



COMPUTER INPUT DEVICE TRACKING

SIX DEGREES OF FREEDOM



166166

Field of the Invention

[0001] The present invention relates to computer input devices.

[0002] In particular, this invention relates to a computer input device having a plurality of tracking points and an imaging device to allow entry of three-dimensional (3D), six-degrees-of-freedom data into a computer or gaming device.

Background of the Invention

[0003] At present, the main input devices for computers are the keyboard and a pointer device commonly called a "mouse". For computer gaming purposes, additional devices such as joysticks and gaming consoles with specialized buttons are available to give more input choices to the user in the game's virtual environment. However, these are two-dimensional input devices used for entering 3D information. More recently, "computer interface apparels" such as gaming gloves, actually worn by gamers, are used to input their choice of movements into the computer games.

[0004] Despite this plethora of input devices, no satisfactory device exists to accept the desired movements in sword fighting, or racquet, club or bat-swinging movements for certain computer games. Another possible application is that of symphony-conducting baton movements or the waving of a "magic wand" to perform magical tricks in games. The combined use of cursor control keys and rolling of trackballs do not satisfactorily convey the user's intended movements for the virtual object or character he controls in some computer games.

[0005] In addition, other applications beyond the realm of computer games such as computer-aided design and computer-aided manufacturing (CAD/CAM) work, require three-dimensional (3D) input when modifying 3D

renditions of designs. Current input devices do not allow ease of use of such applications.

[0006] Therefore, a need clearly exists for an intuitive, direct and ergonomic input device that enables representation and input of 3D movements by users into computer and gaming devices.

Summary of the Invention

[0007] The present invention seeks to provide an input device with three-dimensional, six-degrees-of-freedom data input for a computer, said input device comprising an arrangement of one array of tracking points to define a first axis, a second array of tracking points to define a second axis or plane orthogonal to the first axis, with distances between the tracking points carefully selected such that a two-dimensional image of the tracking points captured by an imaging device can be used by a processor to determine the position and orientation of the input device in three-dimensional space using a provided algorithm.

[0008] Accordingly, in one aspect, the present invention provides an input device comprising a wand with a cross piece, both wand and cross piece having a plurality of tracking points, at least one imaging device, at least one lighting means, and at least a processor.

Brief Description of the Drawings

[0009] A preferred embodiment of the present invention will now be more fully described, by way of example, with reference to the drawings of which:

[0010] FIG. 1 illustrates an embodiment of the present invention;

[0011] FIG. 2 shows the translation and orientation of the wand from the original condition in which its frame of reference coincides with the imaging device's frame of reference;

[0012] FIG. 3 illustrates how the two main axis of the wand is determined in the iterative process of the algorithm of the present invention; and

[0013] FIG. 4 shows two embodiments of the wand of the present invention.

Detailed Description of the Drawings

[0014] A detailed description of the present invention will now be given in accordance with a preferred embodiment of the invention. In the following description, details are provided to describe the preferred embodiment. It shall be apparent to one skilled in the art, however, that the invention may be practiced without such details. Some of these details may not be described at length so as not to obscure the invention.

[0015] It is an object of the present invention to provide an input device for computer and gaming devices that accept 3D input.

[0016] It is another object of the present invention to provide an input device for computer and gaming devices that allow the reproduction of intended 3D, 6 DOF movements of virtual objects in the virtual environment. This type of movement input may be performed with minimal processor demands so that the present invention can be used in simpler, smaller or less expensive computers and gaming devices and not require too much processor resources from other concurrently-running processes.

[0017] The terms "computer" and "gaming devices" include, but are not limited to, any computing device that require 3D input such as CAD/CAM workstations, "personal computers", dedicated computer gaming consoles and devices, personal digital assistants, and dedicated processors for processing images captured by the present invention.

[0018] There are many advantages of the preferred embodiment of the invention. One advantage of the preferred embodiment is that swordfighting thrusts and parries, golf drives, or racquet and bat swinging movements desired by users in the virtual environment of computer games can now be tracked and represented with the present invention.

