(12) INTERNATIONAL APPLICATION PUBLISHED UNDER THE PATENT COOPERATION TREATY (PCT)

(19) World Intellectual Property Organization
International Bureau

(43) International Publication Date
31 May 2007 (31.05.2007)

PCT

(10) International Publication Number
WO 2007/061678 A2

(51) International Patent Classification:
*H04N 1/48* (2006.01)

(21) International Application Number:
PCT/US2006/044096

(22) International Filing Date:
14 November 2006 (14.11.2006)

(25) Filing Language: English

(26) Publication Language: English

(30) Priority Data:
60/738,247 18 November 2005 (18.11.2005) US

(71) Applicants and
(72) Inventors: **SANDBERG, ROY** [US/US]; 1515 Alabama St. #c, Huntington Beach, CA 92648 (US). **SANDBERG, DAN** [US/US]; 163 Fairmount St., San Francisco, CA 94131 (US).

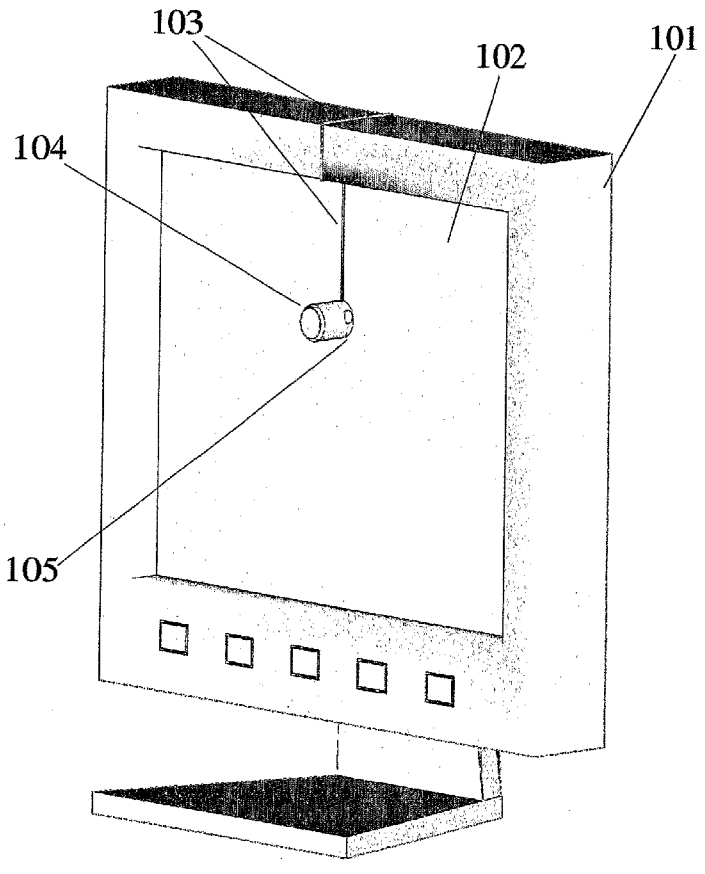
(81) Designated States *(unless otherwise indicated, for every kind of national protection available)*: AE, AG, AL, AM, AT, AU, AZ, BA, BB, BG, BR, BW, BY, BZ, CA, CH, CN, CO, CR, CU, CZ, DE, DK, DM, DZ, EC, EE, EG, ES, FI, GB, GD, GE, GH, GM, GT, HN, HR, HU, ID, IL, IN, IS, JP, KE, KG, KM, KN, KP, KR, 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.*

(54) **Title:** SCREEN MOUNTED VIDEO TELECONFERENCING CAMERA APPARATUS




(57) **Abstract:** An apparatus for reducing gaze-perception discrepancies while video teleconferencing, consisting of a camera mounted directly on a display screen, and preferentially along the same axis as the line of sight between the local user and the remote user's eyes as displayed on the display screen.

WO 2007/061678 A2

# Screen Mounted Video Teleconferencing Camera Apparatus

by Roy Benjamin Sandberg, residing at 1515 Alabama St., Apt. #C, Huntington Beach, CA

and

Dan Ron Sandberg, residing at 163 Fairmount St., San Francisco, CA

## BACKGROUND OF THE INVENTION

(1) Field of Invention

The present invention is related to the field of video teleconferencing, more specifically, the invention is an apparatus for reducing gaze-perception discrepancies while video teleconferencing.

