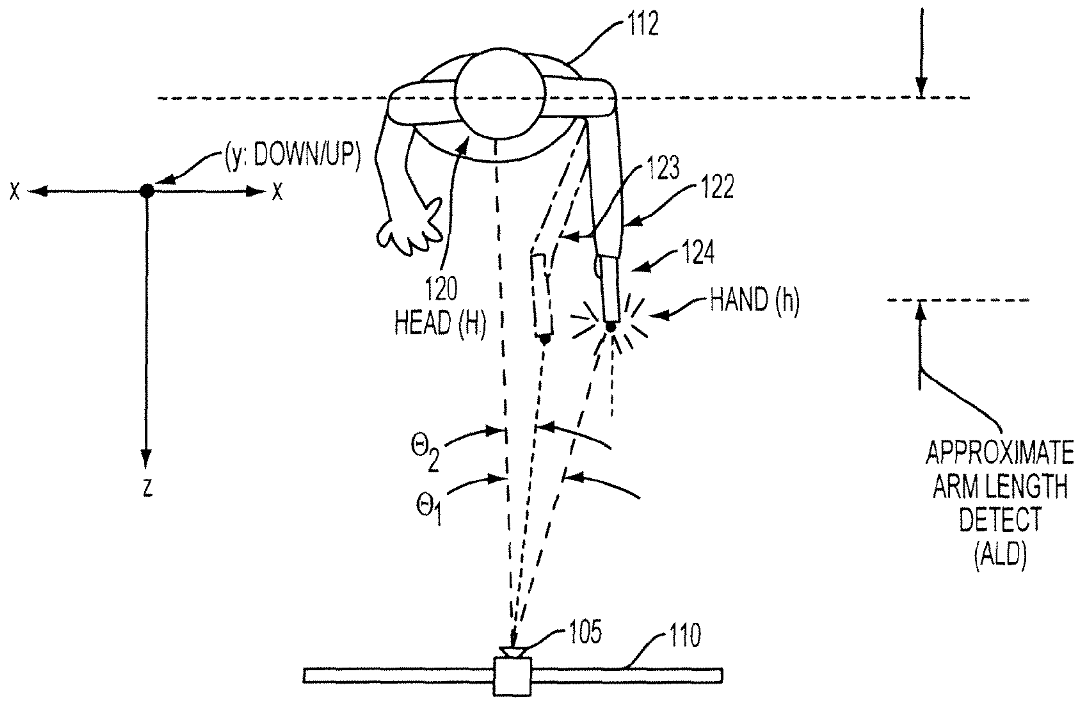

112
(y: DOWN/UP)
x
x
z
123
122
124
120
HEAD (H)
HAND (h)
Θ2
Θ1
APPROXIMATE
ARM LENGTH
DETECT
(ALD)
105
110

FIG. 2

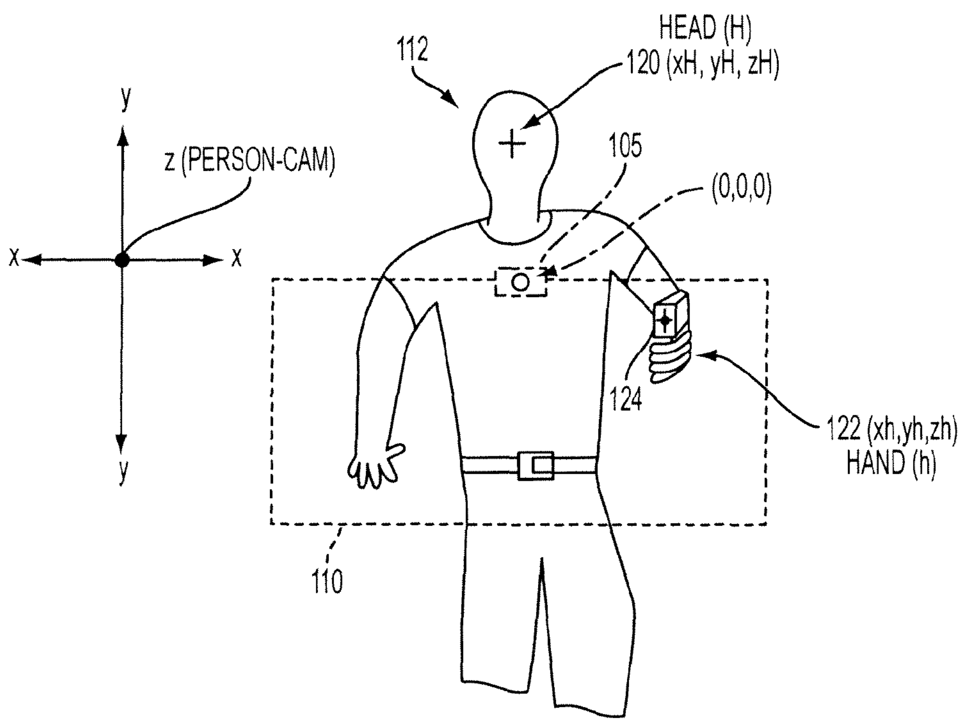

HEAD (H)
120 (xH, yH, zH)
112
y
z (PERSON-CAM)
105
(0,0,0)
x
x
y
124
122 (xh,yh,zh)
HAND (h)
110