[0019] Another advantage is that hitherto unachievable input modalities for applications such as painting and sculpting strokes for virtual artistry, or

even precise manipulations of surgical instruments in telemedicine or surgical training, can now be realized with the present invention.

[0020] The invention includes an input device using an arrangement comprising a first array of tracking points to define a first axis, a second array of tracking points to define a second axis or plane orthogonal to the first axis, with the distances between the points carefully selected such that a two-dimensional image of the tracking points can be used to determine the position and orientation of the input device in three-dimensional space.

[0021] Referring now to FIG. 1, an embodiment in accordance with the present invention is illustrated. The array should not be seen to be limited to the embodiments described. One element of the preferred embodiment of the present invention is the input device is the wand **101**, a longitudinal device that is substantially straight having a handle end **100** and a reference end **103**. The user manipulates the wand by grasping it substantially at the handle end.

[0022] A cross piece **120** is orthogonally disposed near the reference end of the wand. In the preferred embodiment, the plane of the cross piece is substantially planar to the longitudinal axis of the wand. The cross piece of the preferred embodiment is a substantially straight piece of material although cross pieces in the shape of the letter "L" or in form of triangles or polygons may also be used.

[0023] Disposed on the wand, on its reference end, is a first plurality of tracking points. A second plurality of tracking points is also disposed on the cross piece. These tracking points **110** may be active light emitters or passive light reflectors.

[0024] In the case where active light emitters are used for the tracking points, the emitters should be capable of emitting light in all directions. This is so that the tracking points can be captured from all orientations of the wand by an imaging device, an element of the present invention, as long as they are not obscured and are within the field of view of the imaging device.

[0025] The wavelengths of light usable range from infrared wavelengths, through visible light to ultra-violet light.

[0026] In the preferred embodiment, the tracking points are passive infrared reflectors and form a "T" arrangement. The preferred shape of most of these tracking points is that of a globe or ball to maximize reflection of light. The exception is the tracking point of the wand nearest the handle end (the "last" tracking point). Instead of being in a globe or half globe shape, this tracking point is only slightly domed. For most applications, the wand is not more than 30 cm long, a length that caters to a wide range of users. There are at least two tracking points on the wand, with three points being preferable, in a linear arrangement. There are two tracking points at either end of the cross piece which is preferably about 6 cm long (FIG. 1).

[0027] Other elements of the present invention include at least one lighting means **150** providing light in the preferred wavelength range and at least one suitable imaging device **130** such as a camera with complementary metal oxide semiconductor (CMOS) or charge-coupled device (CCD) sensors. In the preferred embodiment, one lighting means and one imaging device is used and the imaging device is connected to a processor.

[0028] In operation, the user grasps the wand of the present invention and wields it as if he is handling the virtual device of choice. The light from the lighting source is reflected from the passive tracking points and is captured as bright spots by the imaging device or camera. The imaging device then in turn transmits the image to the processor.

[0029] The present invention works within the field of view and in line of sight of the imaging device or camera. Thus, the ideal position for both the lighting means and the imaging means is in close proximity to the display screen of the computer or gaming device with both the lighting means and the imaging means pointed substantially at the volume of space in which the user desires to wield the wand.

[0030] To help the user remain in this volume of space covered by the field of view of the imaging means, a pad **160** with a border marking **180** may be used to help the user estimate the field of view of the camera. A circle **170** marked in the center of the pad helps the user to center the wand at the

start of the application or game. The surface material of the pad is also selected to minimize reflection of the light used.

[0031] All transformation parameters relating to the wand's position (translation $T=(t_x, t_y, t_z)$) and orientation such as pitch **201**, roll **203** and yaw **202** are expressed with respect to the frame of reference of the imaging device or camera (X_c, Y_c, Z_c) (FIG. 2).