(2) Related Art

Gaze preservation concerns the problem of appearing to maintain eye contact with a remote user when video teleconferencing. This is advantageous for maintaining the psychological illusion of face-to-face conversation. Standard video teleconferencing systems involve a screen and a camera, where the camera is placed some distance away from the screen. Because a user typically looks at the screen and not the camera, a remote viewer does not perceive the user to be looking into the remote viewer's eyes. This creates a sense of psychological distance between the user and the viewer, and to some degree prevents the video teleconferencing experience from being as emotionally rich as it might be in person.

Other means of gaze preservation across a video teleconferencing link are known, and include: U.S. Patent number 6,889,120 (Jouppi). Complex systems involving software transformations of the camera image or partial mirrors placed in front of the screen also exist. However, these methods require complex and expensive methods and apparatuses and/or do not perfectly preserve proper gaze.

## SUMMARY OF THE INVENTION

The present invention is a new and improved method and apparatus to preserve gaze across a video teleconferencing link. The invention is a low-cost solution to gaze preservation consisting of mounting a camera directly on the display screen, and preferentially along the same axis as the line of sight between the local user and the remote user's eyes as displayed on the display screen. By keeping the frontal area of the camera small, the camera only minimally disrupts the local user's view of the remote user. Additionally, the wire leading away from the camera is designed to be as a narrow as reasonably possible to avoid blocking screen real estate. A suction cup or other means of attaching the camera in front of the screen should be provided. In alternative embodiments, the camera is held in front of the screen using a rigid support, the rigid support in turn connecting to an off-screen location. An optional quick release feature for the suction cup allows a user to press a button on the camera and easily detach the camera from the screen by re-pressurizing the area between the suction cup and the screen.

The device is operated by placing it as near as possible to the image of eyes of the person being video teleconferenced with on the display screen. In the preferred embodiment, the camera is placed on the surface of the display screen between the eyes of the person displayed on the screen, that person being the person who is being video teleconferenced with. By way of example, in an typical embodiment using a 17 inch monitor with full-screen video conferencing, the camera is located at the horizontal center point of the display surface and about 4 inches below the top edge of the display surface. The wires that supply the camera power and data are routed from the camera to the edge of the screen. The device is then used in the same manner as existing video cameras or web cams.

A "web cam" style camera is built out of a lens, lens holder, and CCD or CMOS sensor. In addition to the sensor, other electronic components are often required to make a functioning CCD or CMOS camera. For example, integrated circuit packages may be used that convert the output of the CCD or CMOS camera

into a format suitable for streaming over a Universal Serial Bus ("USB") or other serial interface. In the preferred embodiment of the invention, these additional electronic packages are situated behind the CCD or CMOS chip, and thus do not add to frontal area of the camera, thereby enabling an overall form factor with a smaller footprint on the LCD screen. Techniques know in the art of electronic printed circuit board manufacturing such as double sided printed circuit boards, headers, and daughter boards may be used to implement this embodiment.

In another embodiment of the invention, the camera and LCD form an integrated system. In the preferred embodiment, the camera is attached to a slide or pivot mechanism that allows it to move from a non-overlapping position to a position that overlaps and obscures a portion of the LCD screen. This is useful for applications where the user does not always want a camera overlapping the screen. By sliding or pivoting the camera on or off the screen, the user can configure the system to best suit his immediate needs. In the preferred embodiment, the slide or pivot assembly is composed of a transparent material such as polycarbonate or acrylic, and thin or transparent wires are routing along the slide or pivot assembly to supply power and read data from the camera. This minimizes the degree to which the screen is obstructed by the camera and the slide or pivot assembly.

In another embodiment of the invention, the camera is attached to a slide or pivot mechanism, and the slide or pivot mechanism connects to a hanger section. The hanger section allows the invention to hang over the top of a monitor or LCD panel, and the slide or pivot mechanism allows the camera to be moved from a non-overlapping position to a position that overlaps and obscures a portion of the monitor or LCD screen. This allows the invention to be used with a standard monitor or LCD panel while being easily moved between on-screen and off-screen positions. In the preferred embodiment, the slide or pivot assembly is composed of a transparent material such as polycarbonate or acrylic, and thin or transparent wires are routing along the slide or pivot assembly to supply power and read data from the camera. This reduces the degree to which the screen is obstructed by the camera and the slide or pivot assembly.

The invention provides a low cost solution to the problem of gaze preservation,

and is usable with all forms of video teleconferencing utilizing a glass or plastic display screen. While some portion of the display screen is blocked, that portion (ideally the portion near or at the image of the bridge of the other user's nose) is not critical to gathering visual cues about the other party.

