

(19) World Intellectual Property Organization
International Bureau



(43) International Publication Date
12 April 2007 (12.04.2007)

PCT

(10) International Publication Number
WO 2007/041690 A2

- (51) International Patent Classification:
H04N 13/02 (2006.01) *H04N 13/04* (2006.01)
- (21) International Application Number:
PCT/US2006/039066
- (22) International Filing Date: 4 October 2006 (04.10.2006)
- (25) Filing Language: English
- (26) Publication Language: English
- (30) Priority Data:
60/723,903 4 October 2005 (04.10.2005) US
- (71) Applicant and
(72) Inventor: **ALEXANDER, Eugene, J.** [US/US]; 717
Avenida Presido, San Clemente, CA 92672 (US).

AT, AU, AZ, BA, BB, BG, BR, BW, BY, BZ, CA, CH, CN, CO, CR, CU, CZ, DE, DK, DM, DZ, EC, EE, EG, ES, FI, GB, GD, GE, GH, GM, HN, HR, HU, ID, IL, IN, IS, JP, KE, KG, KM, KN, KP, KR, KZ, LA, LC, LK, LR, LS, LT, LU, LV, LY, MA, MD, MG, MK, MN, MW, MX, MY, MZ, NA, NG, NI, NO, NZ, OM, PG, PH, PL, PT, RO, RS, RU, SC, SD, SE, SG, SK, SL, SM, SV, SY, TJ, TM, TN, TR, TT, TZ, UA, UG, US, UZ, VC, VN, ZA, ZM, ZW.

(84) Designated States (unless otherwise indicated, for every kind of regional protection available): ARIPO (BW, GH, GM, KE, LS, MW, MZ, NA, SD, SL, SZ, TZ, UG, ZM, ZW), Eurasian (AM, AZ, BY, KG, KZ, MD, RU, TJ, TM), European (AT, BE, BG, CH, CY, CZ, DE, DK, EE, ES, FI, FR, GB, GR, HU, IE, IS, IT, LT, LU, LV, MC, NL, PL, PT, RO, SE, SI, SK, TR), OAPI (BF, BJ, CF, CG, CI, CM, GA, GN, GQ, GW, ML, MR, NE, SN, TD, TG).

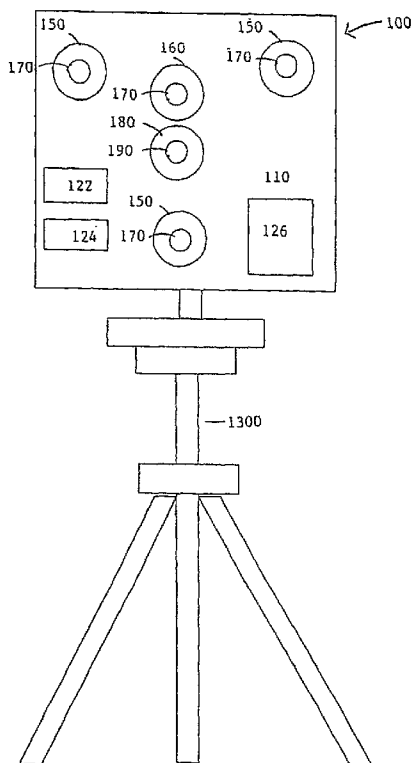
Published:
— without international search report and to be republished upon receipt of that report

For two-letter codes and other abbreviations, refer to the "Guidance Notes on Codes and Abbreviations" appearing at the beginning of each regular issue of the PCT Gazette.

- (74) Agents: **MERKADEAU, Pamela, S.** et al.; MANATT, PHELPS & PHILLIPS, LLP, 1001 Page Mill Road, Building 2, Palo Alto, CA 94304 (US).
- (81) Designated States (unless otherwise indicated, for every kind of national protection available): AE, AG, AL, AM,

(54) Title: DEVICE FOR GENERATING THREE DIMENSIONAL SURFACE MODELS OF MOVING OBJECTS

(57) Abstract: A device and method for producing a model of the surface of one aspect of a three dimensional object. The model is produced by extracting a three dimensional mesh from a pattern projected on the object. The three dimensional object maybe in motion.