FIG. 3

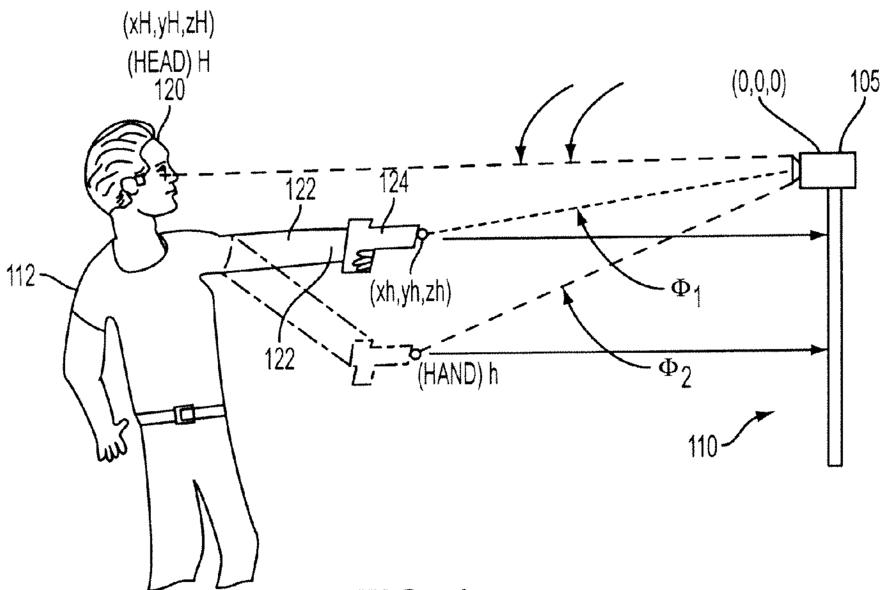


FIG. 4

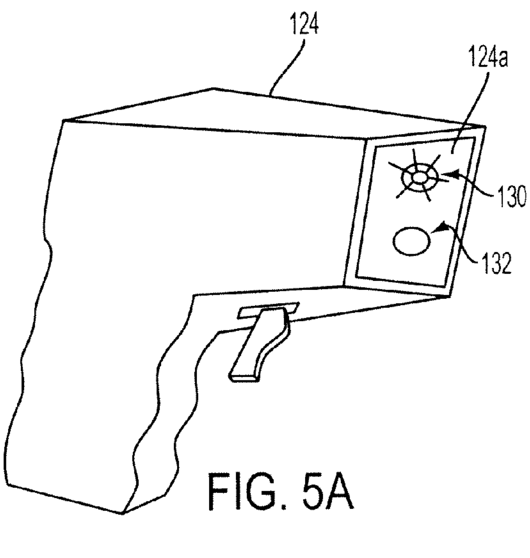


FIG. 5A

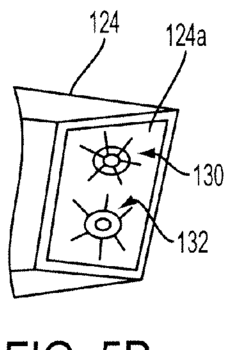


FIG. 5B

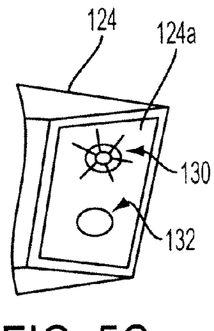


FIG. 5C

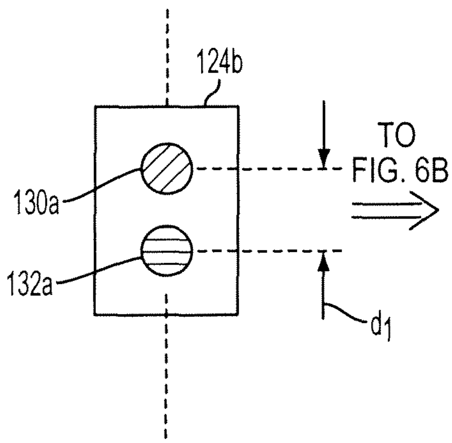


FIG. 6A

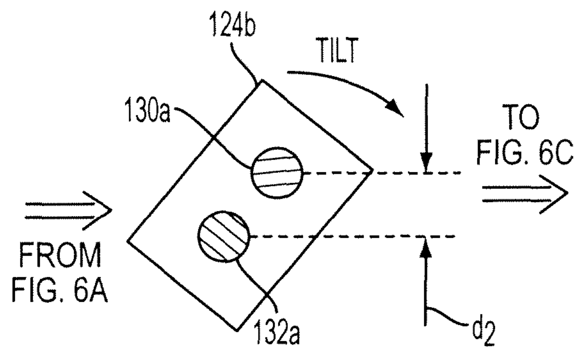


FIG. 6B

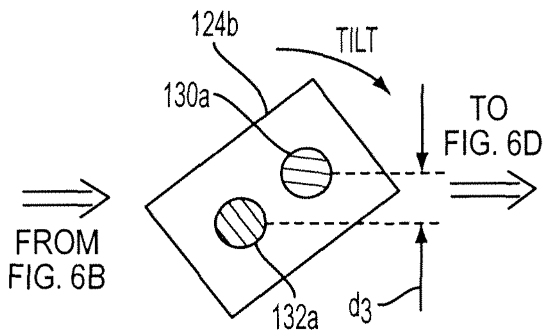


FIG. 6C

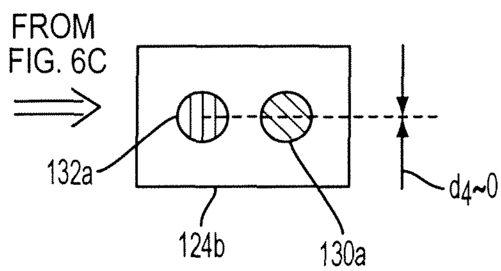


FIG. 6D

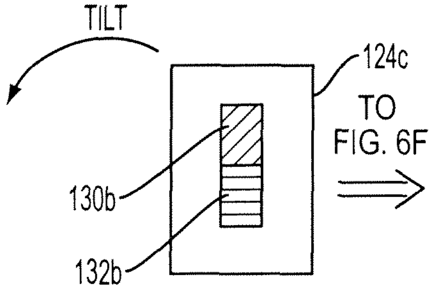


FIG. 6E

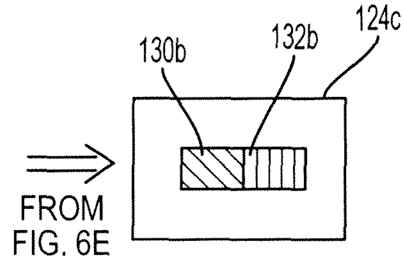


FIG. 6F

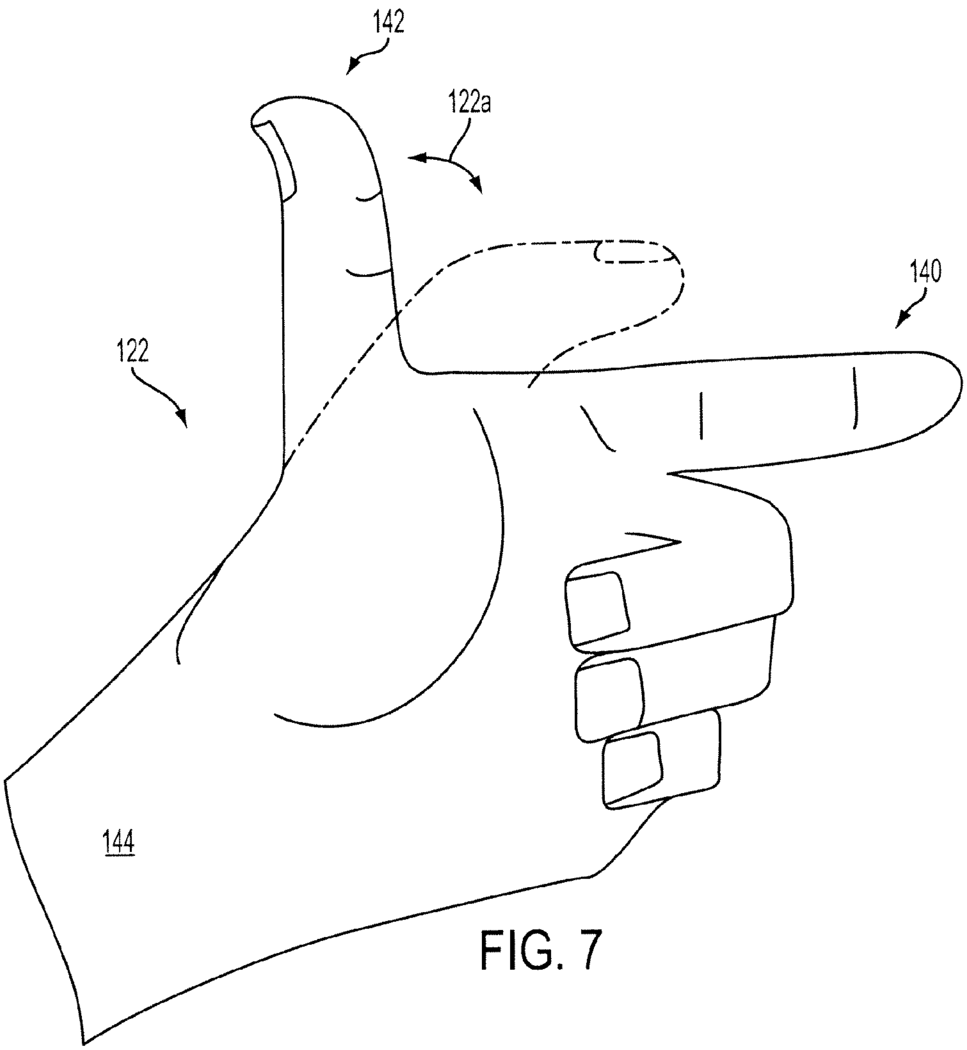


FIG. 7

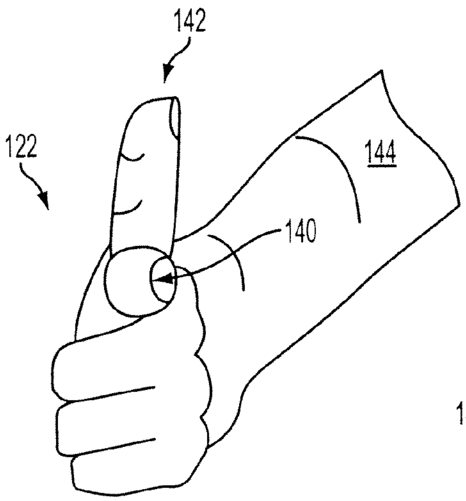


FIG. 8A

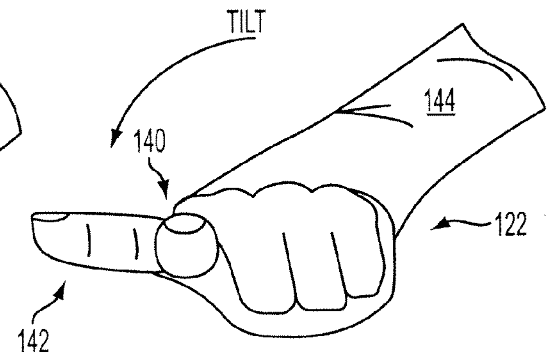


FIG. 8B

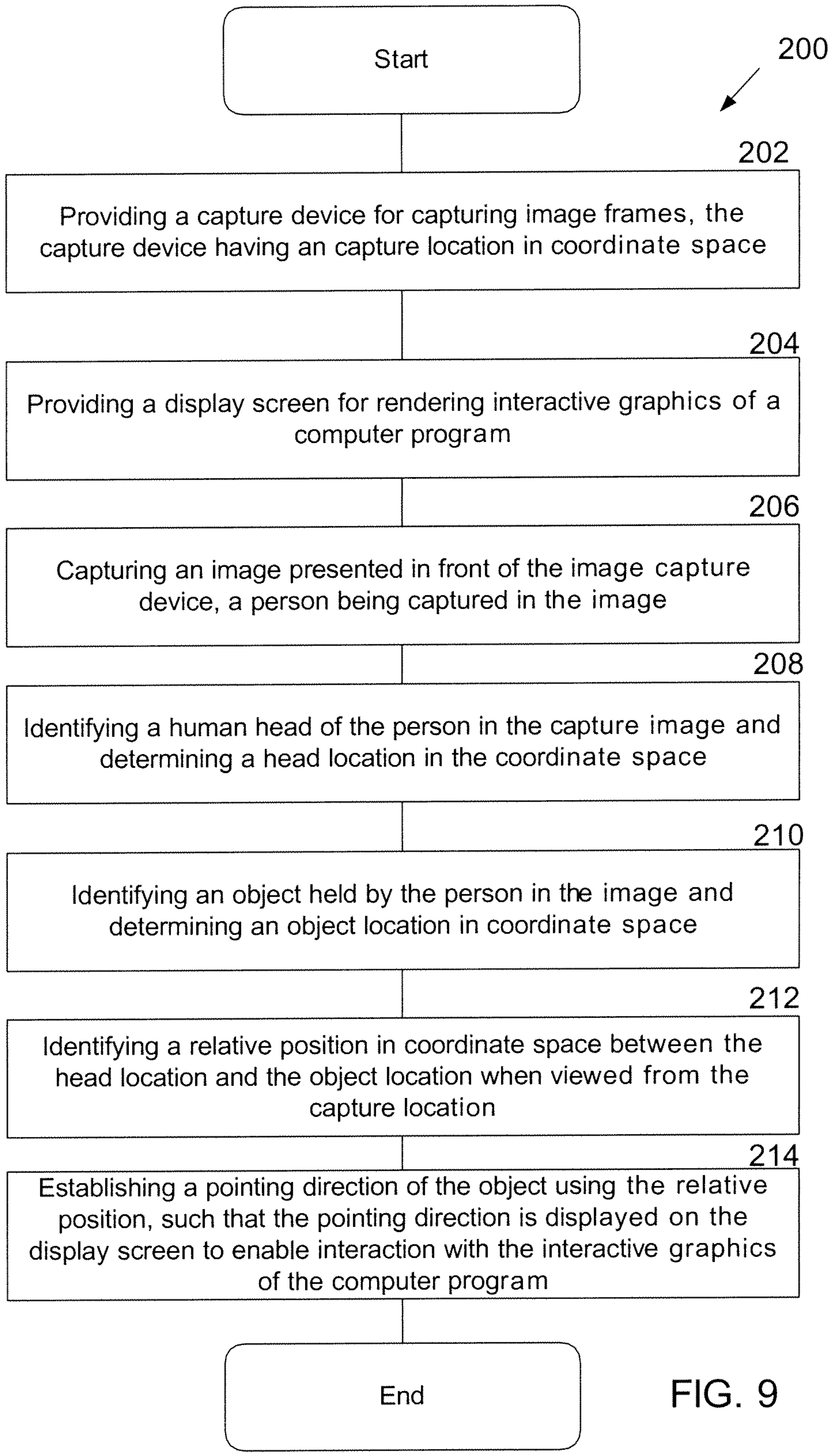

Start
200
202
Providing a capture device for capturing image frames, the capture device having an capture location in coordinate space
204
Providing a display screen for rendering interactive graphics of a computer program
206
Capturing an image presented in front of the image capture device, a person being captured in the image
208
Identifying a human head of the person in the capture image and determining a head location in the coordinate space
210
Identifying an object held by the person in the image and determining an object location in coordinate space
212
Identifying a relative position in coordinate space between the head location and the object location when viewed from the capture location
