Office de la Propriété Intellectuelle du Canada

Un organisme d'Industrie Canada

Canadian Intellectual Property Office

An agency of Industry Canada

CA 2306515 A1 2001/10/25

(21) 2 306 515

(12) DEMANDE DE BREVET CANADIEN
CANADIAN PATENT APPLICATION

(13) A1

(22) Date de dépôt/Filing Date: 2000/04/25

(41) Mise à la disp. pub./Open to Public Insp.: 2001/10/25

(51) Cl.Int.$^{7}$/Int.Cl.$^{7}$ H04N 13/02, H04L 12/16

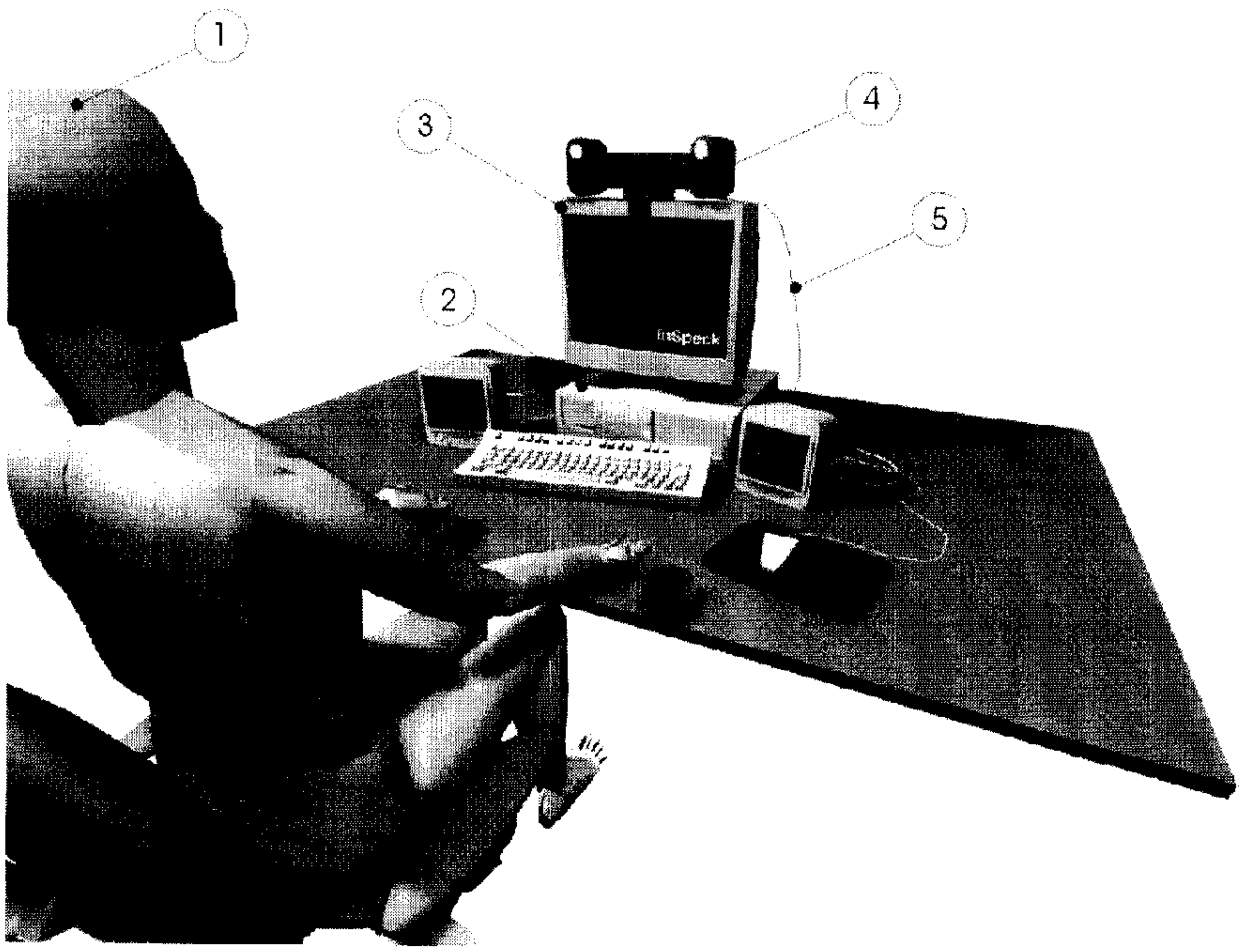
(71) Demandeur/Applicant: INSPECK INC., CA

(72) Inventeur/Inventor: UNKNOWN, ZZ

(74) Agent: ROBIC

(54) Titre : VISION STEREO INTERNET, NUMERISATION 3D ET CAMERA DE SAISIE DE MOUVEMENT

(54) Title: INTERNET STEREO VISION, 3D DIGITIZING, AND MOTION CAPTURE CAMERA



# INTERNET STEREO VISION, 3D DIGITIZING, AND MOTION CAPTURE CAMERA

## FIELD OF THE INVENTION

The present invention relates to internet stereo vision, 3D digitizing, and motion capture camera that may form, for example, a color non-contact optical 3D digitizer or a motion capture system for stereo vision, 3D display, computer assisted 3D vision, internet communication, off-line and on-line electronic games.