WO 2007/041690 A2

DEVICE FOR GENERATING THREE DIMENSIONAL SURFACE MODELS OF MOVING OBJECTS

BACKGROUND OF THE INVENTION

Field of the Invention

[0001] The invention relates generally to apparatus and methods for generating three dimensional surface models of moving objects.

Background of the Invention

[0002] The generation of three dimensional models of moving objects has uses in a wide variety of areas, including motion pictures, computer graphics, video game production, human movement analysis, orthotics, prosthetics, surgical planning, sports medicine, sports performance, product design, surgical planning, surgical evaluation, military training, and ergonomic research.

[0003] Two existing technologies are currently used to generate these moving 3D models. Motion capture techniques are used to determine the motion of the object, using retro-reflective markers such as those produced by Motion Analysis Corporation, Vicon Ltd., active markers such as those produced by Charnwood Dynamics, magnetic field detectors such as those produced by Ascension Technologies, direct measurement such as that provided by MetaMotion, or the tracking of individual features such as that performed by Peak Performance, Simi. While these various technologies are able to capture motion, nevertheless these technologies do not produce a full surface model of the moving object, rather, they track a number of distinct features that represent a few points on the surface of the object.

[0004] To supplement the data generated by these motion capture technologies, a 3D surface model of the static object can be generated. For these static objects, a number of technologies can be used for the generation of full surface models: laser scanning such as that accomplished

by CyberScan, light scanning such as that provided by Inspeck, direct measurement such as that accomplished by Direct Dimensions, and structured light such as that provided by Eyetronics or Vitronic).

[0005] While it may be possible to use existing technologies in combination, only a static model of the surface of the object is captured. A motion capture system must then be used to determine the dynamic motion of a few features on the object. The motion of the few feature points can be used to extrapolate the motion of the entire object. In graphic applications, such as motion pictures or video game production applications, it is possible to mathematically transform the static surface model of the object from a body centered coordinate system to a global or world coordinate system using the data acquired from the motion capture system.

[0006] All of these surface generation systems are designed to operate on static objects. Even when used in combination with a motion capture system, as described above, an object that is not a strictly rigid body is not correctly transformed from a body centered coordinate system, as a single static surface model does not adequately represent the non rigid motion of the object. Therefore, there exists a need for a systems and methods that can produce a model of the surface a three dimensional object, with the object possibly in motion and the object possibly deforming in a non-rigid manner.

[0007] A device is needed that has the capability to operate on an object that is moving relative to the device while the device itself is possibly in motion. In order to achieve this goal, a novel device is provided wherein the optical characteristics of the device are under the control of an operator. In one embodiment a device that projects a pattern onto one aspect of the object is disclosed. The projection is in focus regardless of the distance between the object and the device, within a given range. In addition, the size of the pattern projected onto the object is under operator control. Furthermore, the imaging components of the device may also be under operator control, so that the zoom, focus, and aperture of the imaging components can be set to operate optimally, depending on where the object is in relation to the device.

BRIEF DESCRIPTION OF THE DRAWINGS

[0008] The drawings illustrate the design and utility of preferred embodiments of the invention, in which similar elements are referred to by common reference numerals and in which:

[0009] FIG. 1 is a front view of an imaging device.

[00010] FIG. 2 is a side view of the imaging device of FIG. 1.

[00011] FIG. 3 illustrates an imaging device in operation.

[0010] FIG. 4 illustrates an imaging device in operation with a projector activated.

[0011] FIG. 5 illustrates an image as observed by grey scale cameras.

[0012] FIG. 6 illustrates a reconstructed three dimensional surface mesh of a subject.

DETAILED DESCRIPTION

[0013] Various embodiments of the invention are described hereinafter with reference to the figures. It should be noted that the figures are not drawn to scale and elements of similar structures or functions are represented by like reference numerals throughout the figures. It should also be noted that the figures are only intended to facilitate the description of specific embodiments of the invention. The embodiments are not intended as an exhaustive description of the invention or as a limitation on the scope of the invention. In addition, an aspect described in conjunction with a particular embodiment of the invention is not necessarily limited to that embodiment and can be practiced in any other embodiment of the invention.