214
Establishing a pointing direction of the object using the relative position, such that the pointing direction is displayed on the display screen to enable interaction with the interactive graphics of the computer program
End

FIG. 9

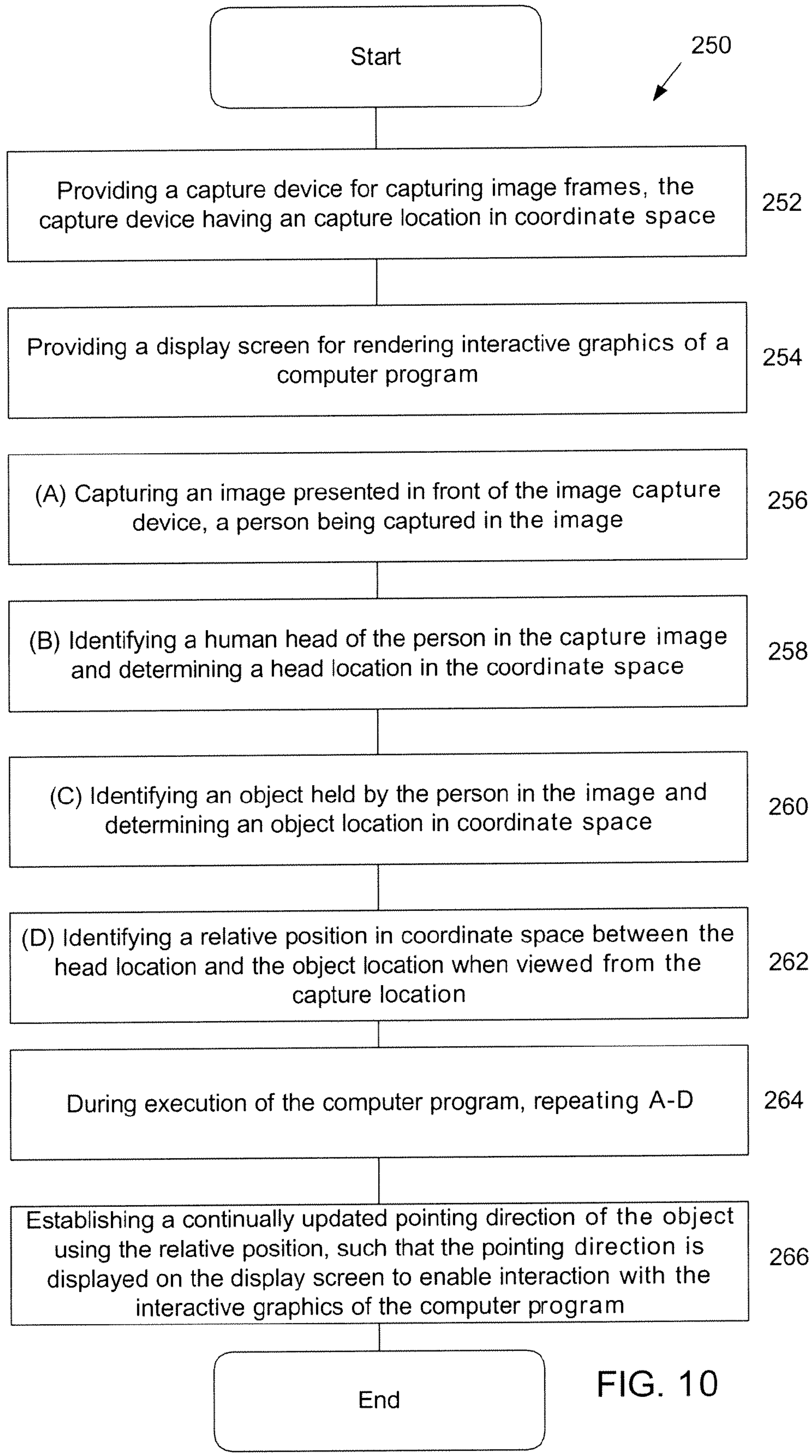

Start
250
Providing a capture device for capturing image frames, the capture device having an capture location in coordinate space
252
Providing a display screen for rendering interactive graphics of a computer program
254
(A) Capturing an image presented in front of the image capture device, a person being captured in the image
256
(B) Identifying a human head of the person in the capture image and determining a head location in the coordinate space
258
(C) Identifying an object held by the person in the image and determining an object location in coordinate space
260
(D) Identifying a relative position in coordinate space between the head location and the object location when viewed from the capture location
262
During execution of the computer program, repeating A-D
264
Establishing a continually updated pointing direction of the object using the relative position, such that the pointing direction is displayed on the display screen to enable interaction with the interactive graphics of the computer program
266
End