## BACKGROUND

3D digitizing, particularly non-contact optical 3D digitizing techniques, have become commercially available during recent years. Most of these techniques are based on the principle of optical triangulation. Despite the fact that passive optical triangulation (stereo vision) has been studied and used for many years for photogrametic measurements, the active optical triangulation technique (particularly laser scanning technique) has gained popularity because of its robustness and simplicity to process the obtained data using a computer. Most of the systems based on the active optical triangulation principle were developed for industrial applications, such as robotics assembly, robot guidance, industrial inspection, reverse engineering, etc. A laser beam or a laser stripe is projected onto a 3D surface of an object, scattering the laser beam or laser stripe on the surface. It is measured using a photo-electronic device. A signal can be measured indicating the position (usually the depth) of the measuring point. In most cases, the basic measurements are either a point or a section profile. A mechanical or optical scanning device is usually used to provide a frame of 3D measurement. Laser is a monochromatic light source which does not provided full color information.

So, an additional camera and light source are used when a color texture is needed.

A new category of optical color 3D digitizers, such as InSpeck's product line, have been developed. These systems use structured white light projection combined with a CCD camera allowing for the measurement of 3D geometry and color texture of a surface. The projected structured light (viewed by a camera from an angle different from the light projection) is deformed due to the 3D surface relief . The 3D coordinates of the surface are calculated by analyzing the deformation. These kind of systems are being used in computer animation, special effects and in electronic game development.

On the other hand, the passive optical triangulation (stereo vision, for example) is largely used for the purpose of motion capture. The correspondence problem (automatically finding one point on the object's surface from two optical sensors, cameras in general) is not a major obstacle for this application because only a limited number of points must be measured. These points are often characterized by using visible markers.

Another application of stereo vision is stereoscopic 3D display. Instead of determining the 3D coordinates of some points of an object in a 3D space, it simply needs to display a pair of stereoscopic images on a monitor (TV or computer monitor) so that the 3D perspective of an image can be seen. One possible configuration is to capture a pair of images using two cameras which observe the parallax effect of an object. Then one of this pair of stereoscopic images will be viewed by the left eye and another will be viewed by the other eye. The human brain can easily merge this pair of images so that the object is viewed as a 3D image.

The existing 3D digitizing systems and optical motion capture systems are, in general, complex and too expensive for the Internet and mass consumer applications. Most of these systems incorporate sophisticated optical, electro-optical, mechanical and electronic components. Special expertise is needed to operate such digitizers. In addition, the existing systems support separately the 3D digitizing and motion capture functions. It would be preferable in fact to incorporate two functions into one system. The 3D model could be first created with some identifiable control points located on the model surface. Then the 3D position of these control points could be captured in real or quasi-real time, so that the whole model could be controlled or animated.

SUMMARY

An object of the invention is to provide an apparatus having a combination of capturing stereoscopic images, color 3D digitizing, and motion capture functions, mainly but not restrictively for internet related application. The apparatus will be connected to a computer and used by the mass consumer for applications including internet, conference via internet, 3D Web, e-commerce, off-line and on-line games and any application which requires affordable 3D digitizing and/or motion capture solution.

The invention includes both a hardware and software solution for the integration of a pair of stereoscopic images at a video rate, color 3D digitizing and 3D motion capture features, into one single apparatus. In order to create an affordable device and solution for the mass consumer, there is developed a way to incorporate elements to capture and transfer a pair of stereo images, to obtain 3D coordinates and the color texture of a surface, and to capture the displacement of a number of given points in a real or quasi-real time. The data capturing process is also simplified to make the operation of the apparatus as automatic as possible.

Multiple issues are examined covering the creation of a complete color texture model and the merge of 3D model and motion control points.

The Internet stereo vision, 3D digitizing and motion capture camera according to the invention comprises at least two electronic image capturing devices (cameras), and at least one projection system combining a miniaturized light projector and an encoded projected pattern. The miniaturized light projector provides the necessary lighting for an active 3D range sensing for each of the cameras. The encoded pattern will be projected on the surface of an object. The two cameras (or more) are mounted on a base that keeps the relative position of the cameras fixed in such a way that the scene captured by one camera shares a common area with the scene captured by another camera. The relative position and angle of the cameras is chosen in such a way that each of the optical axis of the cameras converge through a single point, called the converging point or origin of the 3D space. The optical axis of one light projector intersects with the optical axes of one or more cameras at one fixed point that we define as the origin of the 3D space. All cameras can observe disparity created by the light projector. There is also disparity in a pair of images captured by two cameras. This apparatus is connected to a computer via digital port like a USB port, or other standard high speed connections. The cameras and respective miniaturized light projectors are controlled by a computer and software. It is also possible to launch a 3D measurement process using a snapshot button. In addition, the apparatus can be mounted on a rotational table. The rotation of this rotational table is directly controlled by the computer. It is also possible to place an object on the rotational table so that the angular position of the rotated object can be known.

The apparatus provides at least three functions:

1. A pair of cameras can capture a pair of stereoscopic images at video rate. The stereo 3D image can be created when these two images are displayed on a monitor which sends one image to the left eye and another image to the right eye. These images are transferred via a high speed link (Internet for example) to another computer.

2. Combining the light projectors and both cameras, this apparatus provides measurement of the 3D coordinates of a surface. The encoded pattern is projected on the surface of an object by a light projector and the respective camera captures the scene. With the surface relief of the object, the projected pattern is deformed from the point of view of the camera. With a careful calibration technique, it is possible to determine the 3D coordinates of some points on this surface by measuring the deformation of the projected pattern. In principle, a combination of one camera and one light projector can carry out the measurement of the 3D coordinates. The use of two or more cameras which cover a common space, combined with one light projector, provides three major advantages. First, the weighted average values of the 3D measurements obtained by each of the cameras correspond to a better 3D measurement. Second, this configuration overcomes more problems caused by a shadow effect. Third, the two cameras observe the projected pattern from different views so that a better interpretation of the deformation of the projected pattern on a discontinued surface can be obtained.