An alternative solution consists of mounted a small prism or mirror to the display surface, and reflecting the image of the local user as positioned in front of the display surface from the prism or mirror to a camera mounted near or along the edge of the computer screen. The mirror or prism is oriented such that the scene in front of the display surface is reflected into the camera. Furthermore, the mirror or prism is manufactured with a concave curvature that directs light from at least a 20 degree diagonal field of view in front of the display surface onto the camera's lens element. This solution eliminates the requirement for a wire running along the screen, and allows the screen real estate blocked by the prism or mirror to be reduced relative to that blocked by the previous version of the invention. The prism or mirror is mounted to the screen using a suction cup or other means of positioning the prism or mirror in from of the screen. An optional quick release feature for the suction cup allows a user to press a button on the prism or mirror assembly and easily detach the assembly from the screen by re-pressurizing the area between the suction cup and the screen.

The device is operated by placing the mirror or prism assembly as near as possible to the image of the person being video teleconferenced with on the display screen. In the preferred embodiment, the camera is placed between the eyes of the person being video teleconferenced with. A camera is then mounted coaxially with the reflected image of the local user that has been redirected by the prism or mirror. This typically requires that the camera is mounted along the edge of the screen and attached with glue, hook-and-loop cloth, tape, clamps, or other means of attaching the camera known in the art. A means of aligning the camera with the prism or mirror assembly is necessitated for proper operation of the system. This may involve adjustment screws on the camera or prism, or alignment guides that are removed after proper alignment has occurred. The device is then used in the same manner as existing web cameras.

The preferred embodiment of this invention uses digital image stabilization

techniques to maintain the video image of the remote user in a constant position on the video screen. This assists in maintaining the video image of the remote user's face in a constant position on the local video screen, thereby assisting in maintaining the video camera's position in its preferred location, typically between the eyes of the remote user as displayed on the local video screen. Techniques similar to image stabilization, such as face-tracking, and jitter-reduction algorithms may also be used. In all cases, the techniques may be implemented as software executing on a processor, or using dedicated hardware, or a combination of both techniques. The image stabilization may be done at the remote location or at the local location. In the preferred embodiment the image stabilization would be done at the remote location, as this would have the added benefit of reducing the bandwidth of the transmitted video data due a reduction in frame-to-frame movement when using temporal video compression techniques.

In lieu of placing the video camera on the active display surface, a hole may also be made through the opaque layers of an LCD panel during the LCD manufacturing process. A CMOS or CCD imaging camera may then be located behind this hole, and via this means, the LCD screen appears to be unmodified on casual inspection.

The imaging sensor and the LCD may also both be produced on the same semiconductor substrate, thereby saving the extra step of integrating the imaging sensor and LCD.

Finally, metal wires or other conductive materials are used to transfer power and data to and from the CCD. Conductors can then be routed along the LCD screen or along a support structure without being visible to the viewer.

The invention provides a low cost solution to the problem of gaze preservation, and is usable with all forms of video teleconferencing utilizing a glass or plastic display screen

## DETAILED DESCRIPTION OF THE DRAWINGS

FIG. **1** is an exemplary embodiment of the camera on the screen apparatus.

FIG. **2** illustrates the orientation of the local user and the remote user during a video conferencing session when using the invention.

FIG. **3** illustrates the slide mechanism embodiment of the invention.

DETAILED DESCRIPTION OF THE INVENTION

The present invention is a new and improved video teleconferencing method and apparatus to preserve gaze across a video teleconferencing link by directly attaching a camera, or reflective device such as a mirror or prism to the active video display surface.

FIG. 1 is an exemplary embodiment of the camera on the screen apparatus. The monitor casing 101 is shown in a portrait orientation, but in an alternative embodiment, the landscape orientation may also be used. The monitor active display surface 102 is a conventional glass screen, although other display surfaces, such as rigid plastic or flexible plastic may also be used. The monitor active display surface 102 is shown as being flat, but a curved display surface may be used in alternative embodiments. The data and power cable 103 is shown as a thin wire that has been glued onto the active display surface and routes into the monitor casing 101. In alternative embodiments, the data and power cable may be freely hanging, and need not enter the monitor case. Attachment points for the data and power cable on the monitor casing may be used in alternative embodiments. The video camera 104, is shown as a small tubular chassis with a cylindrical lens that faces away from the display surface. Other video camera configurations and sizes may also be used. In the preferred embodiment a video camera based on charge coupled device technology is used, but other video camera technologies known in the art may also be used. Larger or small video cameras may also be used, but in the preferred embodiment, a video camera with a small surface area as measured on the plane formed by the active display will be used. The video camera to active display surface interface 105 is attached with an epoxy adhesive. Other adhesives known in the art, such as polyurethane adhesives, silicon-based adhesives, anaerobic adhesives, or acrylic adhesives may also be used. Other techniques used to join materials to glass or plastic, such as diffusion welding or vacuum-based bonding, for example, suction cups, may also be used. Mechanical attachment to the monitor bezel and an offsetting mechanism may also be used. For example, a thin steel rod may extend from the edge of the bezel and substantially parallel to the display surface. The far end of the thin steel rod attaches mechanically or by other means known in the art to the video

