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

(19) World Intellectual Property Organization
International Bureau

(43) International Publication Date
28 September 2006 (28.09.2006)

PCT

(10) International Publication Number
WO 2006/100645 A2

(51) International Patent Classification:
G06F 3/01 (2006.01)

(21) International Application Number:
PCT/IB2006/050872

(22) International Filing Date: 21 March 2006 (21.03.2006)

(25) Filing Language: English

(26) Publication Language: English

(30) Priority Data:
60/665,023 24 March 2005 (24.03.2005) US
60/710,950 24 August 2005 (24.08.2005) US

(71) Applicant *(for all designated States except US)*: **KONINKLIJKE PHILIPS ELECTRONICS, N.V.** [NL/NL]; Groenewoudseweg 1, NL-5621 BA Eindhoven (NL).

(71) Applicant *(for AE only)*: **U.S. PHILIPS CORPORATION** [US/US]; 1251 Avenue of the Americas, New York, NY 10020 (US).

(72) Inventor; and

(75) Inventor/Applicant *(for US only)*: **CORTENRAAD, Hubertus M. R.** [NL/NL]; Groenewoudseweg 1, NL-5621 BA Eindhoven (NL).

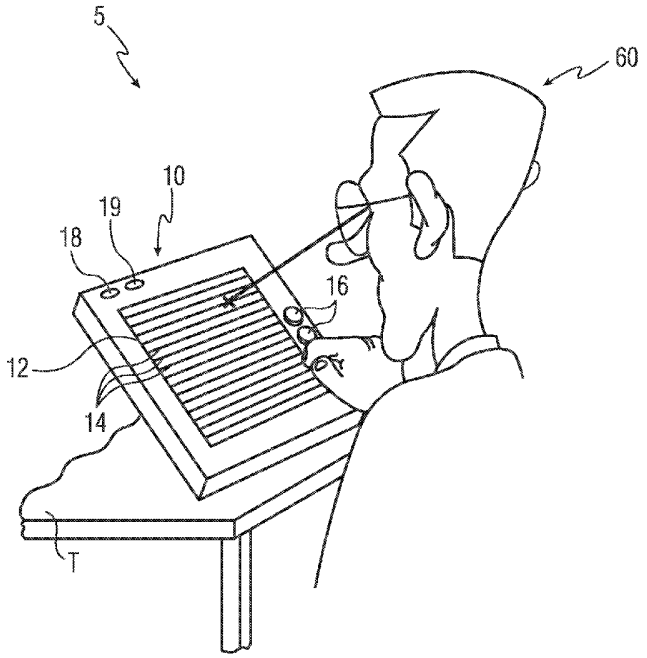
(74) Common Representative: **KONINKLIJKE PHILIPS ELECTRONICS, N.V.**; c/o Edward W. Goodman, P.O. Box 3001, Briarcliff Manor, NY 10510-8001 (US).

(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, HR, HU, ID, IL, IN, IS, JP, KE, KG, KM, KN, KP, KR, KZ, LC, LK, LR, LS, LT, LU, LV, LY, MA, MD, MG, MK, MN, MW, MX, MZ, NA, NG, NI, NO, NZ, OM, PG, PH, PL, PT, RO, RU, SC, SD, SE, SG, SK, SL, SM, SY, TJ, TM, TN, TR, TT, TZ, UA, UG, US, UZ, VC, VN, YU, 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),

[Continued on next page]

(54) Title: IMMERSIVE READING EXPERIENCE USING EYE TRACKING



(57) Abstract: An immersive reading experience is provided for a reader. In a system (5, 5a, 5b) for providing such an immersive reading experience, at least one camera (18, 18a) captures images of the user reading a literary work (10a, 14). At least one processor (20, 20a) receives the images from the camera (18, 18a), and the processor (20, 20a) determines the location (X) of the text (14, 14a) currently being read by the user (60) using the images. At least one immersive effect is output corresponding to the current text being read by the user (60).