[0032] The wand's frame (X_w, Y_w, Z_w) is assumed to be aligned with the camera's frame (X_c, Y_c, Z_c) as an original condition. Any point on the wand $w = (x, y, z, 1)$ is transformed to another point $w' = (x', y', z', 1)$ in the camera's space by the following "XYZ fixed angle" transformation:

$$w' = T + R_z R_y R_x w$$

where R_x is the pitch rotation matrix about the X_c axis;

R_y is the yaw rotation matrix about the Y_c axis and

R_z is the roll rotation matrix about the Z_c axis

[0033] The relative positions of each bright spot to each other in each captured image is used by the accompanying algorithm of the present invention to reconstruct the position and orientation of the wand at each time interval. This information is then used to make corresponding movements in the object controlled by the user in the virtual environment.

[0034] Thus, the relative positions of the tracking points of the wand in the image as captured by the camera are used in the present invention to reconstruct the movements of the wand.

[0035] Another piece of information that is used is the distance between the cross piece's two tracking points $d1$ (the value of which is indicated by the line **323** in FIG. 3) and the distance between the two tracking points $d2$ of the wand nearest the cross piece (the value of which is indicated by the line **322**). The ratio of these two apparent distances captured by the camera is used in the reconstruction of the position of the object in the virtual environment represented by the wand.

[0036] The tracking point **210** on the wand nearest the handle end may not be captured by the imaging device, depending on whether the wand is

oriented towards or away from the imaging device. However, the algorithm of the present invention is capable to resolving captured images with either four or five tracking points.

[0037] The four or five tracking points present in the captured image, are used by the algorithm of the present invention to determinate the longitudinal axis and the included angle θ **324** as illustrated in FIG. 3.

[0038] FIG. 3A shows an example of the image of the five tracking points in 3D space captured by the camera in 2D representation. It is possible to join any two points to obtain 10 lines (FIG. 3B)

[0039] The algorithm will begin by comparing a first **320** of these 10 possible lines with another of these 10 possible lines to determine if the first line **320** bisects the other line **321** (FIG. 3C). This process is repeated (an unsuccessful example is shown in FIG. 3D) until such two lines with this relationship are found (FIG. 3E). Thus, the line **322** that bisects another line is the longitudinal axis of the wand. The line **323** being bisected is the line defined by the tracking points of the cross piece, that is, the axis of the cross piece.

[0040] When the two axes have been identified by the algorithm, the intersection of the two axes defines the position of the reference origin **104** on the wand (FIG. 2), the junction of the wand and its cross piece. This is represented by the position **326** on the captured image.

[0041] In addition, the included angle θ of these two axes as well as the ratio of d_1 and d_2 as captured by the camera can also be determined. From line **323** and the x-axis of the image, δ **325**, the angle formed by the projected axis of the cross piece with the horizontal axis of the image can also be determined. The degree of roll can then be directly assigned by the angle δ .

[0042] Concomitantly, the translations t_x and t_y can be directly interpreted from the position of the wand's reference origin **326** in the captured image.

[0043] The last bit information required is whether the last tracking point **210** on the wand (nearest the handle end) can be seen in the captured image.

This information is represented by P_5 which is 1 when the point is present and 0 when the point is absent.

[0044] Once these essential parameters are known, the wand's position may then be reconstructed or obtained from the 6 DOF information stored in a prepared lookup table that performs the mapping:

$$\langle \theta, d_1/d_2, P_5 \rangle \rightarrow \langle \text{pitch, yaw, } l_1, l_2 \rangle$$

where l_1 and l_2 are the projected distances of either:

- a. the two tracking points of the cross piece, if $d_1/d_2 \geq 1$; or
- b. the distance of the fourth tracking **300** on the wand from the wand's origin **326**, if $d_1/d_2 < 1$;

in the captured image when the wand's origin is at distances from the camera at the lower limit z_l and the upper limit z_u respectively, of the usable operating range from the camera, and where the wand's orientation is at the specific pitch and yaw given in the mapping above.

[0045] The usable operating range thus defined is important as the lookup table has been generated based on this usable operational range from the camera. The usable operational range do not limit the present invention as the operating range can be varied and scaled to different applications with accompanying lookup tables. For most desktop applications, the two limits z_l and z_u are set at 25 cm and 50 cm respectively.