[0014] Turning to the drawings, FIG. 1 is a front view of an imaging device 100. The imaging device is supported by a stand 1300. The imaging device 100 has a mounting panel 110. Contained within the mounting panel 110 of the imaging device 100 are grey scale digital video cameras 150. As illustrated in the present embodiment, there are three grey scale digital video cameras 150, however this number is not intended to be a limitation of the number of grey scale digital video cameras that may be contained in the mounting panel 110. There may be as few as two grey scale digital video cameras 150 and as many grey scale digital video cameras 150 as can be mounted on the mounting panel 110. The more digital video cameras that are incorporated, the more detailed the model generated is. The grey scale digital video cameras 150 may be time synchronized. The grey scale digital video cameras will be used in pairs to generate a 3D surface mesh of the subject. The mounting panel 110 may also contain a color digital video camera 160. The color digital video camera may be used to supplement the 3D surface mesh generated by the grey scale camera pair with color information. This color information allows for the direct generation of three dimensional models with color information derived from the actual object.

[0015] Each of the video cameras 150, 160 have lenses 170 with electronic zoom, aperture and focus control. Also contained within the mounting panel 110 is a projection system 180.

The projection system 180 has a lens 190 with zoom and focus control. The projection system 180 allows an image, generated by the imaging device, to be cast on the object of interest, such as an actor or an inanimate object. The projection system could be a slide projector, which includes a light source, slide holder, slide or patterned glass, and lens(es), a liquid crystal display (LCD) projector, which includes a light source, one or three liquid crystal displays, and lens(es), a digital light processing (DLP) projector, which includes a light source, an array of micro-mirrors, and lens(es), or any other type of projector that allows an image to be projected on an object.

[0016] Control signals are transmitted to the imaging device 100 through a communications channel 122. Data is downloaded from the imaging device 100 through another communications channel 124. Power is distributed to the imaging device 100 through a power system 126. The power system 126 may be internal to the imaging device 100 or an external power system may be used. The imaging device 100 is controlled by a computer (not shown). The computer may be a remote computer for example a laptop, desktop or workstation, or the imaging device 100 may have an internal computer.

[0017] FIG. 2 is a side view of the imaging device 100 of FIG. 1. The imaging device 100 has orientation motors (not shown) that control the roll, pitch and yaw of the imaging device 100. One orientation motor may control multiple placements of the imaging device, or separate orientation motors may be utilized to control the various movements of the imaging device 100. As illustrated, an orientation motor controls the rotation of the imaging device 100 about its y-axis 220. The rotation about the y-axis is also known as the roll of the imaging device 100. Another orientation motor controls the rotation 230 of the imaging device 100 about its z-axis; this movement is known as the pitch of the imaging device 100. Still another orientation motor may control the rotation 240 of the imaging device 100 about its x-axis. The movement of the imaging device 100 about its x-axis is known as the yaw of the imaging device 100.

[0018] FIG. 3 shows a subject 310 within the viewing range of an imaging device 320. The subject 310 is shown moving, alternatively, the subject 310 may be a moving object. For each frame, multiple control signals are sent to the imaging device 320. These control signals specify the roll, pitch, and yaw of the imaging device 320. Control signals also specify the zoom, focus, and aperture for each of the gray scale digital video cameras 350, and the zoom, focus, and aperture for the color digital video camera if one is included in the imaging device (not shown). Still another control signal specifies the zoom, focus, and aperture for the projection system 380. In addition the pattern to be projected by the projection system is also specified by a control signal.

[0019] After the imaging unit 320 is positioned and digital video cameras 320 and projection system 380 zoom, focus and aperture are set, a trigger signal is sent through the control channel. The trigger signal causes the projector to turn on, projecting the desired pattern 420 out into the viewing volume and onto a subject 420 as illustrated in FIG. 4. A delayed trigger signal is then sent to each of the digital video cameras 450 and each of the digital video cameras captures an image. The image will be of the three dimensional object with the projected pattern superimposed on the object as illustrated in FIG. 5. These images are then transmitted from the imaging unit to a computer. This completes a single frame and the imaging device is now is now ready for the next cycle of operation. The process of capturing data continues, depending on the desired application. For example, if a model of the subject is being built, the camera would be moved around the subject, acquiring data, until a complete mesh was acquired. If the application involves tracking the subject in real-time, the data acquisition would run continuously. If the activity is over some specified time interval, say for example a scene, the imaging device would record until stopped by the operator.