3. The third function is to make a motion capture of a limited number of points in 3D space. When using some markers on a surface, it is relatively easy to determine the 3D positions of these points. It is also possible to use some known points on the object, like the features of the skin, lips, eyelids, eyes, etc. Of course, the 3D space observed by the cameras must be calibrated and a disparity of

a given point captured by the cameras can be evaluated and its 3D position can be calculated. When the number of points to be measured is low, it is even possible to determine the 3D positions of these points several times per second. This data can be used to control the motion of an object or model.

The invention also includes a motion tracking method which analyzes the dynamic motion of a subject in a scene captured by one or more cameras. A servo control device will control (in real time) the rotation of the rotational table in a such way that the apparatus can follow the dynamic motion of the subject. A user, either present or at a distance, can also send a command to the computer in order to orient the apparatus to a desired direction. Since the control device provides the exact position of the rotational table, it is evident that the whole 3D space covered by the apparatus mounted on the rotational table is calibrated as a known geometric space with respect to the apparatus. This function provides the possibility to cover a larger space to perform the three basic functions of this apparatus.

Instead of mounting the apparatus on the rotational table, sometimes, it is convenient to place an object on this table. This configuration simplifies the operation to merge several views of a 3D object to create a complete 3D model. A 3D digitizer can measure the 3D surface of one single view of an object. In order to create a complete object model, it is necessary to capture different views of an object. When the object is placed on the controlled rotational table, the precise position of each view with respect to the other views is known. So it is easy to register several views in a common 3D coordinate and to merge them to create a complete 3D model.

## BRIEF DESCRIPTION OF THE DRAWINGS

A detailed description of preferred embodiments will be given herein below with reference to the following drawings, in which like numbers refer to like elements:

Figure 1 shows one of the possible working environments for the apparatus.

Figure 2 shows a simplified inside view of the apparatus according to the invention.

Figure 3a and 3b show respectively two embodiments of the apparatus according to the invention.

Figure 4 shows the use of a rotational table for digitizing different views of an object, according to the invention.

Figure 5 show a webcam 3D incorporated in the monitor, according to the invention.

## DETAILED DESCRIPTION OF THE PREFERRED EMBODIMENTS

Referring to Figure 1, the apparatus according to the invention can be mounted either on top of a monitor (3), or beside the computer, or anywhere accessible by the cables. The apparatus is directed to a subject (1), a human for example, so that the functions such as stereo vision, 3D digitizing, or motion capture can be carried out on the subject.

Referring to Figure 2, the apparatus consists of a light projector (8) coupled with a projection lens (9), a encoded projected pattern (10), the first camera (11) with its lens (12), and the second camera (13) with its lens (14). The light projector (8), camera (11) and camera (13) are linked respectively to the connector (15) so that the signals are communicated to computer (2) by the cable (5).

The encoded projected pattern (10) includes one or a series of 2D graphic patterns. These patterns may be binary or gray scaled lines or blocks. The orientation of the lines or blocks can be vertical, horizontal or any given angle. The shape and size of the lines or blocks can be constant everywhere on the pattern. They can also be variable. The optical axis of the projection lens should be perpendicular to the surface of the projected pattern. The aperture of this lens and its focus should be well adjusted so that the projected pattern will be sharp enough over all of depth of the measurement.

The camera (11) and the camera (13) are placed symmetrically on both sides of the light projector (8), as shown in Figure 2. The optical axes of lens (12) and (14) intersect with the optical axis of the projection lens at identical distances from the apparatus. These three optical axes intersect at the same point, which is usually referred as the optical center of the apparatus. The angle between the optical axis of the camera and the optical axis of the projection lens determines the sensitivity of the distance measurement, referred to as depth, in the direction of the optical axis of the projection lens. The bigger the angle is, the more sensitive the depth measurement will be. On the other hand, the angle should not be too big, in order to keep an overlapped area over the depth of measurement covered by the two cameras at least equal to 80-90% of the field of view of each camera.

To perform stereo vision display, the image captured by the two cameras are first transferred to a host computer, then the images are compressed before being sent to another computer. The images received by the second computer will be decompressed and displayed on a screen. In order to see a 3D stereo image, one of the two images need to be seen by the left eye and the other image seen by the right eye. This can be accomplished by synchronizing the display of the two images using an LCD eye glass.

Once the apparatus is assembled, the optical and mechanical configuration of the system are fixed. A conversion table is generated using parameters of the fixed configurations. Each time a 3D digitizing is performed, the deformation of the projected pattern captured by the two cameras are converted to geometrical measurements of the 3D surface. The color texture can be generated either by capturing a separate image or by removing the deformed pattern from the images. When the rotational table (18) is used as shown in Figure 4 for the purpose of capturing multiple views, the computer controls precisely the position of the rotation so that the relative position between each view is known. It is very straight forward to put multiple partial models together to create a complete model.

For motion capture function, the disparity observed by two cameras (11) and (13) is calibrated in the common 3D space of these cameras. The 3D position of a corresponding point captured by two cameras can then be determined. A method for real time disparity measurement is developed to ensure a motion capture of limited points at quasi video rate.

An algorithm to track an object in a dynamic motion is needed to control the apparatus 4 mounted on the rotational table (18) to follow the object in motion. The difference between subsequent video images is used to determine the motion of the object and a servo control algorithm is used to control the rotation of the rotational table. The quantity of the rotation is used to bring each new coordinate to a common starting coordinate.