WO 2006/100645 A2

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).

Declarations under Rule 4.17:

— *as to applicant's entitlement to apply for and be granted a patent (Rule 4.17(ii))*

— *as to the applicant's entitlement to claim the priority of the earlier application (Rule 4.17(iii))*

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.

IMMERSIVE READING EXPERIENCE USING EYE TRACKING

The entire contents of the following two commonly owned U.S. provisional patent applications are hereby incorporated by reference herein: (1) U.S. provisional patent application ser. no. 60/665,023, entitled "Immersive Reading Experience By Means Of Eye-Tracking" by Hubertus M. R. Cortenraad, filed March 24, 2005; and (2) U.S. provisional patent application ser. no. 60/665,025, entitled "Orientation and Position Adaptation for Immersive Experiences" by Hubertus M. R. Cortenraad and Anthonie H. Bergman, filed March 24, 2005.

The invention generally relates to the creation of immersive experiences in the context of reading.

Systems for delivering immersive experiences for traditional televisions have recently been developed. The Philips Ambilight system has been developed for traditional home television sets that are at a fixed location in the home, for example. Built-in lighting around the periphery of the television set is coordinated with the content on the television to give the viewer a more immersive experience. For example, if a soccer game is being viewed (where much of the television display is green grass), the peripheral lighting projects a green hue on the walls or other surfaces immediately surrounding the television, thus giving the viewer the feeling that the display is actually larger and more continuous.

There are also various enhancements that have been directed at reading of books. For example, "Parent and Child Reading, Designing for an Interactive, Dimensional Reading Experience" by Herve Gomez describes sound interactivity that includes a default background for each page, consistent with the general context of that page, and refers to the sound events being triggered according to the properties of the room that the reader is in. An example given by Gomez for a scene in a book where bears are sleeping in a cave is augmenting the reader's environment using snoring sounds, water drops falling and reverberating any sounds in the reader's room through a loudspeaker.

When considering default sound background that corresponds to an entire page, however, should scenes in the book change within a page, for example, the sound for the page will no longer correspond to the text currently being read.

An object of the present invention comprises providing a reading experience that provides a well-correlated immersive experience. An object of the present invention also

comprises providing an immersive reading experience where the immersive experience is based upon the current content of text being read on a page by the user.

In accordance with one aspect of the present invention, immersive reading effects and immersive reading experiences correspond to particular text that the reader is reading. In one of many embodiments, eye tracking is used to determine the location of the material currently being read by the user, and data for generating immersive effects for the material is available and is correlated to the location in the material. Using the detected location in the material, the particular data for generating immersive effects corresponding to the material currently being read is identified. The particular data is used to generate an immersive experience corresponding to the material currently being read.

As applied herein to a reading experience, an "immersive effect" generally comprises an evocation of one or more senses of the user separate from the reading itself and in a manner that corresponds to the content of the material being read by the user. An immersive effect generally comprises an intelligent re-presentation, supplementation and/or enhancement of the content of the reading material that is scaled and presented to the user in the real-world frame of reference. An "immersive experience" presented to the user is the totality of one or more immersive effects. For reading material, immersive effects (and the resulting immersive experience) will generally serve to give the user the sensation of being immersed or physically involved in the material he/she is reading. For example, sound and light may be used to provide an immersive experience corresponding to a particular passage that the user is reading and may suggest or give the user the feeling that they are in the scene being read in the book, instead of just reading about it.

In addition, embodiments of the invention also comprise an immersive reading experience that uses the content of the material currently being read and is also generated based on the current orientation and/or position of the material and/or user in a room.

In one embodiment of a system for providing an immersive reading experience falling within the scope of the invention, at least one camera captures images of the user reading a literary work. At least one processor receives the images from the camera, and the processor determines the location of the text currently being read by the user using the images. Using the determined location, at least one immersive effect is output corresponding to the current text being read by the user. In an embodiment of a method falling within the scope of the invention, the reader's eyes are tracked, and the location of

the portion of text currently being read by the reader is determined from the tracking. Based on the determined location, at least one immersive effect is output to the reader corresponding to the portion of text currently being read by the reader.