[0020] A computer, either external or internal to the imaging device as described above, reads these images of the object with the superimposed pattern. For each of the various camera images, the pattern features are detected. One approach that may be used for the pattern

detection is a mesh extraction technique such as that described in Bayesian Grid Matching, *IEEE Transactions on Pattern Analysis and Machine Intelligence*, Vol. 25, No. 2, February 2003, pp. 162 – 173 by K. Hartelius, and J.M. Carstensen, which is hereby incorporated by reference in its entirety. In this algorithm, the projection pattern is a grid of black lines on a clear background. Since the pattern is known, it is possible using computer vision techniques to quickly extract these features.

[0021] Once the features are extracted from all of the digital video cameras on the imaging device, the features can be corresponded (that is, features on the object observed in one camera image are related to the same features as observed in another camera) across the digital video cameras. Corresponding the features can be accomplished in a variety of ways. One such approach is to make a number of the features distinct from the rest of the feature field. These features are detected in each of the digital video cameras and serve as a template for corresponding the rest of the features. Another approach uses a technique known as maximization of mutual information as detailed in "Multi-modality image registration by maximization of mutual information", *Mathematical Methods in Biomedical Image Analysis*. IEEE Computer Society Press, 1996, by A. Maes, et al., (the "Maes technique") which is incorporated herein by reference in its entirety. Using the Maes technique, equation 1 below is calculated for each of the possible correspondence patterns.

$$MI(X, Y) = H(X) + H(Y) - H(X, Y) \quad (1)$$

[0022] Where MI is the mutual information between image X and image Y, H(X) and H(Y) are the entropies of images X and Y respectively, and H(X,Y) is the joint entropy of the two images. The correspondence pattern that maximizes the mutual information indicates the correct correspondence pattern. The mutual information in this case is taken as a measure for the efficacy of the correspondence between images. When the two images are correctly matched, the mutual information will be maximized.

[0023] Once the features have been corresponded, the known relative positions of the digital video cameras can be used to calculate the three dimensional location of the feature points detected in more than one digital video camera. There are multiple approaches to performing this three dimensional location determination. One such approach as explained below is commonly known as the DLT technique and is detailed in Direct Linear Transformation from Comparator Coordinates into Object Space Coordinates in Close-Range Photogrammetry (1971) by Y.I. Abdel-Aziz, Y.I. and H.M. Karara. Proceedings of the Symposium on Close-Range Photogrammetry (pp. 1-18). Falls Church, VA: American Society of Photogrammetry.

[0024] Using the DLT technique, for each of the feature points detected in more than one digital video camera, a set of simultaneous equations is set up. These equations are:

$$\begin{bmatrix} \frac{v^{(1)}L_9^{(1)} - L_1^{(1)}}{R^{(1)}} & \frac{v^{(1)}L_{10}^{(1)} - L_2^{(1)}}{R^{(1)}} & \frac{v^{(1)}L_{11}^{(1)} - L_3^{(1)}}{R^{(1)}} \\ \frac{w^{(1)}L_9^{(1)} - L_5^{(1)}}{R^{(1)}} & \frac{w^{(1)}L_{10}^{(1)} - L_6^{(1)}}{R^{(1)}} & \frac{w^{(1)}L_{11}^{(1)} - L_7^{(1)}}{R^{(1)}} \\ \vdots & \vdots & \vdots \\ \frac{v^{(m)}L_9^{(m)} - L_1^{(m)}}{R^{(m)}} & \frac{v^{(m)}L_{10}^{(m)} - L_2^{(m)}}{R^{(m)}} & \frac{v^{(m)}L_{11}^{(m)} - L_3^{(m)}}{R^{(m)}} \\ \frac{w^{(m)}L_9^{(m)} - L_5^{(m)}}{R^{(m)}} & \frac{w^{(m)}L_{10}^{(m)} - L_6^{(m)}}{R^{(m)}} & \frac{w^{(m)}L_{11}^{(m)} - L_7^{(m)}}{R^{(m)}} \end{bmatrix} \begin{bmatrix} X \\ Y \\ Z \end{bmatrix} = \begin{bmatrix} \frac{L_4^{(1)} - v^{(1)}}{R^{(1)}} \\ \frac{L_8^{(1)} - w^{(1)}}{R^{(1)}} \\ \vdots \\ \frac{L_4^{(m)} - v^{(m)}}{R^{(m)}} \\ \frac{L_8^{(m)} - w^{(m)}}{R^{(m)}} \end{bmatrix} \quad (2)$$