[0046] The distance t_z , the distance of the translated wand's origin from the camera can then be approximated from the following formula:

$$t_z = z_l + ((z_u - z_l) \times (l_1 - d_1)/(l_1 - l_u)), \text{ if } d_1/d_2 \geq 1; \text{ or}$$

$$t_z = z_l + ((z_u - z_l) \times (l_1 - d_2)/(l_1 - l_u)), \text{ if } d_1/d_2 < 1$$

[0047] With the present invention, tracking user movements of the wand simply entails capturing the relative positions of the tracking points, determining the essential parameters of the wand's position and orientation and then reconstructing the movement in the virtual environment concerned. This process is repeated at desired time intervals for as long as the present invention is in use.

[0048] It will be appreciated by one skilled in the art that there will be positions in which the other objects such as the part of the user's body or the wand itself may mask one or more of the tracking points from the line of sight of the camera. When this happens, a simple algorithm may not be able to accommodate such situations. To overcome this limitation, any driver software implementing this algorithm can be programmed to predict the intended trajectory or movement of the wand so as to ignore such situations and to resume processing when sufficient tracking points are detected again. With continual motion of the user's hand, such situations are not expected to occur with sufficient frequency to even cause any annoyance with users.

[0049] The minimum number of tracking points for the preferred embodiment is four. In the case when the minimum number of tracking points cannot be captured by the imaging means after a predetermined number of time intervals, the driver software used would alert the user that an intolerable obscuration of the tracking points has occurred.

[0050] Such situations may arise when the wand is out of the field of view of the camera or prolonged obscuration by some object. The pad 160 of FIG. 1 will help avoid such situations by reminding the user to keep the wand within the field of view of the imaging device.

[0051] The distances between the tracking points have been carefully determined to minimize ambiguity when interpreting them on the images captured by the imaging device so that the position and orientation of the wand can be rapidly reconstructed.

[0052] Another possible way to reconstruct the position of the wand from the relative positions of the tracking points in the captured image is by standard 3D reconstruction algorithms found in most textbooks on robotic vision and computer graphics techniques. However such algorithms are computationally expensive and discourage real-time operations in typical computers. As such, implementation of such algorithms is also not desirable at the time of this application as it goes against the aim of the present invention of implementation in simple computing devices. As such,

a look up table is used in the interim to provide a reasonably fast response time.

[0053] Despite this, the elements of the present invention will still work with algorithms that reconstruct 3D movements when processor capabilities are able to accommodate such algorithms.

[0054] In the preferred embodiment, the lighting source is a ring light emitting light substantially in 750-1000 nm, mounted in a co-axial arrangement around the lens of the camera.

[0055] If light of a narrow wavelength range is used, suitable materials such as reflectors for the tracking points, filters for the lighting means, and lenses for the imaging device will also be required for optimal results. As the preferred embodiment uses light in the infrared range, the tracking points has coatings such as "Laser Gold" plating optimized for maximal reflectance at this range of wavelengths.

[0056] While a grayscale or color camera is usable under the present invention, cameras optimized to represent images using 1 bit per pixel (that is, in strict black and white representation) is preferred as data transmission requirements are minimized.

[0057] It will be appreciated by one skilled in the art that many embodiments are possible under the scope of the present invention. For example, the cross piece may be a V or L-shaped piece (FIG. 4A), equilateral triangle (FIG. 4B) or regular polygon. The present invention will work as long as the tracking points of the cross piece all lie on a plan perpendicular to the longitudinal axis of the wand.

[0058] Each tracking point of the present invention may also emit or reflect different unique colors or wavelengths distinguishable by the imaging device. In such an embodiment, speed of reconstruction of the wand's position may be greatly increased as it will be easier and faster to distinguish the essential parameters of the wand's position, when used with a color imaging device.

[0059] In this embodiment, the lighting source may illuminate with visible light and the reflective coating of the tracking points are selected to reflect

the color of choice. Alternatively, small radiation sources such as light-emitting diodes (LEDs) may be used in embodiments with active, light-emitting tracking points.

[0060] The present invention is also scalable, depending on the application. The sensitivity of the present invention can be increased, say for tracking a surgeon's hand, by bringing the imaging device nearer the wand or by using imaging devices with higher resolution.