Fig. 1 is a perspective view of an electronic book in accordance with an embodiment falling within the scope of the invention;

Fig. 1a depicts a number of components included in the e-book of Fig. 1;

Fig. 1b represents a particular technique for calibration of eye tracking that may be used with embodiments of the invention;

Fig. 1c represents a portion of the calibration of Fig. 1b;

Fig. 1d depicts the e-book of Fig. 1 in a setting for generating an immersive experience in accordance with an embodiment falling within the scope of the invention;

Fig. 2 is a perspective view of a paper book and related components in accordance with an embodiment falling within the scope of the invention;

Fig. 2a depicts a number of internal components of a component shown in Fig. 2; and

Fig. 3 is a flow chart of a method in accordance with an embodiment of the invention.

Fig. 1 depicts an electronic book 10 (e-book) according to one embodiment. (As will become clearer from the description below, electronic book 10 may be viewed as a system 5 in itself, or may be included as part of a larger system, such as shown in Fig. 1d.) As shown, the e-book 10 is supported at an angle on a table T and is being read by a user 60. The e-book 10 has internal electronic components that generate and display textual pages of a book electronically on a display 12. The particular page of the book currently being shown on the display of Fig. 1 is represented by the lines 14 of text of the page. User 60 is shown reading a particular portion of text on the current page, designated by "X" in Fig. 1. Electronic book 10 also includes user controls 16. These controls may include keys that provide for navigation of a menu for selecting a written work (e.g., a fictional book) from a memory of the book 10 for display on display 12. It may also include keys that allow the user 60 to page forward or backward through pages of the selected work, scrolling through the text of the selected book, or for other navigation. E-book 10 also includes camera 18 and infrared (IR) LED 19 that is used to track the user's eye movement

to determine the location on the displayed page currently being viewed, as explained further below.

Fig. 1a is a representative depiction of a number of underlying components of the e-book 10 of Fig. 1. E-book 10 includes a processor 20 and memory 22, which typically includes working memory and storage. Processor 20 is configured for the processing hereinafter described, for example, via programming provided by appropriate software stored in memory 22, firmware, or other programming sources. Among other things, memory 22 may contain a book (or generally a literary work) whose pages are currently being displayed, as well as store a library of other literary works that may be selected by the user 60. (However, e-book 10 may receive material comprising a work over an interface from remote storage when called for by a user.) As represented in Fig. 1a, processor 20 interfaces with display 12 to display menus, pages of a selected work, or other displayed material of e-book 10. Processor 20 also interfaces with controls 16 to select a work for display, move through pages in the selected work, and like control of the e-book 10.

In addition, processor 20 receives images from camera 18 and processes them using software stored in memory 22 that in the embodiment applies an available eye tracking technique in a manner to determine what portion of the display 12 the user 60 is viewing. In a preferred embodiment, processor 20 controls LED 19 to output IR light, at least some of which is directed toward a user 60 reading the e-book 10, as in Fig. 1. Due to the curvature of the cornea, in the images of the user 60 captured by camera 18 the IR light incident on the eye will appear as a small distinct white spot on the eye. The spots (known as Purkinje images) are easily distinguishable from the dark pupil of the eye.