[0025] Where X,Y,Z are the 3D coordinates of the point, L₁ – L₁₁ are the direct linear transform parameters determined from digital video camera calibration, m is the number of digital video cameras, u⁽ⁱ⁾ and w⁽ⁱ⁾ are the pixel coordinates of the feature in digital video camera i, and R⁽ⁱ⁾ is given by

$$R^{(i)} = L_9^{(i)}X + L_{10}^{(i)}Y + L_{11}^{(i)}Z + 1 \quad (3)$$

[0026] A simultaneous equation solving technique (such as the linear least squares technique) is used to determine the unknowns in these equations. The solution to these equations is the three dimensional location of the feature point. This procedure is repeated for

all of the points detected by more than one digital video camera. This produces a cloud of three dimensional points 610 representing a model of the surface of the three dimensional object as illustrated in FIG. 6.

[0027] Any of a large number of data processing algorithms can be applied to this cloud of points 610 to calculate a surface model more amenable to the specific task. One approach would be to fit a spline surface through this cloud of points 610. Another approach is to use a cluster of super-quadrics. Regardless of which of these point cloud to surface algorithms are used, the essential task of constructing a model of one aspect of the three dimensional surface is complete.

[0028] The embodiments described herein have been presented for purposes of illustration and are not intended to be exhaustive or limiting. Many variations and modifications are possible in light of the forgoing teaching. The system is limited only by the following claims.

CLAIMS

What is claimed:

1. A device for generating a three-dimensional surface model of an object, comprising:
a mounting panel;
at least two video cameras;
a projection system;
a communication channel for controlling the device;
another communications channel for downloading data from the device; and
a power source.
2. The device of claim 1, further comprising at least one orientation motor;
3. The device of claim 1, wherein the video cameras are grey scale digital video cameras.
4. The device of claim 2, further comprising a color digital camera.
5. The device of claim 2, wherein the grey scale digital video cameras are time synchronized.
6. The device of claim 1, wherein the object is a moving object.
7. The device of claim 1, further comprising another object.
8. The device of claim 1, wherein orientation motor controls the pitch of the device.
9. The device of claim 1, wherein orientation motor controls the roll of the device.
10. The device of claim 1, wherein orientation motor controls the yaw of the device.
11. The device of claim 1, wherein the focus of the projection system is controlled electronically.
12. The device of claim 1, wherein the zoom of the projection system is controlled electronically.
13. The device of claim 1, wherein the aperture of the projection system is controlled electronically.

14. The device of claim 1, wherein the focus of the video cameras are controlled electronically.
15. The device of claim 1, wherein the zoom of the video cameras are controlled electronically.
16. The device of claim 1, wherein the aperture of the video cameras are controlled electronically.
17. The device of claim 5, wherein the moving object is a mammal.
18. An imaging device comprising:
 - a means for generating a three-dimensional surface mesh of a subject; a means for casting an image on the subject;
 - a means for controlling the imaging device; and
 - means for obtaining data from the imaging device
19. The imaging device of claim 18, further comprising a means for generating three-dimensional surface mesh with color information.
20. The imaging device of claim 18, further comprising a means for adjusting the roll, pitch and yaw of the imaging device.
21. A method for generating a three dimensional surface model of an object comprising:
 - placing an object within a viewing range of an imaging device;
 - projecting a pattern onto the object
 - obtaining an image of the object with the pattern superimposed on the object;
 - transmitting the image to a computer;
 - detecting pattern features;
 - corresponding the pattern features;
 - calculating coordinates of points on the pattern features;
 - generating a three dimensional model from the coordinates.

22. The method of claim 21, wherein at least two images of the object with the pattern superimposed on the object are obtained from different perspectives.
23. The method of claim 22, wherein the steps of detecting pattern features, corresponding the pattern features and calculating the coordinates of the pattern features are repeated for the at least two images.
24. The method of claim 22, further comprising sending control signals to the imaging device.