Referring to Figures 3a and 3b, the apparatus (4) can be mounted on a rotational table (18), which is linked to a computer by a cable (19).

Referring to Figure 4, an object to be digitized (20) can be placed on the rotational table (18) in a such way that the different views of this object can be digitized with the known angular positions.

Referring to Figure 5, the apparatus according to the invention may be embodied as webcam 3D incorporated in the monitor.

While embodiments of this invention have been illustrated in the accompanying drawings and described above, it will be evident to those skilled in the art that changes and modifications may be made therein without departing from the essence of this invention.

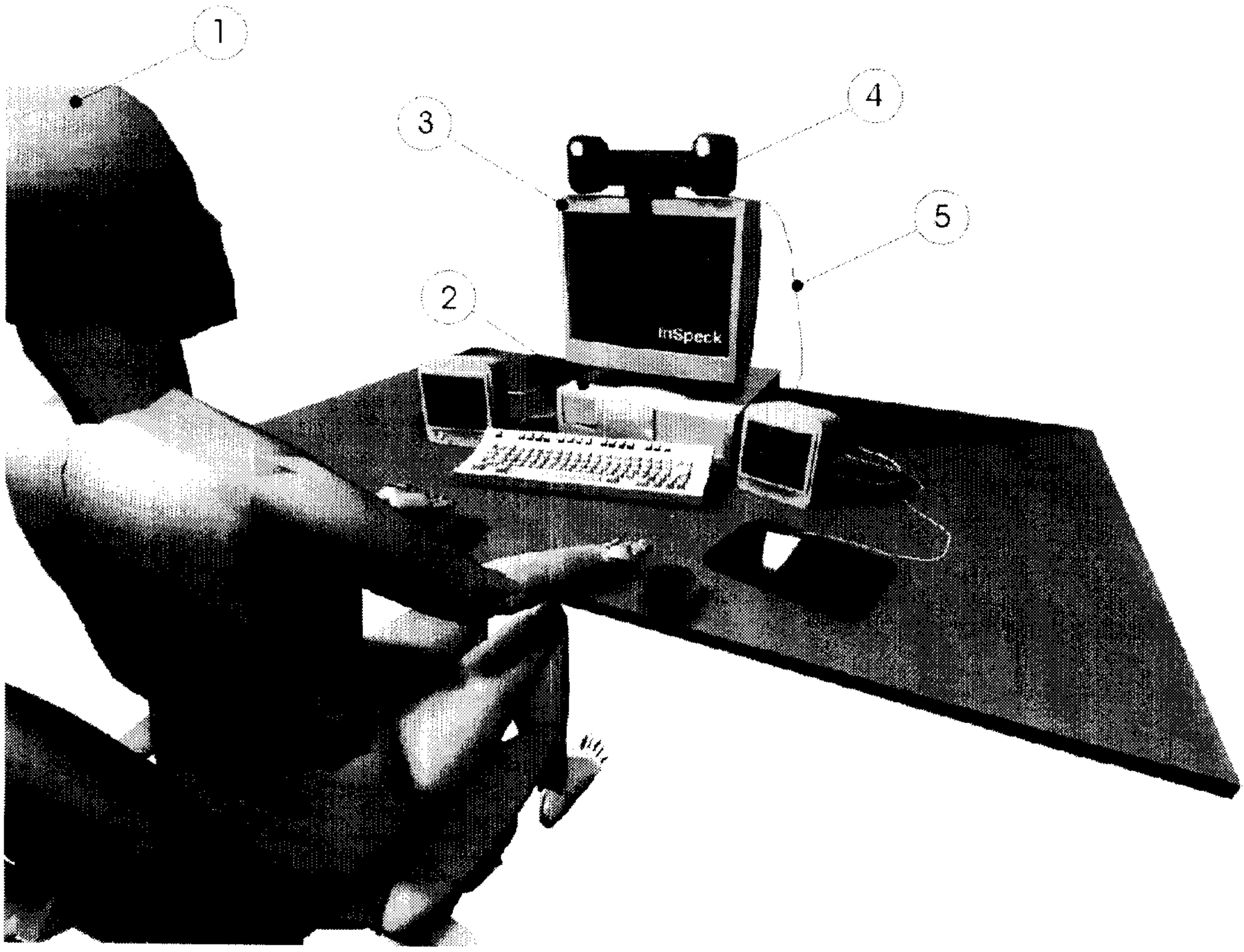



**Figure 1**

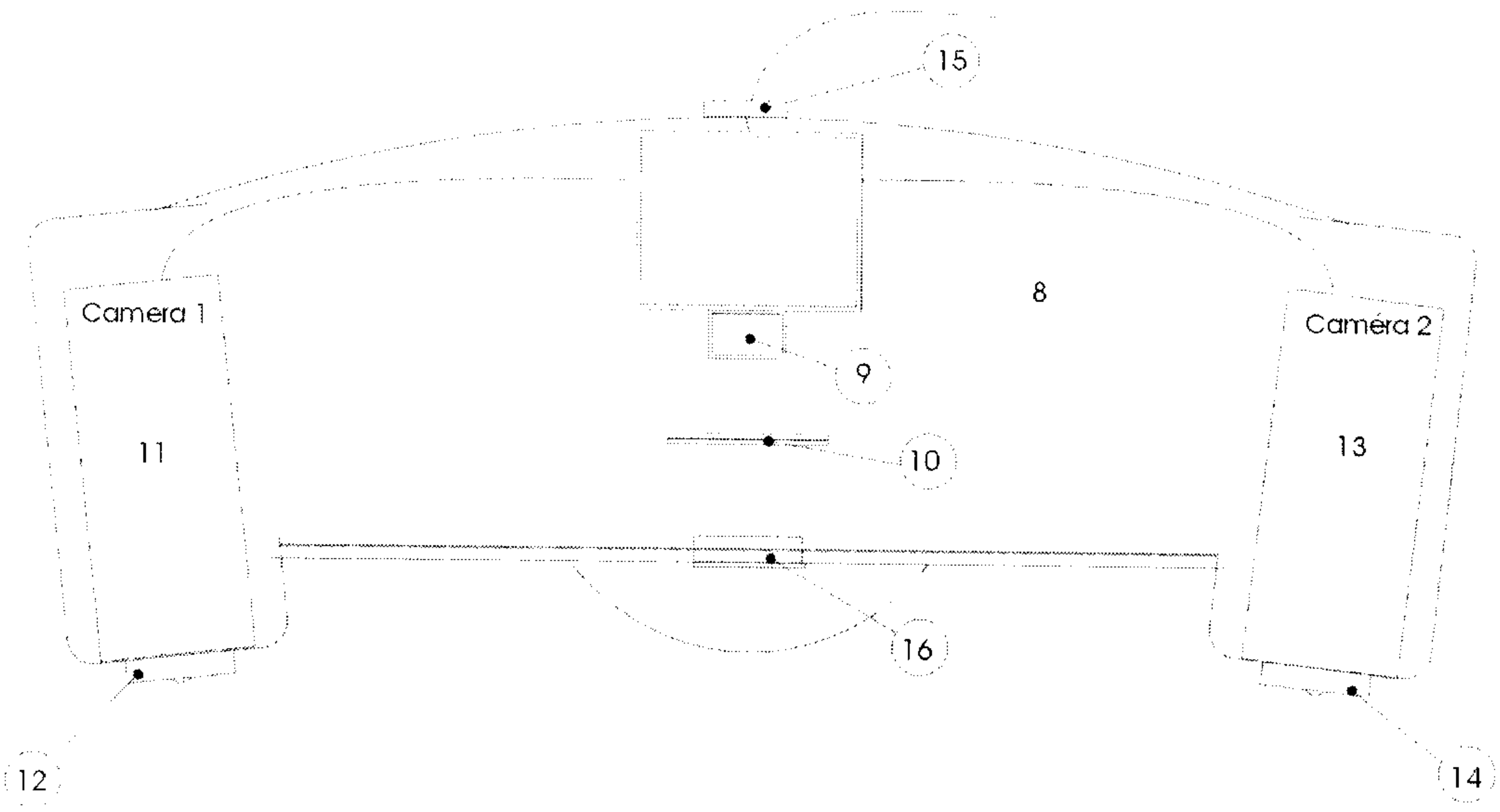



**Figure 2**

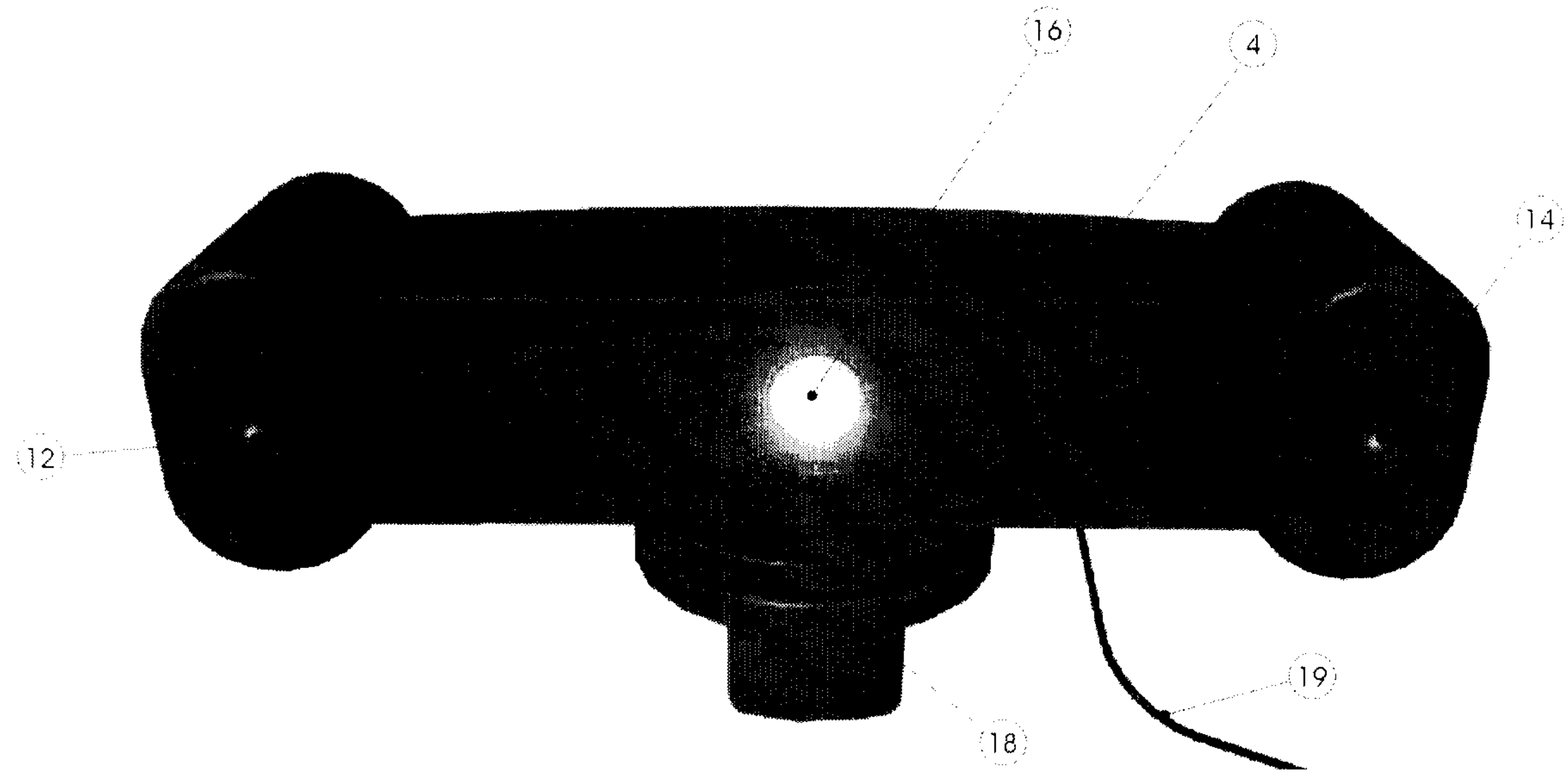