camera. Alternatively, a glass or transparent plastic rod, sheet, or plate may be used to position the video camera at the appropriate location on the active portion of the video display. These methods alleviates the requirement that a direct bond between the video camera and the display surface exist.

In an alternative embodiment, a mirror is placed on the screen in lieu of the camera. The video camera is oriented so that it faces the mirror assembly. Optic properties of the mirror and the camera enlarge the apparent size of the mirror's image such that substantially all of the field of view of the video camera consists of the view as seen by the mirror. The mirror's orientation allows imagery located substantially in front of the display screen to be reflected into the lens of the video camera. By this design, a person located in front of the display will appear to be located in front of the video camera. Because reflection through a mirror reflects the orientation of the image as seen in the video camera, the image may need to be digitally manipulated so that it is reflected back to its proper orientation. Techniques to accomplish the manipulation are well known in the art. Furthermore, a concave reflecting surface is required to form an image with a sufficiently large field of view on the image sensor. Digital imaging techniques known in the art may be used to correct for any optical aberrations due to the curvature of the reflecting surface. The mirror assembly may be bonded to the display surface with an epoxy adhesive. Other adhesives known in the art, such as polyurethane adhesives, silicon-based adhesives, anaerobic adhesives, or acrylic adhesives may also be used. Other techniques used to join materials to glass or plastic, such as diffusion welding or vacuum-based bonding, for example, suction cups, may also be used. Mechanical attachment to the monitor bezel and an offsetting mechanism may also be used. For example, an thin steel rod may extend from the edge of the bezel and substantially parallel to the display surface. The far end of the thin steel rod may attach mechanically or by other means known in the art to the mirror assembly. Alternatively, a glass or transparent plastic rod, sheet, or plate may be used to position the mirror assembly at the appropriate location on the active portion of the video display. These methods alleviates the requirement that a direct bond between the mirror assembly and the display surface exist.

In lieu of the mirror, an alternative method of optical reflection such as a prism

may also be used.

This patent application incorporates by reference copending application 11/223675 (Sandberg). Matter essential to the understanding of the present application is contained therein. Specifically, by mounting a teleconferencing apparatus of the sort described in this invention on a remotely controlled platform as per pending patent 11/223675 (Sandberg) a user is able to project his presence to a remote location in a more realistic manner.

By being able to maintain direct eye contact with users in the apparatus' environment, the user's sense of presence, as well as the perception of presence by those near the apparatus is heightened. The apparatus described herein, combined with electro-mechanical actuators allowing the orientation and position of the apparatus to be changed (as described in pending patent 11/223675) can thus be used to greatly enhance the perception of presence in a remote telepresence device.

Ideally, both sides of a video teleconferencing call will use the apparatus described herein. If both sides are using the device, both sides will feel that the other user is looking them straight in the eye. This is preferable to when only side A of a conversation has the apparatus, which results in side B feeling like she is being looked at directly while side A will feel that side B is looking elsewhere.

When one side of a video teleconferencing call is aware that the other side is using the apparatus described herein he should make an effort to align his own head so that his eyes will be displayed near the camera on the remote user's apparatus. A visual cue can be provided to facilitate this. By showing the user his own image with a visual cue superimposed on top of it showing where his eyes should be, the user is easily able to move his head ( or the camera if the camera is movable ) so that his eyes are at the correct position for the remote user.