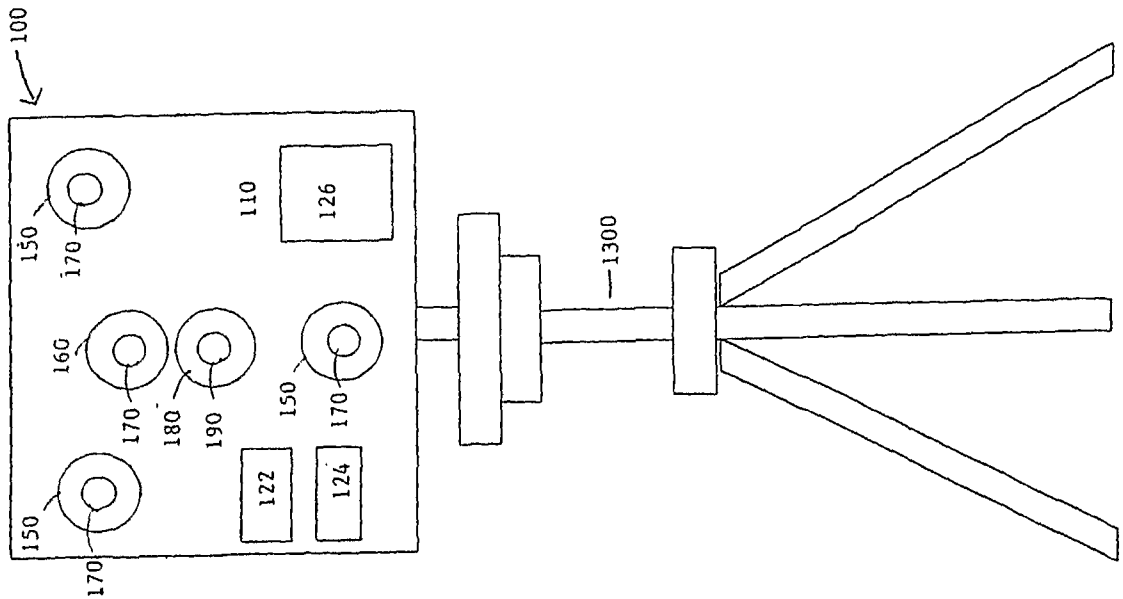


Figure 1.