FIG. 10

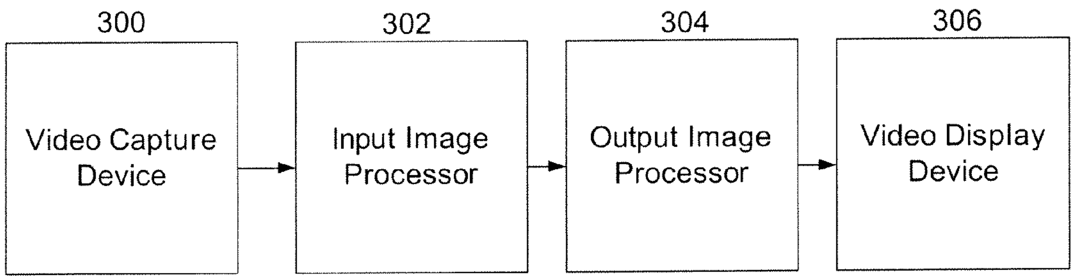

300
Video Capture Device
302
Input Image Processor
304
Output Image Processor
306
Video Display Device

FIG. 11

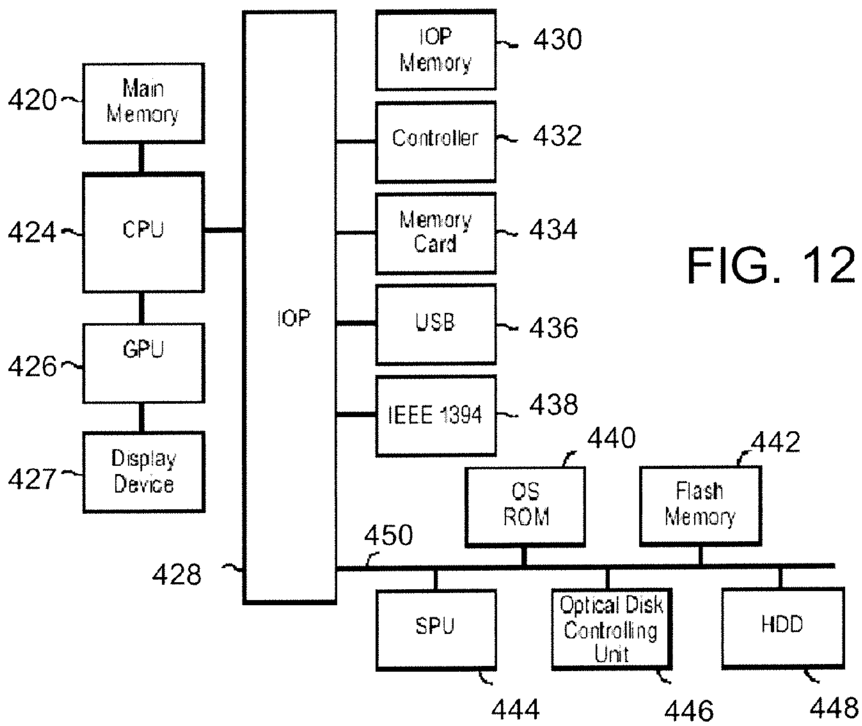

420
Main Memory
424
CPU
426
GPU
427
Display Device
IOP
428
IOP Memory
430
Controller
432
Memory Card
434
USB
436
IEEE 1394
438
OS ROM
440
Flash Memory
442
450
SPU
444
Optical Disk Controlling Unit
446
HDD
448

FIG. 12

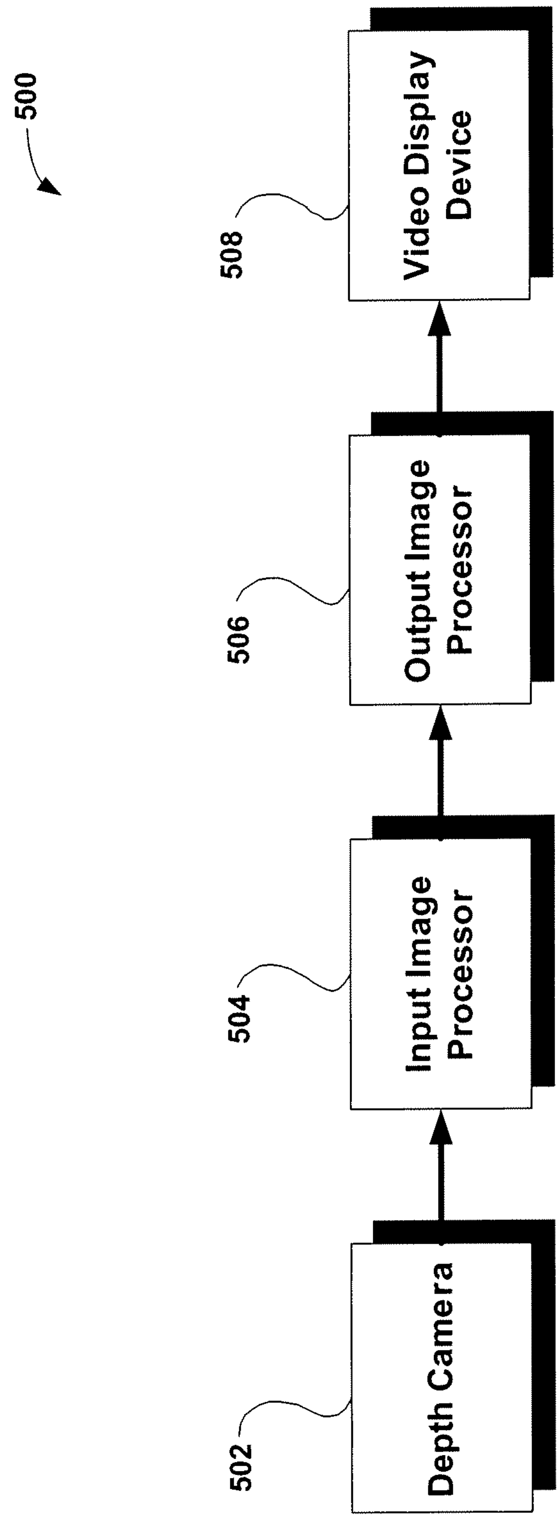

500
502
Depth Camera
504
Input Image Processor
506
Output Image Processor
508
Video Display Device

FIG. 13A

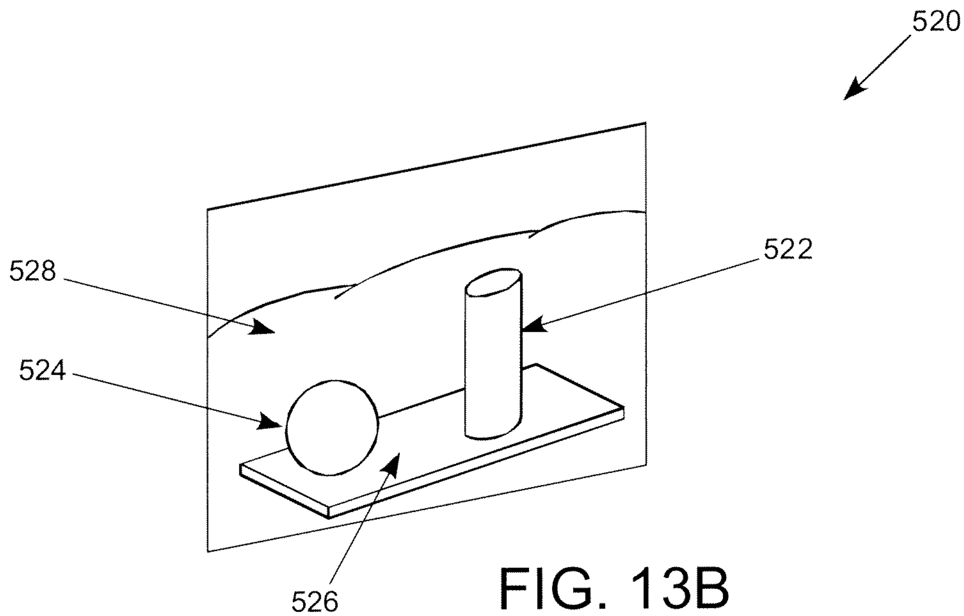


FIG. 13B

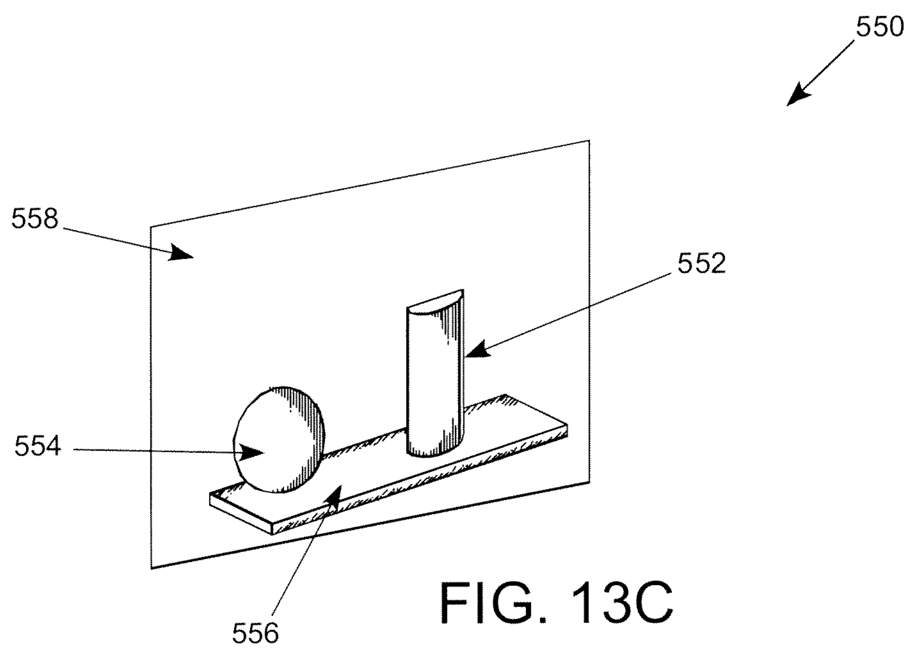


FIG. 13C

METHODS FOR DIRECTING POINTING DETECTION CONVEYED BY USER WHEN INTERFACING WITH A COMPUTER PROGRAM

CLAIM OF PRIORITY

This application is a Continuation application of U.S. patent application Ser. No. 11/301,673, entitled "Methods and Systems for Enabling Direction Detection When Interfacing With A Computer Program", filed on Dec. 12, 2005 now U.S. Pat. No. 7,646,372, which is a continuation in part (CIP) of U.S. patent application Ser. No. 10/663,236, entitled "METHOD AND APPARATUS FOR ADJUSTING A VIEW OF A SCENE BEING DISPLAYED ACCORDING TO TRACKED HEAD MOTION", filed on Sep. 15, 2003 now U.S. Pat. No. 7,883,415, and is also a continuation in part (CIP) of U.S. patent application Ser. No. 10/759,782, entitled "METHOD AND APPARATUS FOR LIGHT INPUT DEVICE", filed on Jan. 16, 2004 now U.S. Pat. No. 7,623,115. Each of the above identified applications is hereby incorporated by reference.