FIG. **2** illustrates the orientation of the local user and the remote user during a video conferencing session when using the invention. The monitor casing 201 is shown in a landscape orientation, but in an alternative embodiment, the portrait orientation may also be used. The removable video camera 203, is shown as a small tubular chassis with a cylindrical lens that faces away from the monitor active display surface 202. The video camera 203 is attached to the monitor active display surface

202 with a suction cup concealed within the tubular chassis 2023 A vacuum release button 204 is used to create and release vacuum within the suction cup, facilitating removal of the invention when not being used. The video camera power and data wire 205 is shown stretched over the top of the monitor. The wire continues until it connects to a personal computer or other computational device capable of accepting a video camera input, not shown. The local user 206 is shown with his gaze directed at the video camera. Due to the small angular distance between the video camera 203 and the remote user's eyes 207 depicted on the monitor active display surface 203 , the remote user perceives the local user 206 as looking substantially into his eyes.

In an alternative embodiment, a hole is made through the opaque layers of an LCD panel during the LCD manufacturing process. A CCD or CMOS imaging camera is located behind this hole, and via this means, the LCD screen appears to be unmodified. This also has the benefit of blocking less LCD screen "real-estate" as the hole in the LCD panel need only be small enough to provide enough light for the aperture of the imaging device's lens. The aperture surface area is, of course, smaller than the surface area required to house an entire video camera.

In yet another alternative embodiment, the imaging sensor and the LCD are both produced on the same glass screen substrate, thereby saving the extra step of integrating the imaging sensor and LCD.

In yet another alternative embodiment, thin metal wires or other transparent conductive materials that are invisible to the naked eye are used to transfer power and data to and from the imaging device. These wires can be routed along the LCD screen without being visible to the viewer. In an alternative embodiment, the wires can be routed along a transparent support structure that holds the camera assembly in a fixed location on the active portion of the monitor. In addition to extremely thin wires, conductive materials such as Poly(3,4-ethylenedioxythiophene)/poly(styrenesulfonate), which is available under the tradename Baytron® P, may also be used. Carbon nanotubes may also be used to form a transparent conductive electrical conduit.

FIG. **3** illustrates the slide mechanism embodiment of the invention. The monitor casing 301 is shown in a landscape orientation, but in an alternative

embodiment, the portrait orientation may also be used. The video camera 302, is shown as a small tubular chassis with a cylindrical lens that faces away from the monitor active display surface 303. The video camera 302 is attached to a slide mechanism 304. The slide mechanism 304 slides vertically through a slide mechanism retainer 305. In one embodiment of the invention, the slide mechanism retainer is permanently attached to the monitor casing using fabrication techniques known in the art. In the preferred embodiment, the slide mechanism retainer can be removed from the monitor casing. Various removable means of attachment such as vacuum bonding (for example, suction cups), hook-and-loop fastener, ledge and ledge grabber, or a friction interface may be used to removably attach the slide mechanism retainer. A user can position the video camera 302 either on-screen or off-screen by sliding the slide mechanism 304 up and down within the slide mechanism retainer 305. In the figure, the video camera 302 is shown on-screen, covering the video image of a remote user's forehead 306. Because of the proximity of the video camera 302 to the image of the remote user's eyes 307, the remote user will perceive the local user's gaze as being directed at him to a greater degree than if the camera 302 is located off-screen.

The preferred embodiment of this invention uses digital image stabilization techniques to maintain the video image of the remote user in a constant position on the video screen. This allows the image of the remote user to be stable in its preferred position, where the eyes of the remote user, when displayed on the local video screen, straddle the on-screen video camera. Techniques known in the art, similar to image stabilization, such as face-tracking, and jitter-reduction algorithms may also be used. In all cases, the techniques may be implemented as software executing on a processor, or using dedicated hardware, or a combination of both techniques. The image stabilization may be done at the remote location or at the local location. In the preferred embodiment the image stabilization would be done at the remote location, as this would have the added benefit of reducing the bandwidth of the transmitted video data due a reduction in frame-to-frame movement when using temporal video compression techniques. In the preferred embodiment, a face tracking algorithm known in the art, such as the algorithm presented by R.J. Qian, M.I. Sezan, and K.E.

Matthews in their paper, A robust real-time face tracking algorithm, (ISBN: 0-8186-8821-1, 1998) is used to determine the center of the location of a user's face in a video frame. This location is representing as an (x,y) offset from the upper left corner of the video frame. A preferred camera position on the local users face (typically at the bridge of the nose or slightly higher) is represented as an offset (u,v) from the center location of a user's face. When a video frame is displayed on the remote side, it is translated by (x-u, y-v) relative to the center of the video frame. This maintains the local user's face in a constant position, with the camera located in its preferred location as viewed by the remote user.