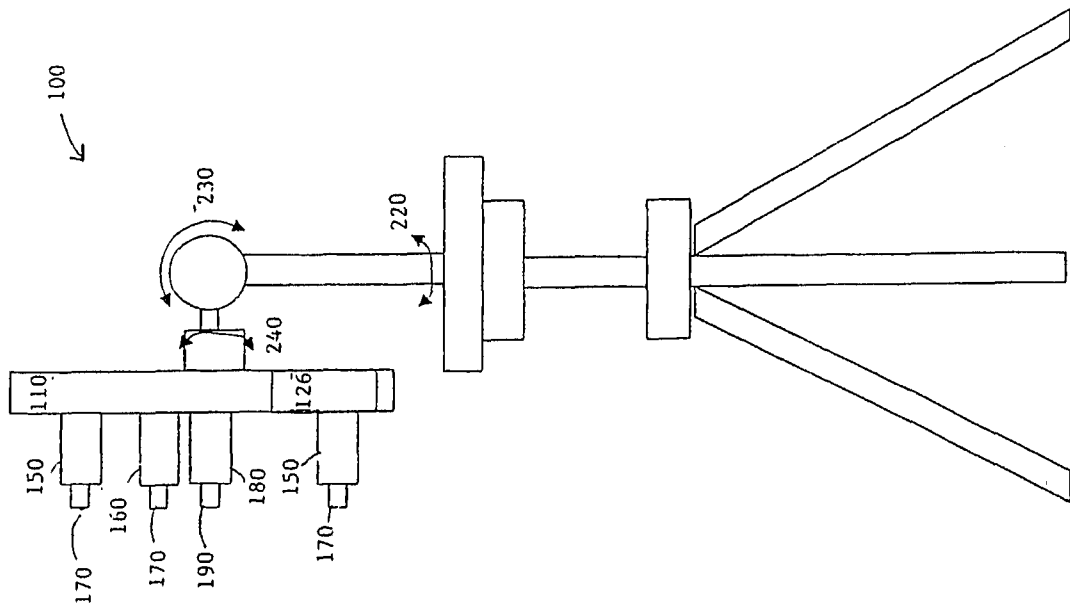
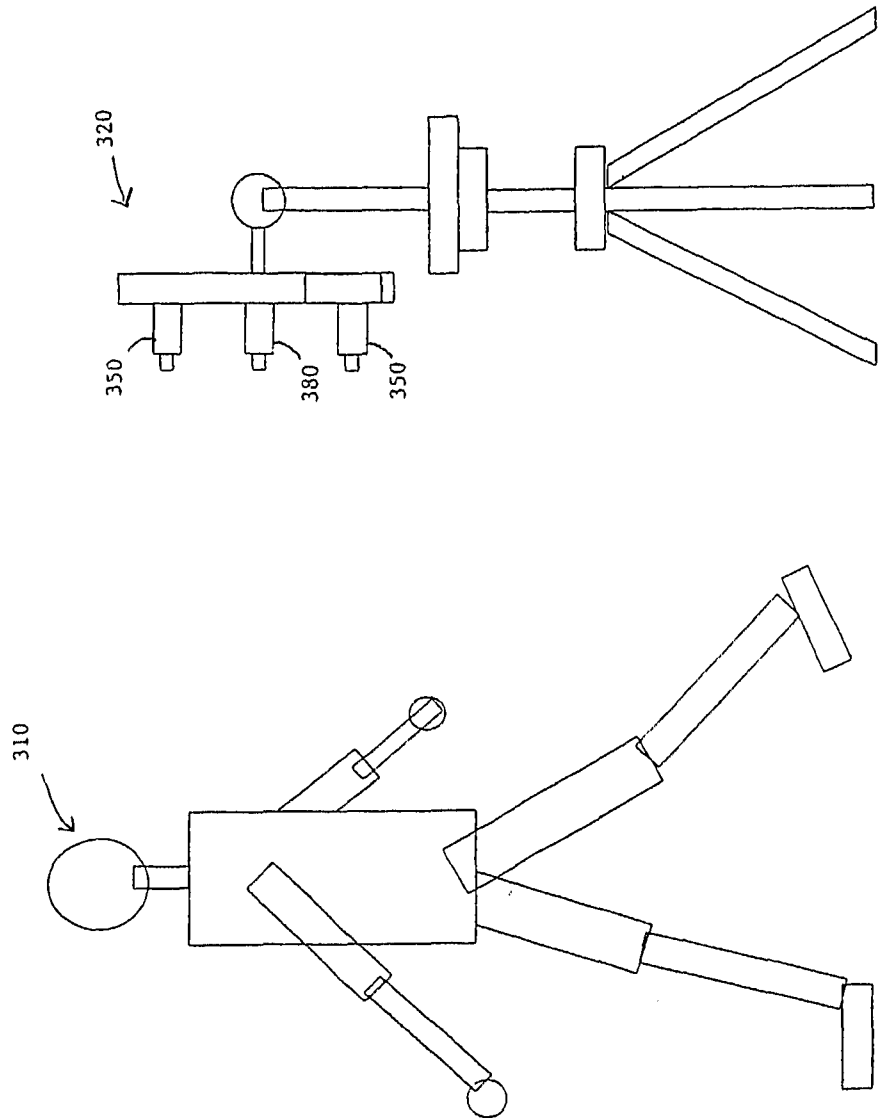


Figure 2.