CROSS-REFERENCE TO RELATED APPLICATION

This application is related to U.S. patent application Ser. No. 11/302,511, filed on Dec. 12, 2005, entitled "METHODS AND SYSTEMS FOR ENABLING DEPTH AND DIRECTION DETECTION WHEN INTERFACING WITH A COMPUTER PROGRAM" to inventors Richard L. Marks and Hrishikesh R. Deshpande, which is hereby incorporated by reference.

BACKGROUND

Description of the Related Art

The video game industry has seen many changes over the years. As computing power has expanded, developers of video games have likewise created game software that takes advantage of these increases in computing power. To this end, video game developers have been coding games that incorporate sophisticated operations and mathematics to produce a very realistic game experience.

Example gaming platforms, may be the Sony Playstation or Sony Playstation2 (PS2), each of which is sold in the form of a game console. As is well known, the game console is designed to connect to a monitor (usually a television) and enable user interaction through handheld controllers. The game console is designed with specialized processing hardware, including a CPU, a graphics synthesizer for processing intensive graphics operations, a vector unit for performing geometry transformations, and other glue hardware, firmware, and software. The game console is further designed with an optical disc tray for receiving game compact discs for local play through the game console. Online gaming is also possible, where a user can interactively play against or with other users over the Internet.

As game complexity continues to intrigue players, game and hardware manufacturers have continued to innovate to enable additional interactivity. In reality, however, the way in which users interact with a game has not changed dramatically over the years. Commonly, users still play computer games using hand held controllers or interact with programs using mouse pointing devices.

In view of the foregoing, there is a need for methods and systems that enable more advanced user interactivity with game play.

SUMMARY OF THE INVENTION

Broadly speaking, the present invention fills these needs by providing an apparatus and method that facilitates interactivity with a computer program. In one embodiment, the computer program is a game program, but without limitation, the apparatus and method can find applicability in any consumer electronic device that will require a user to interact therewith. The present invention simplifies user interaction experience through machine recognizable gestures based on pointing to the interface and discriminating commands based on factors including trigger cues and position determination of a hand or object under user control.

In one embodiment, a method for detecting direction conveyed by a user when interfacing with a computer program executed on a computing device. The computing device being interfaced with an image capture device. The method includes detecting a human head of a person in the image taken with the image capture device and assigning the human head a head location. And, detecting an object held by the person in the image and assigning the object an object location. The method determines a relative position in space between the head location and the object location from the capture location, such that the relative position is used to identify a pointing direction of the object. In one embodiment, the object is a body part of the user, other than the head of the user.

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

The invention, together with further advantages thereof, may best be understood by reference to the following description taken in conjunction with the accompanying drawings.

FIGS. **1**A and **1**B illustrate the environment of one embodiment of present invention, which takes advantage of the pointing direction determination described herein.

FIGS. **2**-**4** illustrate the analysis of the relative positions between a person's head and a person's hand (or object held by the person), when determining pointing direction.

FIGS. **5**A-**5**C illustrate tracking and triggering embodiments, which can be identified by the computer program by the analysis of the captured image data, in accordance with one embodiment of the present invention.

FIGS. **6**A-**6**F illustrate alternate embodiments of detecting an object and detecting changes in the object based on relative orientation of the object itself, in accordance with one embodiment of the present invention.

FIGS. **7** and **8**A-**8**B illustrate examples of hand positions, when the hand positions and shapes are analyzed to determine a desired trigger or tracking response by the computer program, in accordance with one embodiment of the present invention.

FIGS. **9** and **10** provide exemplary flow charts of the operations that can be executed in determining a pointing direction, in accordance with one embodiment of the present invention.

FIGS. **11** and **12** provide exemplary embodiments of hardware that may be used in processing the computer code necessary to executed the claimed operations, in accordance with one embodiment of the present invention.

FIGS. **13**A-**13**C illustrate exemplary embodiments for when depth information is considered when identifying objects to track, when interfacing with a system that needs to ascertain pointing direction, in accordance with one embodiment of the present invention.

DETAILED DESCRIPTION

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 obscure the present invention.

FIG. **1**A illustrates an interactive game setup **100**, in accordance with one embodiment of the present invention. The interactive game setup **100** includes a computer **102** that is coupled to a display screen **110**. An image capture device **105** is placed on top of the display screen **110** and is coupled to the computer **102**. Computer **102** is, in one embodiment, a gaming system console which allows users to play video games and interface with the video games through controllers **108**. The image capture device **105** is shown placed on top of the display screen **110**, but it should be understood that the image capture device **105** can be placed in any other proximate location that will allow it to capture images that are located about in front of the display screen **110**. Techniques for capturing these movements and interactions can vary, but exemplary techniques are described in United Kingdom Applications GB 0304024.3 (PCT/GB2004/000693) and GB 0304022.7 (PCT/GB2004/000703), each filed on Feb. 21, 2003, and each of which is hereby incorporated by reference.

In a specific example, but not limited to any brand, the game console can be a one manufactured by Sony Computer Entertainment Inc., Nintendo, Microsoft, or any other manufacturer. The image capture device **105** can be as simple as a standard web cam or can include more advanced technology. In one embodiment, the image capture device should be capable of capturing images, digitizing the images, and communicating the image data back to the computer **102**. In some embodiments, the image capture device will have logic integrated therein for performing the digitizing and another embodiment the image capture device **105** will simply transmit the captured data back to the computer **102** for digitizing. In either case, the image capture device **105** is capable of capturing either color or black and white images of any object located in front of the image capture device **105**.

FIG. **1**B illustrates an exemplary embodiment of the present invention in which the computer **102** processes image data provided by the image capture device **105** to ascertain a pointing direction of an object placed in front of the image capture device **105**. As shown, the computer **102** is connected to the image capture device **105**. The image capture device **105** is designed to focus onto a capture region **105***a*. In this example, a person **112** is intending to interact with a computer program being executed by the computer **102**. The computer program, in this example, is a video game which is rendered and displayed by the display screen **110**.

For example purposes only, the video game is a target shooting game in which the person **112** wishes to aim at a target and earn points commensurate with his or her performance. As illustrated on the display screen **110**, an image **112**' of the person **112** may also be placed on the display screen **110** during game play. Alternatively, the person's image **112**' may be omitted from the display screen, depending on the particular device under control or game being played. In this example, the user experience may be enhanced by illustrating an image **112**' of the person **112** during the target shooting exercise to present more reality during game play. A feature of the target shooting game is the ability for person **112** to point or direct an object **124** at particular interactive graphics on the display screen **110**.

To achieve accurate pointing direction of the object **124**, which in this case and for example purposes is a gun, the person **112** will hold the object **124** with his or her hand **122**. The hand **122** will be directionally pointed toward the display screen **110**. The image capture device **105** will at this point, analyze the digital image capture of the person **112** to determine the location of the person's **112** head **120**, and the location of the person's **112** hand **122**. As shown, the person's **112** hand is extended in front of his body and the image capture device will identify the object **124** when examining the captured digital image. The captured digital image will also be examined by code executed at the computer **102** to ascertain the location of the person's **112** head **120**. In one embodiment, head tracking is completed with a combination of a template matching (for speed performance), coupled to a face detection code. The face detection code will essentially identify the location of the user's face by locating the user's eyes and other facial features. For additional information on head and face detection, reference may be made to co-pending U.S. patent application Ser. No. 10/663,236, entitled "METHOD AND APPARATUS FOR ADJUSTING A VIEW OF A SCENE BEING DISPLAYED ACCORDING TO TRACKED HEAD MOTION", filed on Sep. 15, 2003.

The object **124** will, in one embodiment, have an identifier which may be color or lights (e.g., light emitting diodes "LEDs") coupled to the object so that the program analyzing the captured digital image will easily identify the location of the object **124**. Once the computer program has identified the location of the person's head **120** (H) and the location of the person's hand **122** (*h*), the computer program will perform computations to determine a relative angle from the image capture device position, and between the detected object **124**, and the head **120**.

As illustrated in FIG. **1**B, the relative position of the object **124** and the head **120** will be calculated relative to the image capture device **105**. This will produce two angle parameters (theta θ and phi Φ). The azimuth angle θ will define the horizontal positioning between the head **120** and the object **124** along an X axis. The phi angle Φ will produce the altitude angle which is the relative angle between the height of the head **120** and the height of the hand **122**. In one embodiment, an initial calibration operation may be performed before a gaming operation begins to calibrate the object's pointing location on the display screen **110**. For instance, the user may be prompted to calibrate the pointing algorithm by having the user point the object **124** at a specific location on the display screen **110**. Once the calibration has been completed, the computer **102** will be able to calculate the azimuth angle and the altitude angle (theta and phi) which define the relative positions of the person's head **120** and the person's hand **122**, for each successive frame being captured by the image capture device **105**. The relative positioning between the head and the hand may be calculated for each captured frame or may be captured every other frame, or after a number of frames are captured, depending on the accuracy required for the pointing operation. For example, if the game is a shooting gallery game, it would be important for the relative positioning of the head **120** and the hand **122** to be computed for each frame so that the person **112** will have accurate aiming and triggering capabilities when attempting to secure a good performing score in the video game contest.