Advantages

What has been described is a new and improved video teleconferencing method and apparatus to preserve gaze across a video teleconferencing link by attaching a camera to the active video display surface.

While certain exemplary embodiments have been described in detail and shown in the accompanying drawings, it is to be understood that such embodiments are merely illustrative of and not restrictive on the broad invention, and that this invention is not to be limited to the specific arrangements and constructions shown and described, since various other modifications may occur to those with ordinary skill in the art.

What is claimed:

1. An apparatus for preserving gaze across a video teleconferencing link consisting of:

A video teleconferencing device consisting of a video screen, video camera, and video teleconferencing processing element, wherein the video camera is mounted on an active portion of the video screen.

2. The apparatus of claim 1, further comprising:

a video camera data wire for receiving data from the video camera, the video camera data wire being substantially invisible to the unaided human eye.

3. The apparatus of claim 1, wherein the video camera is mounted on a location of the video screen that is displaying the bridge of the nose of a remote video teleconferencing user.

4. The apparatus of claim 1, wherein image stabilization algorithms are used to stabilize a remote video teleconferencing user's position on the video screen.

5. The apparatus of claim 1, wherein the remote user can maintain eye contact with anyone in the apparatus' space by controlling the location or orientation of the apparatus through the use of electro-mechanical actuators.

6. The apparatus of claim 1, further comprising:

An eye level alignment visual cue presented to the remote user whereby the remote user may position his head such that the image of his eyes on the remote screen are coincident with the camera on the apparatus.

7. An apparatus for preserving gaze across a video teleconferencing link consisting of:

A video teleconferencing device consisting of a video screen, video camera, and video teleconferencing processing element; and

A reflective device mounted on the active portion of the video screen, wherein the reflective device reflects an image that is substantially orthogonal to the video

screen surface into the lens of the video camera.

8. The apparatus of claim 7, wherein the reflective device is a mirror.

9. The apparatus of claim 7, wherein the reflective device is a prism.

10. The apparatus of claim 7, wherein the reflective device is mounted on a location of the video screen that is displaying the bridge of the nose of a remote video teleconferencing user.

11. The apparatus of claim 7, wherein image stabilization algorithms are used to stabilize a remote video teleconferencing user's position on the video screen.

12. The apparatus of claim 7, wherein the remote user can maintain eye contact with anyone in the apparatus' space by controlling the location or orientation of the apparatus through the use of electro-mechanical actuators.

13. An image display and capture device for preserving gaze across a video teleconferencing link, consisting of:

a video screen, the video screen composed of multiple layer of materials, wherein all of the opaque layers of material include a small hole sufficient to allow visible light frequencies to pass through to the back of the video screen; and

a video camera mounted behind the video screen, such that the aperture of the video camera is substantially aligned with the small hole on the video screen, whereby an image in front of the video screen may be detected through the small hole on the video screen by the video camera;

14. The apparatus of claim 13, wherein the remote user can maintain eye contact with anyone in the apparatus' space by controlling the location or orientation of the apparatus through the use of electro-mechanical actuators.

15. An video camera apparatus for preserving gaze across a video teleconferencing link consisting of:

a lens;

a lens holder holding the lens;

an imaging sensor integrated circuit capturing light from the lens;

a video screen attachment mechanism attaching the video camera to an active display surface; and

a translating mechanism permitting movement of the lens to two positions, the first position obscuring a users view of a portion of the video screen, and the second position not obscuring the users view of the video screen.

16. The apparatus of claim 15, wherein the video screen attachment mechanism removably attaches the video camera to an active display surface.

17. The apparatus of claim 15, wherein the remote user can maintain eye contact with anyone in the apparatus' space by controlling the location or orientation of the apparatus through the use of electro-mechanical actuators.