Figure 3



4/6

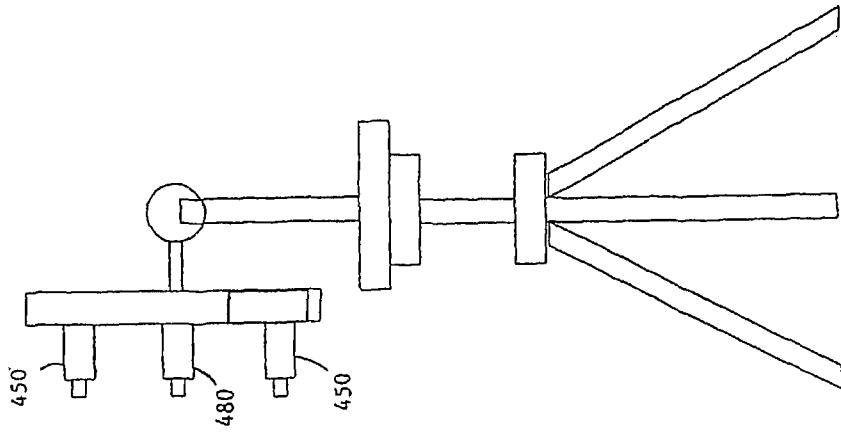
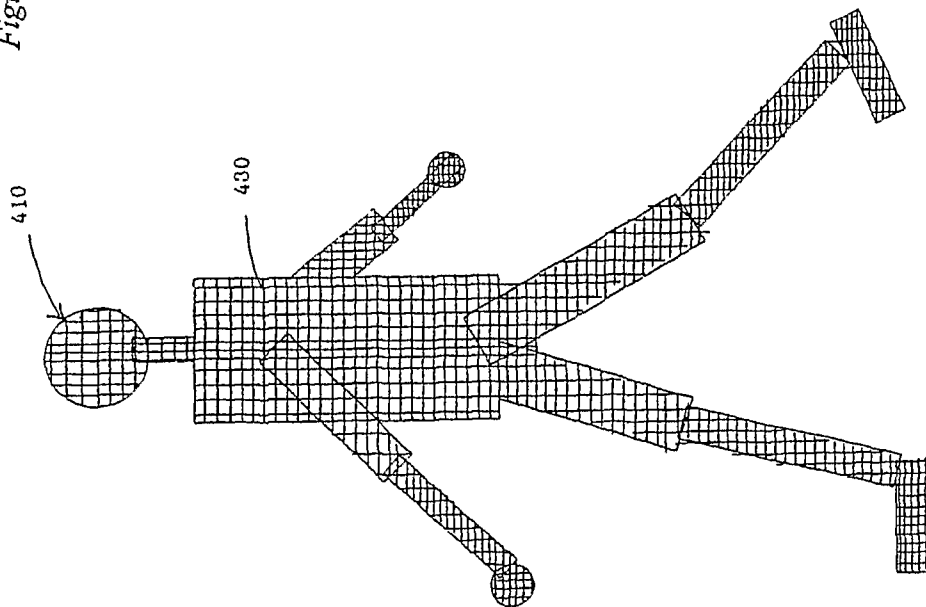


Figure 4.



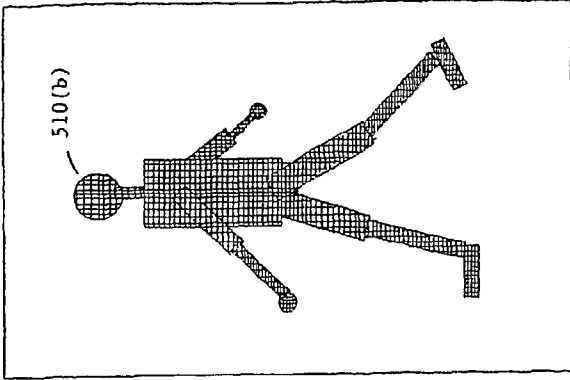


Figure 5.

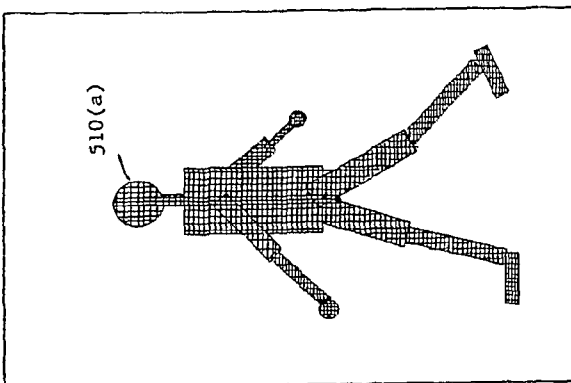
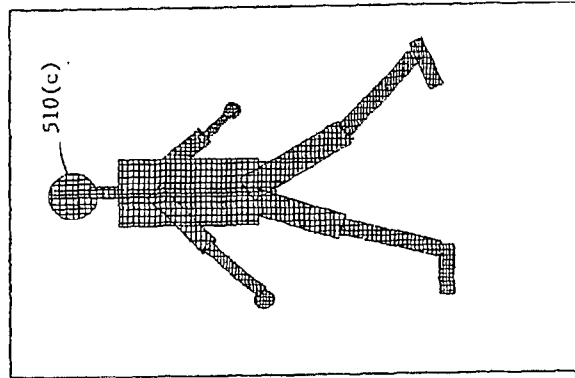


Figure 6.