FIG. **2** illustrates a top view of the person **112** who is positioned in front of the display screen **110**, in accordance with one embodiment of the present invention. In this example, an initial azimuth angle ($theta_1$) is shown being determined as the relative angle between the position of the hand **122** (which is holding object **124**), and the head **120**. The person **112**, during interactive play with the computer program, will be facing the display screen **110** and most likely, will maintain his body substantially parallel to the display screen **110**. When the person **112** maintains his body substantially parallel to the display screen **110**, movement of the hand **122** in direction **123** will cause the azimuth angle to be recomputed and produce a new azimuth angle $theta_2$.

In this example, the person **112** is holding the object **124** out in front of his body at about arm's length. This distance is shown to be the approximate arm's length detect (ALD) that defines the location of the approximate place where the image capture device **105** will attempt to capture the position of the hand **122** and associated object **124**. The approximate arm's length detect (ALD) can vary, depending upon the user's arm length, but generally should be provided to allow a distance relationship between the location of the head **120** and the hand **122**. For instance, there should be at least a minor projection of the hand **122** in front of the person's body to point to different locations of the display screen **110**.

FIG. **3** illustrates a front view of the person **112** where the head **120** is identified and the hand **122** is identified, from the captured digital image. In this example, a dashed line represents the display screen **110** and the image capture device **105** that is directionally pointed at the person **112**. In this example, the image capture device **105** will be illustrated to be at a coordinate space of (0, 0, 0), representing the reference point of the image capture device **105**, and its position in coordinate space. The approximate location of the head **120** will also have an associated coordinate space (x H, y H, z H). Likewise, the hand **122** and the associated object **124** that is held by the hand **122** will have a coordinate space (x h, y h, z h) that is relative to the image capture device **105**.

FIG. **4** illustrates the person **112** pointing the object **124** at the display screen **110**. In this example, the coordinate space of the head **120** is identified in the digital image captured by the capture device **105**. The location in coordinate space of the hand **122** is also captured in the digital image captured by the image capture device **105** when the person **112** is pointing at the display screen **110**. The image capture device **105** is the reference point, which is at the coordinate space (0, 0, 0). The altitude angle phi is therefore calculated between the position of the head **120** relative to the position of the hand **122**. In one example, the angle is calculated as:

$$\text{Altitude angle}=\arctan((yh-yH)/(zh-zH))$$

In a similar manner, the azimuth angle theta of FIG. **2** is calculated as:

$$\text{Azimuth angle}=\arctan((xh-xH)/(zh-zH))$$

When the user moves his hand down (e.g., as captured in a later frame) as illustrated in FIG. **4**, a new angle phi_2 will be produced to define the new relative position between the hand **122** and the head **120** of the person **112**. Based on this new relative positioning of the head and the hand, the computer **102** will re-position the pointing direction on the display screen.

FIG. **5**A illustrates an embodiment where the object **124** is a gun. The gun may be a simplified gun object that is configured to be pointed in the direction of the display screen **110** to hit particular objects or interact with objects on the display screen **110**. In this example, the gun **124** will include a detection region **124***a*. Detection region **124***a* is the region which is directionally pointed toward the display screen **110**. The detection region **124***a* is also the region that is captured by the image capture device **105** for analysis of the digital image by the computer **102**. In one embodiment, the detection region **124***a* is configured to include a pair of lighting objects that will assist in the interactivity with the computer program being executed by the computer **102** and displayed on the display screen **110**. In this example, a tracking indicator **130** is provided as a light or color object that is present on the detection region **124***a*. Based on the tracking indicator **130**, the image capture device **105** will produce a digital image that will be analyzed by the computer **102** to identify the position in coordinate space of the object **124**. In this example, by providing the tracking indicator **130**, the computer program being executed on the computer **102** is able to quickly identify the location of the object **124** and in relation to the head **120** of the person interacting with the computer program.

The tracking indicator **130** may be provided by way of a number of implementations. One implementation might be a light indicator that can be tracked by the computer program that analyzes the captured digital images, or may be in the form of a color indicator that the computer can identify quickly from the captured digital images. The hand itself may be the tracking indicator **130**. In still another embodiment, the tracing indicator **130** may be provided as a reflective tape that will have different coloring or intensity depending on the angle that it might be displayed when shown to the image capture device **105**. In this example, the object **104** is tracked as the user moves his hand **122** to different regions pointed to on the display screen **110**.

In one embodiment, while the user moves his hand **122** relative to the head **120**, the tracking indicator **130** will allow the computer program to provide a visual indicator on the display screen **110**. This visual indicator on the display screen **110** will allow the user to understand where the object is currently pointing to on the display screen **110**.

In another embodiment, the detection region **124***a* will also include a trigger indicator **132**. The trigger indicator **132** may be in the form of a light that is triggered ON and OFF when the user pulls the trigger of the object **124**. For instance, the detection region **124***a* is shown in FIG. **5**B after the trigger has been pulled and the trigger indicator **132** is lit. When the trigger indicator **132** is lit as shown in FIG. **5**B, the computer program executing on the computer **102** will provide an indictor on the display screen **110** so that the user can identify whether his or her pointing has accurately hit an object of the computer game. In FIG. **5**C, the trigger indicator **132** is shown to be in the OFF position to signify that the object **124** will still remain actively tracked, but the shooting which can be continuous or intermittent, can be discontinued when the user removes his finger from the trigger of the object **124**. The trigger indicator **132** may operate in any frequency range, including audio, ultrasonic, visible light wave, infrared and radio. Passive trigger indicator **132** may be achieved through a mechanical sound generated upon actuating a trigger and receiving and decoding the audio input to the device and determining whether the trigger was actuated.

FIG. **6**A illustrates another embodiment of the present invention where tracking and trigger indicators **130***a* and **132***a* are provided. In this example, the track/trigger indicators **130***a* and **132***a* are provided so that determinations can be made of the relative distances between the two indicators as shown by distance (d_1). In one example, the object, when pointed at the image capture device **105** may respond by having the computer program that is executed on the computer **102** to ascertain the distance d_1 and perform an interactive action on the display screen. When the object **124***b* is tilted relative to the starting position of FIG. **6**A, a second distance (d_2) is computed. This distance is the new distance between the track/trigger indicators **130***a* and **132***a*. As the user continues to tilt the object **124***b* as shown in FIG. **6**C, the distance continues to shrink as shown by distance (d_3). Once the object **124***b* has been placed in the horizontal position relative to the vertical position of FIG. **6**A, the distance between the track and trigger indicators **130***a* and **132***a* is brought to approximately zero. At this point, the program may read that the user intends for a trigger action to occur, or any other action that can be triggered when the detected distance from d_1 to d_4 has been detected.

In another embodiment, the response by the computer program may be gradually changed, depending on the angle at which the detection region **124***b* is tilted. For instance, the user may immediately begin to shoot the gun (or trigger the shooting of the gun) when the tilting begins and is executed between the tilt of FIG. **6**B to the tilt of FIG. **6**D. When the user tilts the gun back to the original position, the gun may discontinue the shooting activity. Consequently, the trigger activity caused by analyzing the patterns or colors of the tracking and trigger indicators of **130***a* and **132***b* can cause the computer program to react in different interactive ways.

An example of this interactivity may be to trigger a reloading operation to occur for a gun that is being used in a video game, or a change of gun type being used on the video game program. Once these changes are processed, the video display screen **110** will produce a different graphical animation for the user, depending upon the control being provided and detected by the image capture device.

Commands and trigger states are not limited to an ON and OFF parameters, but can be incrementally changed depending on the position of the relative state and angles of the trigger and track indicators. For example, the state of the trigger may be determined in a linear or variable state as opposed to ON or OFF. Any technique can be used to determine the relative trigger position including a resistive types used to control acceleration in remote control race tracks. The device, or gun in this example, can communicate the state of its trigger by encoding and transmitting its value in any of a multitude of ways known in the art. A variety of commands and gestures may be formulated based on the state of the trigger and the position of the device, including those based on all machine recognizable gestures.

FIGS. **6**E and **6**F provide yet another embodiment where different colors may be used to track tilt or relative tilt between positions of the detection regions **124***c*. In this example, the track and trigger indicators **130***b* and **132***b* are square or rectangular in dimension and can be defined by colored tapes, bar codes, light indicators, LEDs, or the like. As a user flips or tilts the detection region **124***c* from the position of FIG. **6**E to FIG. **6**F, for example, the reaction by the computer game as displayed on the computer display screen will change.