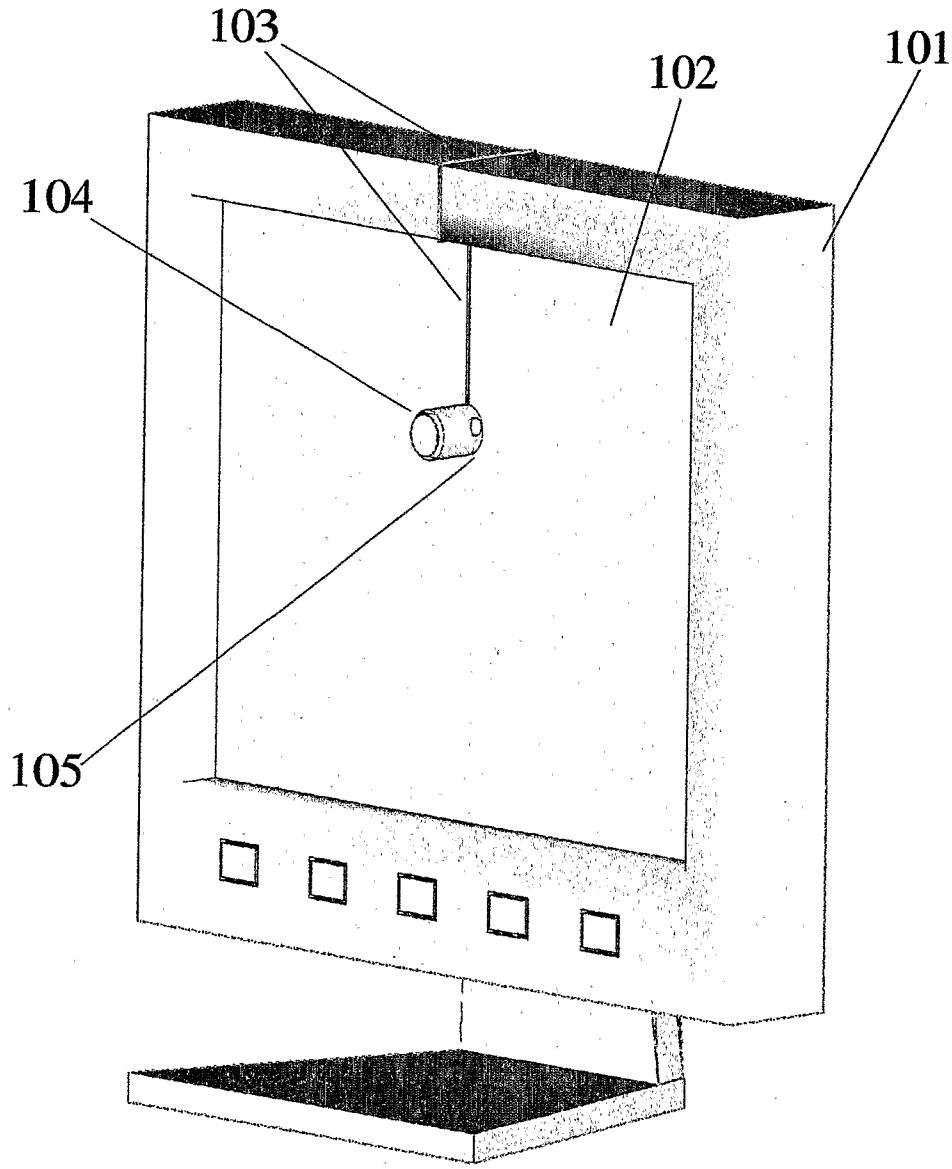



Fig. 1

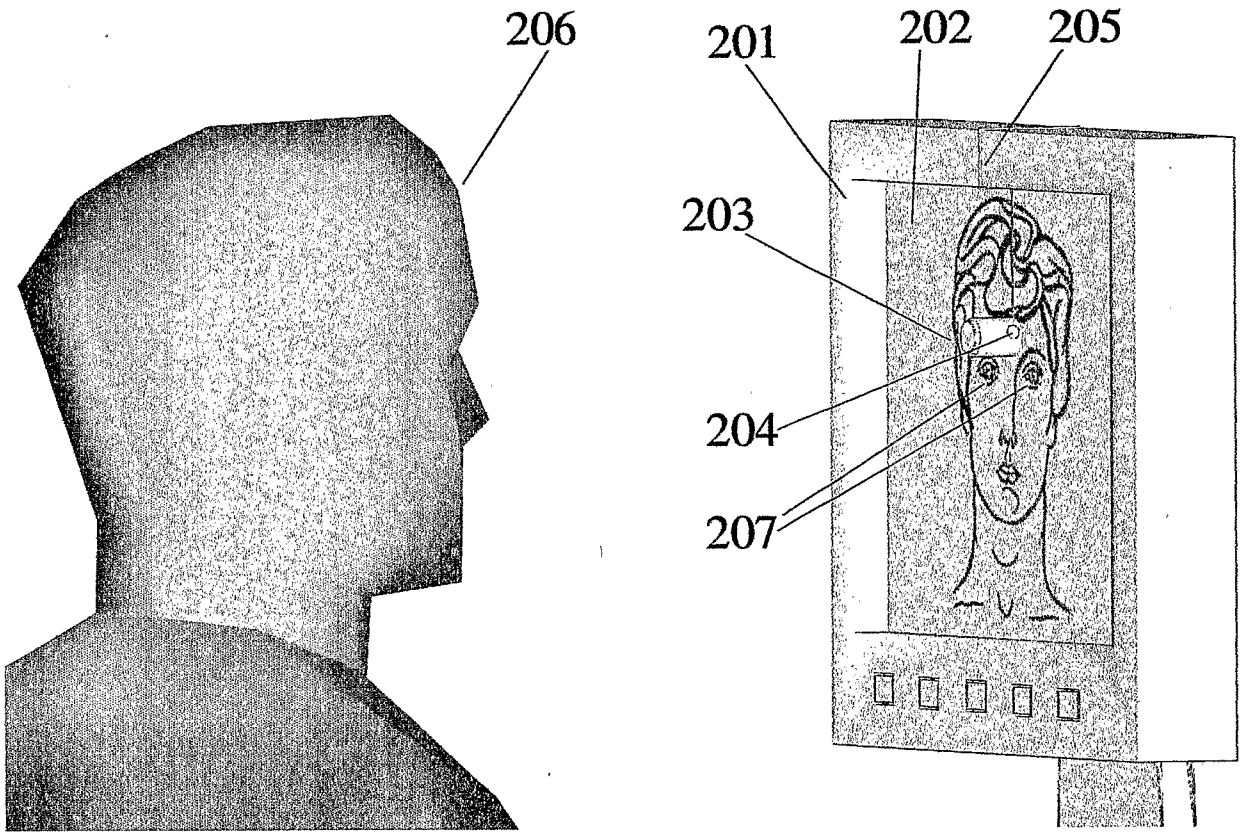