**Figure 3a**

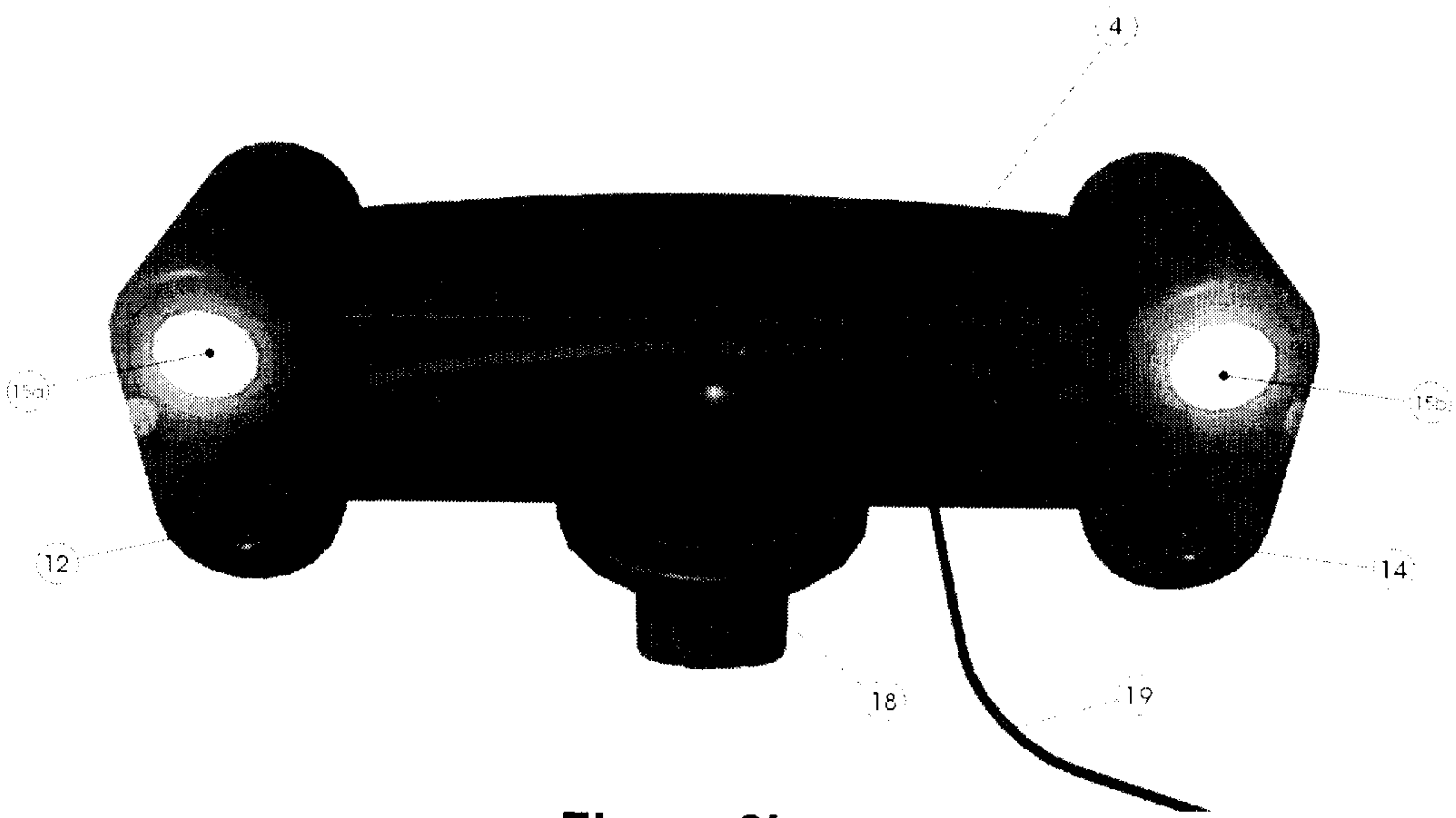



**Figure 3b**

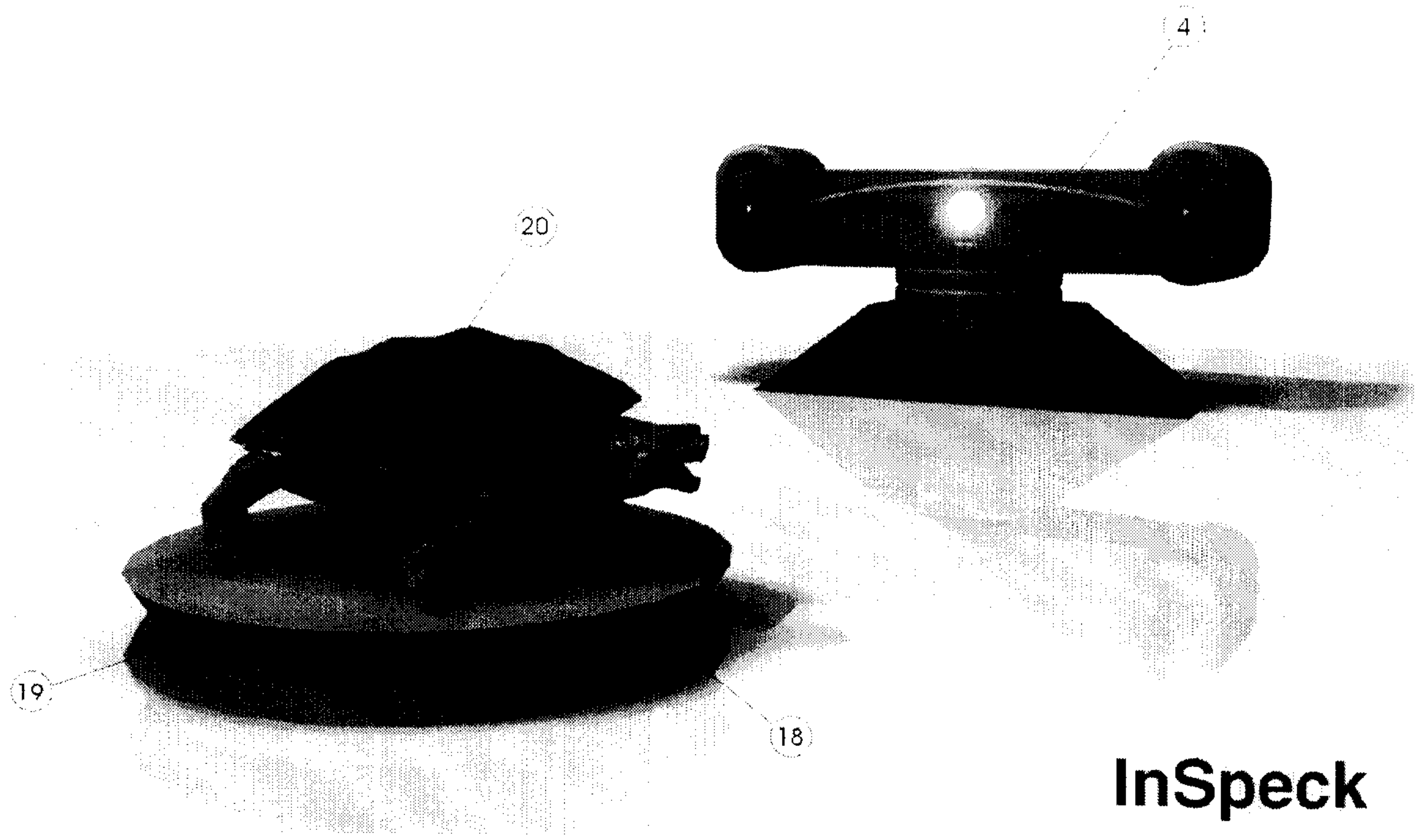



**Figure 4**

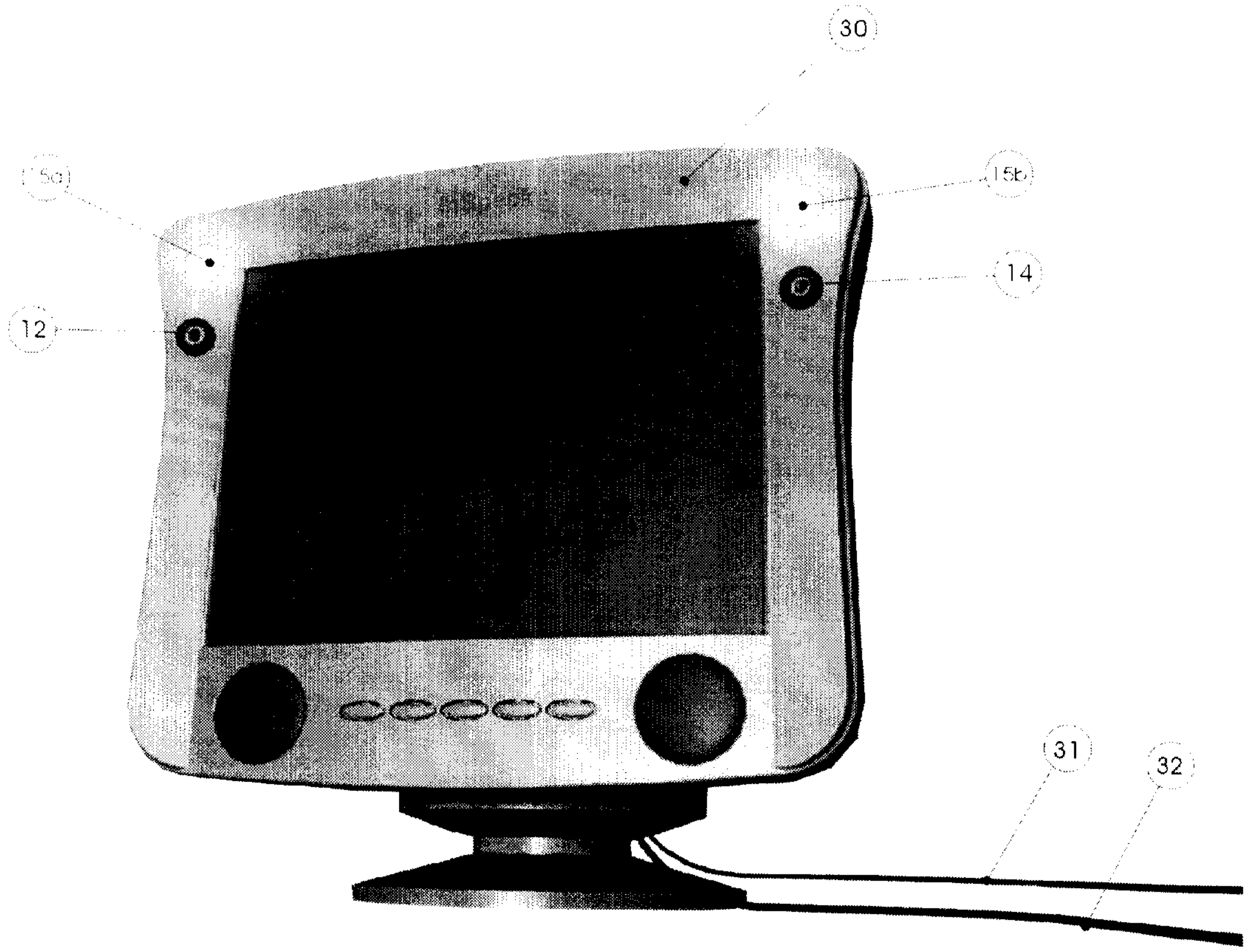
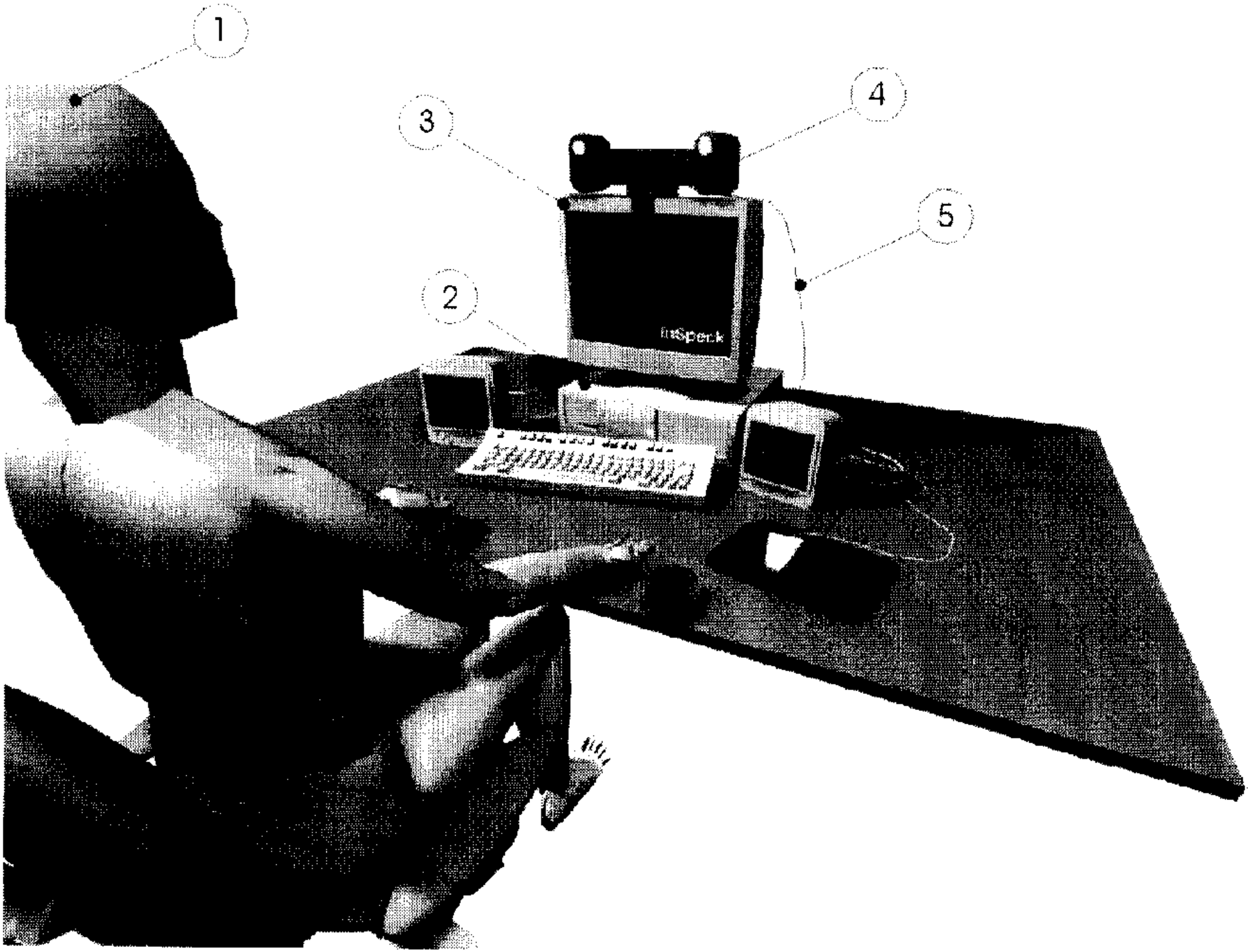



**Figure 5**


1
2
3
4
5
inSpeck