FIG. **7** illustrates another embodiment of the present invention in which a hand **122** is used to control the interactivity on the display screen. The user may place the hand **122** in front of the image capture device **105** and it is tracked relative to the head of the person **112**. In this example, the user's hand is extended having of an index finger **140** pointing at the display screen **110**. The user's thumb **142** may be pointing upright to indicate to the computer program executing on the computer **102** that the trigger device has not been activated. When the user's thumb **142** is moved down toward the index finger **140** in the direction **122***a*, the computer program executing on a computer **102** may detect from the captured digital image that the user intends to shoot or trigger or interactively point to a specific region on the display screen **110**. Thus, the user's hand being placed in a different position can trigger an event or cause the interactivity of a command with a computer program being executed and shown on the display screen **110**. For example, the user may be able to shoot by different hand gestures, may be able to reload the gun with different gestures, and the different positions or orientations of the user's hand may cause different graphical renderings of the user or gun on the display screen when the user is interacting with a particular game program. In this embodiment, the state of the trigger may be determined ON or OFF as described above, or, the trigger may be determined in a variable state. In the latter, the relative position of the user's thumb may range from, for example, the position in which the thumb is substantially perpendicular to the pointing finger and the position where the thumb is substantially parallel to the pointing finger and where the system performs image analysis of the hand to yield a relative state of the thumb. This state may be mapped to various control schemes, including those relating to a scroll-wheel on mouse.

FIGS. **8**A and **8**B illustrate another example where the user's hand **122** may be used to identify or trigger different activity on the display screen. In this example, FIG. **8**A shows the user pointing directly at the display screen **110** and the computer program being executed to identify this hand orientation on the digital image that was captured by the capture device. When a subsequent capture period occurs at a different point in time, the image of FIG. **8**B might be captured and analyzed by the computer program. In this example, the user's hand is shown tilted from where the thumb **142** is pointing upright to where the thumb **142** is pointing sideways. In this orientation, the user may be able to trigger the shooting of a gun on the display screen, or the selection of an icon on a graphical display.

Consequently, the detection of the user's hand and the orientation of the user's hand can be used to provide the interactivity necessary when analyzing the position of the user's hand relative to the user's head, and the pointing activity on a display screen. The pointing activity will allow the user to select icons, control the device under operation, shoot at graphical objects, select or scroll graphical objects, de-select graphical objects, turn ON and OFF graphical objects, disperse graphical objects, or simply interface with the graphics icons and features of a computer program being displayed on a display screen **110**. However, in certain configurations, it may be desirable, that the system employing the present invention may operate with minimal or no icons on a display screen. Instead, the system may simply just recognize the gestures of the user and provide a control input to the device under operation. For example, a television or related peripheral configured or integrated with the present invention may be controlled by the present invention. Changing a channel, for example, may not necessarily involve interacting with an icon as opposed to recognizing a gesture command according to the scheme presented in the present invention.

FIG. **9** illustrates a flowchart diagram **200** where a process for identifying a pointing direction is described, in accordance with one embodiment of the present invention. In operation **202**, the method begins by providing a capture device for capturing image frames. The frame capture rate will depend on the particular implementation, and will not limit the invention described herein. The capture device will, in one embodiment, include a capture location in coordinate space. As shown in FIG. **4**, the coordinate space of the image capture device is (0, 0, 0). In operation **204**, a display screen is provided for rendering and displaying the interactive graphics of a computer program.

The interactive graphics of the computer program may be associated with a computer game and with any other program or interface that may require interactivity by a user. For instance, the interactivity may include the selection or de-selection of objects, opening files, changing channels, recording TV shows, closing files, accessing files over a network, or interactively communicating with users by way of the interne, electronic mail or by electronic video mail, selecting a consumer electronic device, turning a device ON or OFF. Next, the operation **206** will include the capturing of an image that is presented substantially in front of the image capture device. In one embodiment, the image will include a person that is part of the captured space, and who is in front of the display screen and the capture device.

Once the image has been captured in operation **206**, operation **208** will include the identification of a human head of the person that is found in the captured image. The identified human head will therefore be analyzed to ascertain the head location in the coordinate space relative to the capture location. The method then moves to operation **210** where an object held by the person in the image is identified. The object's location is identified such that the coordinate space of the object is identified relative to the coordinate space of the capture location. Having the identified head location and the identified object location in memory, at operation **212**, the computer program can identify a relative position in coordinate space between the head location and the object location when viewed from the capture location reference point (e.g., coordinate (0,0,0)). As mentioned above, an azimuth angle and an altitude angle can be computed for the relative locations of the head and the hand relative to the image capture device. This relative position in coordinate space is calculated for the captured frame.

In operation **214**, a pointing direction is established for the object using the relative position identified between the object location and the head location. The pointing direction is displayed on the display screen to enable interaction with the interactive graphics provided by the computer program and displayed on the display screen.

FIG. **10** illustrates a more detailed process diagram **250** that can be implemented when determining the pointing direction of an object that is directionally pointed at a display screen during an interactivity with a computer program. The method begins at operation **252** where the capture device for capturing image frames is provided. The capture device will have a capture location in coordinate space. The coordinate space of the capture location will be the reference point for performing operations to determine relative locations in the process of identifying directionality pointing.

The method moves to operation **254** where a display screen is provided for rendering interactive graphics of the computer program. The interactive graphics may be a computer game or may be any other program as defined above. In operation **256**, an image is captured in front of the image capture device and a person is captured in the image. The captured image may be that of a digital frame of video. In one embodiment, the digital frame of video may be a JPEG frame or may be part of a compressed video frame (e.g., MPEG or the like).

Next, the operation moves to the identification of a human head of the person in the captured image in operation **258**. The human head is analyzed on the captured image to determine a head location and the coordinate space relative to the image capture device. In operation **260**, the method moves to the identification of an object held by the person in the image and determining an object location in the coordinate space. In operation **262**, a relative position is identified in the coordinate space between the head location and the object location when viewed from the capture location of the capture device. The relative position will include a calculation of an azimuth angle and an altitude angle relative to the image capture device.

In operation **264**, during execution of the computer program, the operations identified as A, B, C and D corresponding to operations **256**, **258**, **260**, and **262** will be performed iteratively and continuously, depending on a rate desired for the performance of a computer program. For instance, the execution of operations A through D will occur at the rate of one time for each frame that is captured or only after a certain number of frames are captured. The rate at which operations A through D are performed will therefore depend on the specific environment and the need for accurate detection of the pointer location and the select ability of interactive objects on the display screen. If the display screen is processing a video game that has objects that are moving at rapid rates, the tracking operation may require that operations A through D be performed for each frame that is displayed on the video display screen.

In operation **266**, the method indicates a continual update of the pointing direction of the object using the relative position. The pointing direction is displayed on the display screen to enable interaction with the interactive graphics of the computer program. It should again be understood that the pointing direction can be to enable a user to select icons, de-select icons, move icons, open objects, open files, save files, move files, and interact with files that may be part of a file database, or part of a graphical user interface on a computer desktop or the like.

FIG. **11** is a block diagram of an exemplary user input system for interaction with an object on a graphical display that can be used to implement embodiments of the present invention. As shown in FIG. **11**, the user input system is comprised of a video capture device **300**, an input image processor **302**, an output image processor **304**, and a video display device **306**. Video capture device **300** may be any device capable of capturing sequences of video images, and, in one embodiment, is a digital video camera (such as a "web-cam"), or similar image capturing device. As mentioned above, the video capture device may be configured to provide depth image. Input image processor **302** translates the captured video images of the control object into signals that are delivered to an output image processor. In one embodiment, input image processor **302** is programmed to isolate the control object from the background in the captured video image through the depth information and generate an output signal responsive to the position and/or movement of the control object. The output image processor **304** is programmed to effect translational and/or rotational movement of an object on the video display device **306** in response to signals received from the input image processor **302**.

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 **302** and an output image processor **304**. 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. **11** are shown only for convenience of description.

FIG. **12** is a simplified block diagram of a computer processing system configured to implement the embodiments of the invention described herein. The processing system may represent a computer-based entertainment system embodiment that includes central processing unit ("CPU") **424** coupled to main memory **420** and graphical processing unit ("GPU") **426**. CPU **424** is also coupled to Input/Output Processor ("IOP") Bus **428**. In one embodiment, GPU **426** includes an internal buffer for fast processing of pixel based graphical data. Additionally, GPU **426** 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 display device **427** connected external to the entertainment system 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.

IOP bus **428** couples CPU **424** to various input/output devices and other busses or device. IOP bus **428** is connected to input/output processor memory **430**, controller **432**, memory card **434**, Universal Serial Bus (USB) port **436**, IEEE1394 (also known as a Firewire interface) port **438**, and bus **450**. Bus **450** couples several other system components to CPU **424**, including operating system ("OS") ROM **440**, flash memory **442**, sound processing unit ("SPU") **444**, optical disc controlling **4**, and hard disk drive ("HDD") **448**. In one aspect of this embodiment, the video capture device can be directly connected to IOP bus **428** for transmission therethrough to CPU **424**; where, data from the video capture device can be used to change or update the values used to generate the graphics images in GPU **426**. Moreover, embodiments of the present invention can use a variety of image processing configurations and techniques, such as those described in 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 hereby incorporated by reference in its entirety. The computer processing system may run on a CELL™ processor.

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 **448**, flash memory **442**, OS ROM **440**, or on memory card **432**. Alternatively, the program can be downloaded to the processing unit through one or more input ports coupled to CPU **424**. The program modules defining the input method can be provided with the game or application program that is executed by CPU **424** and displayed on display device **427** or they may be provided separately from the application program, such as for execution from local main memory **420**.