[0061] Alternatively, more gross movements such as the swing of a modified racquet, bat or club may be captured by moving the camera further away. Capturing of images from these sports equipment may be done by placing tracking points on them or on representative simulated equipment.

[0062] As technology advances, the refresh rate for capturing the images and the sensitivity of tracking the movements, may also be enhanced as processor speeds and capabilities increase.

[0063] While the use of one imaging device has been described, it is possible to implement the present invention with more than one imaging device. If more than one imaging devices are used, triangulation techniques may be used to better reconstruct the distance of wand from the imaging devices.

[0064] It will be appreciated by one skilled in the art that the algorithm described can be readily implemented in a variety of programming languages at various levels from high level programming languages, to assembly languages to machine languages. The integration of the present invention to various gaming and CAD/CAM applications may also be readily performed by one skilled in the art.

[0065] The present invention therefore provides an input device comprising a wand with a cross piece, a lighting means, an imaging device, a processor and an algorithm and lookup table to allow users to overcome, or at least alleviate, the problems of the prior art.

[0066] Although in the preferred embodiment, the processor used is that of the computer to which the imaging device is connected to, a separate,

dedicated processor for interpreting the captured images to obtain the position and orientation of the wand may also be used under the scope of the present invention.

[0067] In addition, while the present invention has capability to determine the 3D, 6 DOF input data, it will be apparent to one skilled in the art that the present invention can also be readily modified to track either 3DOF input data alone or 6 DOF input data alone, as desired for other applications.

[0068] It will be appreciated that although one preferred embodiment has been described in detail, various modifications and improvements can be made by a person skilled in the art without departing from the scope of the present invention.

Claims

1. An input device with three-dimensional, six-degrees-of-freedom data input for a computer, said input device comprising:
an arrangement of a first array of tracking points to define a first axis, a second array of tracking points to define a second axis or plane orthogonal to the first axis, with distances between the tracking points carefully selected such that a two-dimensional image of the tracking points can be used to determine the position and orientation of the input device in three-dimensional space using a provided algorithm;
at least one imaging device; and
at least one processor.
2. The input device of Claim 1, said provided algorithm further comprising:
identifying positions and orientations of the tracking points from a captured two-dimensional image;
identifying the first axis and the second axis or plane from the captured image;
determining essential parameters of the arrangement from the identified first axis and second axis or plane;
obtaining position and orientation of the input device from a lookup table;
and
reconstructing the three-dimensional position and orientation of the input device for the computer.
3. The input device of Claim 1, said input device capable of being modified to only determine its own position in three-dimensional space.
4. The input device of Claim 1, said input device capable of being modified to only determine its own pitch, roll and yaw in three-dimensional space.

5. The input device of Claim 1, said input device further comprising at least one lighting means.
6. A computer input device providing three-dimensional, six-degrees-of-freedom data input for a computer, said input device comprising:
a wand, said wand further comprising a cross piece orthogonally-disposed with longitudinal axis of the wand; the wand and cross piece further comprising a plurality of tracking points;
at least one imaging device;
at least one processor; and
an algorithm able to determine position and orientation of the wand from two-dimensional images of tracking points on the wand captured by the at least one imaging device from a lookup table.
7. The input device of Claim 6, wherein said cross piece is disposed perpendicular to the wand.
8. The input device of Claim 6, the plurality of tracking points of the wand comprising passive light reflectors and where said input device further comprising at least one lighting means.
9. The input device of Claim 8, wherein said at least one lighting means is in a ring form co-axially disposed around lens of said imaging device.
10. The input device of Claim 6, the plurality of tracking points of the wand comprising active light emitters.
11. The wand of Claim 10, said plurality of tracking points comprising active light emitters for different wavelengths.

12. The input device of Claim 6, wherein distances between the plurality of tracking points are carefully selected to avoid ambiguity in the determination of the position and orientation of the wand by the algorithm.
13. The input device of Claim 6, wherein the imaging device further comprising zero or more filters for selected wavelength range of light used.
14. The input device of Claim 6, said input device further comprising a pad to help a user estimate the field of view of said imaging device.