Fig. 2

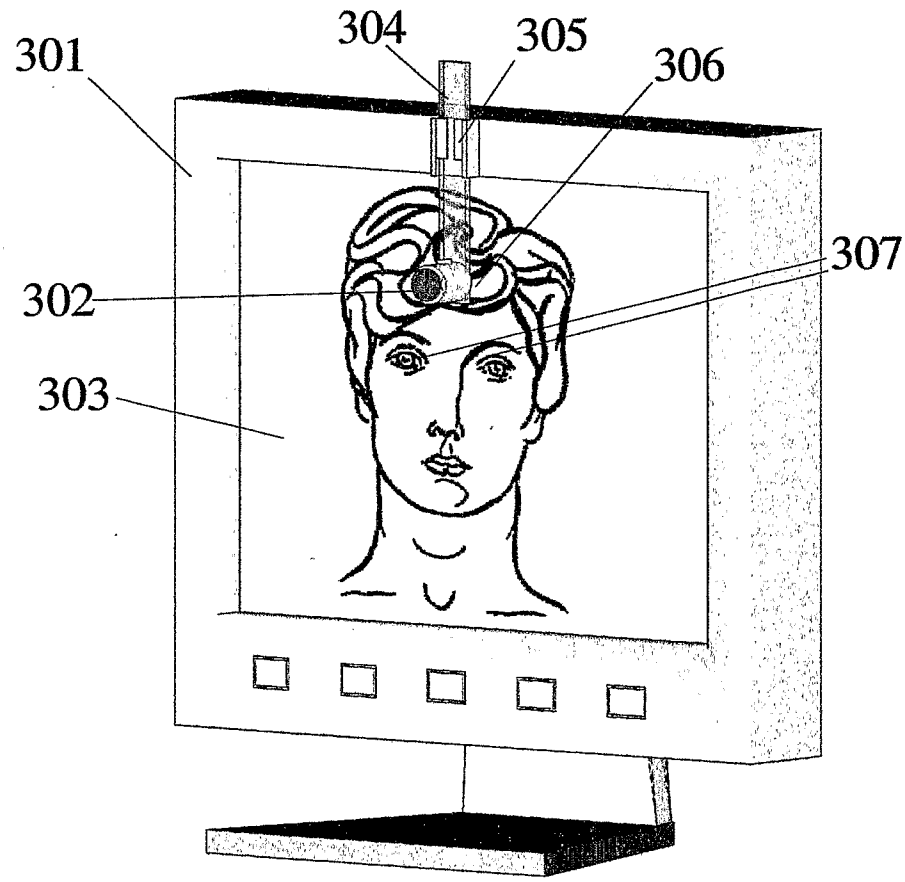



Fig. 3