In still another embodiment, the program may be executed partially on a server connected to the internet and partially on the local computer (e.g., game console, desktop, laptop, or wireless hand held device). Still further, the execution can be entirely on a remote server or processing machine, which provides the execution results to the local display screen. In this case, the local display or system should have minimal processing capabilities to receive the data over the network (e.g., like the Internet) and render the graphical data on the screen. The user's input, by way of the capture device can be communicated back to the server and then the response represented on the screen.

FIGS. **13A-13C** illustrate embodiments where depth data is taken into consideration in order to better identify the object used to perform the directional pointing. The object can be something the person is holding or can also be the person's hand. In the this description, the terms "depth camera" and "three-dimensional camera" 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" refers to any type of device that is capable of obtaining distance information as well as two-dimensional pixel information.

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.

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.

FIG. **13**A is a block diagram of an exemplary system **500** for providing a real-time three-dimensional interactive environment, in accordance with an embodiment of the present invention. As shown in FIG. **13**A, the system **500** includes a depth camera **502**, an input image processor **504**, an output image processor **506**, and a video display device **508**.

As mentioned above, the depth camera **502** provides the ability to capture and map the third-dimension in addition to normal two-dimensional video imagery. FIGS. **13**B and **13**C illustrated the images generated by a typical depth camera **502**. In particular, FIG. **13**B is an illustration showing two-dimensional data **520** 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. **13**B, the exemplary scene includes a cylinder object **522** and a sphere object **524** disposed on a table **526**, which may be situated among hills **528**.

However, unlike a conventional camera, a depth camera also captures depth values for the scene. FIG. **13**C is an illustration showing depth data **550** captured using a typical depth camera. As illustrated in FIG. **13**B, 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. **13**C, 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.

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. **13**C, z-values are illustrated for the cylinder object **552**, the sphere object **554**, and part of the table **556**. In addition, a maximum detection range is defined beyond which depth values will not be detected. For example, in FIG. **13**C the maximum depth range **558** 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. For more information on depth analysis, reference may be made to U.S. patent application Ser. No. 10/448,614, entitled System and Method for Providing a Real-time three dimensional interactive environment, having a filing date of May 29, 2003, which is incorporated herein by reference.

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. The embodiments described herein are 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.

With the above embodiments in mind, it should be understood that the invention may employ various computer-implemented operations involving data stored in computer systems. These operations include operations requiring physical manipulation of physical quantities. Usually, though not necessarily, these quantities take the form of electrical or magnetic signals capable of being stored, transferred, combined, compared, and otherwise manipulated. Further, the manipulations performed are often referred to in terms, such as producing, identifying, determining, or comparing.

The above described invention may be practiced with other computer system configurations including hand-held devices, microprocessor systems, microprocessor-based or programmable consumer electronics, minicomputers, mainframe computers and the like. The invention may also be practiced in distributing computing environments where tasks are performed by remote processing devices that are linked through a communications network.

The invention can also be embodied as computer readable code on a computer readable medium. The computer readable medium is any data storage device that can store data which can be thereafter read by a computer system, including an electromagnetic wave carrier. Examples of the computer readable medium include hard drives, network attached storage (NAS), read-only memory, random-access memory, CD-ROMs, CD-Rs, CD-RWs, magnetic tapes, and other optical and non-optical data storage devices. The computer readable medium can also be distributed over a network coupled computer system so that the computer readable code is stored and executed in a distributed fashion.

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 method for detecting direction conveyed by a user when interfacing with a computer program executed on a computing device, the computing device being interfaced with an image capture device, comprising:

(a) detecting a human head of a person in an image taken with the image capture device and assigning the human head a head location;

(b) detecting an object held by the person in the image and assigning the object an object location;

(c) determining a relative position in space between the head location and the object location from a capture location of the image capture device, the relative position used to identify a pointing direction of the object;

(d) displaying the pointing direction of the object on a display screen; and

(e) updating the pointing direction of the object on the display screen based on a change in the relative position between the head location and the object location, wherein the pointing direction of the object is updated on the display screen when the head location changes independent of a change in the object location,

wherein the method is executed by one or more processors.

2. The method for detecting direction conveyed by a user when interfacing with a computer program executed on a computing device as recited in claim **1**, further comprising:

capturing a plurality of frames through the image capture device, each of the frames being analyzed to process (a)-(c); and

determining changes in the pointing direction of the object over the plurality of analyzed frames.

3. The method for detecting direction conveyed by a user when interfacing with a computer program executed on a computing device as recited in claim **2**, wherein the determined changes in the pointing direction of the object are translated to detected motion of the pointing direction of the object over the plurality of frames captured through the image capture device.

4. The method for detecting direction conveyed by a user when interfacing with a computer program executed on a computing device as recited in claim **1**, wherein the pointing direction of the object is translated as a pointing location on the display screen, the display screen being located below the location of the image capture device.

5. The method for detecting direction conveyed by a user when interfacing with a computer program executed on a computing device as recited in claim **1**, wherein the relative position is identified by computing a first angle and a second angle between the head location and the object location in relation to the capture location.

6. The method for detecting direction conveyed by a user when interfacing with a computer program executed on a computing device as recited in claim **1**, wherein the display screen is configured to render interactive graphics.

7. The method for detecting direction conveyed by a user when interfacing with a computer program executed on a computing device as recited in claim **1**, further comprising:

enabling interactivity with particular interactive graphics presented on the display screen using a graphic indicator of the pointing direction of the object.

8. The method for detecting direction conveyed by a user when interfacing with a computer program executed on a computing device as recited in claim **7**, wherein the interactivity includes selection of a graphic, shooting of a graphic, touching a graphic, moving of a graphic, activation of a graphic, triggering of a graphic, acting upon or with a graphic, or a combination thereof.

9. The method for detecting direction conveyed by a user when interfacing with a computer program executed on a computing device as recited in claim **1**, wherein detecting the human head is processed using face detection.

10. The method for detecting direction conveyed by a user when interfacing with a computer program executed on a computing device as recited in claim **1**, wherein detecting the object is facilitated using colors, lights, or a combination thereof, or a combination of identifying differences in colors, lights, and shapes, or identifying on or off states.

11. The method for detecting direction conveyed by a user when interfacing with a computer program executed on a computing device as recited in claim **1**, wherein the computer program is a video game.

12. The method for detecting direction conveyed by a user when interfacing with a computer program executed on a computing device as recited in claim **1**, wherein the computer program is either local to the image capture device or networked over the internet.

13. A method for detecting direction conveyed by a user when interfacing with a computer program executed on a computing device, the computing device being interfaced with an image capture device, comprising:

(a) detecting a human head of a person in an image taken with the image capture device and assigning the human head a head location;

(b) detecting a body part of the person, other than the human head, in the image and assigning the body part a body part location;

(c) determining a relative position in space between the head location and the body part location from a capture location of the image capture device, the relative position used to identify a pointing direction of the body part;

(d) displaying the pointing direction of the body part on a display screen; and

(e) updating the pointing direction of the body part on the display screen based on a change in the relative position between the head location and the body part location, wherein the pointing direction of the body part is updated on the display screen when the head location changes independent of a change in the body part location,

wherein the method is executed by one or more processors.

14. The method for detecting direction when interfacing with a computer program executed on a computing device as recited in claim **13**, further comprising:

capturing a plurality of frames through the image capture device, each of the frames being analyzed to process (a)-(c); and

determining changes in the pointing direction of the body part over the plurality of analyzed frames.

15. The method for detecting direction conveyed by a user when interfacing with a computer program executed on a computing device as recited in claim **14**, wherein the determined changes in the pointing direction of the body part are translated to detected motion of the pointing direction of the body part over the plurality of frames captured through the image capture device.

16. The method for detecting direction conveyed by a user when interfacing with a computer program executed on a computing device as recited in claim **13**, wherein the pointing direction of the body part is translated as a pointing location on the display screen, the display screen being located below the location of the image capture device.

17. The method for detecting direction conveyed by a user when interfacing with a computer program executed on a computing device as recited in claim **13**, wherein the relative position is identified by computing a first angle and a second angle between the head location and the body part in relation to the capture location.

18. The method for detecting direction conveyed by a user when interfacing with a computer program executed on a computing device as recited in claim **13**, further comprising enabling interactivity with particular interactive graphics presented on the display screen using a graphic indicator of the pointing direction of the body part.

19. The method for detecting direction conveyed by a user when interfacing with a computer program executed on a computing device as recited in claim **18**, wherein the interactivity includes selection of a graphic, shooting of a graphic, touching a graphic, moving of a graphic, activation of a graphic, movement of a graphic, triggering of a graphic, acting upon or with a graphic, or a combination thereof.

20. The method for detecting direction conveyed by a user when interfacing with a computer program executed on a computing device as recited in claim **13**, wherein detecting the human head is processed using face detection and detecting the body part is facilitated by identifying shapes.

21. The method for detecting direction conveyed by a user when interfacing with a computer program executed on a computing device as recited in claim **13**, wherein the computer program is a video game or an interactive communication program.

22. The method for detecting direction conveyed by a user when interfacing with a computer program executed on a computing device as recited in claim **13**, wherein the computer program is either local to the image capture device or networked over the internet.

* * * * *