In the case of many typical persons reading textual material, their head movement is generally negligible and reading is principally done by rotation of the eyes alone (in the eye sockets). As the eye moves (i.e., changes viewing direction by rotating in the eye socket), the position of the pupil changes with respect to the IR spot in the images. Generally, the pupil will appear to be moving, whereas the reflected position of the IR spot appears relatively stationary in the images. For a camera and IR source having fixed positions, and a relatively steady head, the position of the IR spot with respect to pupil is substantially a function of the viewing direction (or eye movement) alone. In addition, basic geometric consideration of the convex curvature of the cornea also demonstrates that

the position of the IR spot with respect to the pupil will be unique for any given viewing direction. Thus, for the determinations made herein in the context of a user reading a book (such as e-book in Fig. 1), IR spot position with respect to the pupil is regarded as uniquely correlated to the location on the page being viewed. (For persons having larger head movement during reading, more complex eye tracking techniques can be applied, such as using higher order Pukinje reflections.)

Processor 20 and eye tracking software are initially calibrated based on the positions of spots in the images at various known reference points on the page on display 12. For example, a manual calibration may initially be undertaken where the user 60 is directed to look at various locations of the page (e.g., the four corners). The positions of the four reflected IR spots with respect to the pupil in the images captured by camera correspond to the four corners of the text and are used as reference points. (Typically, the center of the IR spot position with respect to the center of the pupil is used in the images, although other points may be used.) When the viewer is reading text, the current position of the spot on the eye as detected in the images by processor 20 will fall within the four reference spot positions on the eye. Processor 20 determines a current reading location on the page for the current spot position by interpolation using one or more of the reference IR spot positions and the corresponding corners of the text.

Fig. 1b is a representative depiction of the four reference IR spots with respect to the pupil on an eye when looking at the four corners of a page, such as displayed page of e-book 10. In each representation of the eye, the pupil is represented by the dark circle and the IR spot by the light circle. The eye labeled UL represents the relative IR spot position when looking at the upper left corner of the page, LL represents the position when looking at the lower left corner, UR the upper right corner, and LR the lower right corner. These reference positions for the IR spots in the image corresponding to the four page corners are stored in memory 22. The stored positions of Fig. 1b are shown together as points with respect to the pupil on the eye in Fig. 1c, and approximately outline a box on the eye. (The squareness of the box will vary with the tilt of the display, but for convenience is depicted as a square in Fig 1c.) For a viewer viewing a location on the page in Fig. 1b (e.g., position X), the IR spot will lie within this box on the eye (e.g., at position x in Fig. 1c). The relative location X on the page currently being read is readily interpolated from one or more of the corners of the page using the position of x in the current image in relation to

one or more of the corresponding reference IR spot positions (LL, LR, UL and/or UR). Depending on the system, the relative position on the page may then be translated to an absolute position (e.g., coordinates) or other position measurement (e.g., line number).

Preferably, the calibration points are selected so as to "frame" the text on the page. Location points on the text for the eye tracking calibration may be chosen so that they align with the position parameter used to reference corresponding immersive effects for a data set, described below. Additional points on the page may also be used to provide a more comprehensive and accurate calibration. It is noted that for a paper book (as in the embodiment described below) or other like reading material, when the book is opened there are two successive pages are viewable at a time. Thus, the calibration will generally include more reference points, for example, four for the corners of each opposing page.

Current eye tracking software for the reading setting as described above (for a substantially steady head) is able to calculate a viewing direction on the order of 1°. At average reading distances, this provides for an accuracy in determination of the user's reading location on display 12 on the order of 1 cm. For standard text of a normal reading size displayed on display 12, such eye tracking is thus capable of determining the location approximately to an accuracy of the line and words within the line that the user is currently reading, if needed. (As noted below, many immersive reading effects will be based on larger passages of text and such a high level of detection is not necessary. Moreover, even for small text where the accuracy of the eye tracking is within two or three lines, the immersive effect will still substantially correspond to the reading location, considering the speed at which most people read a line.) As noted, additional calibration points (e.g., a grid on the display of e-book) may be used to improve the tracking even further. Also, additional cameras and/or additional LEDs may be utilized along with additional processing in analogous manner to provide a more accurate determination.

As noted, processor 20 of e-book 10 electronically provides the text of work being displayed on display 12. Processor 20 in this particular embodiment also generates the reference points used in the calibration of the eye tracking, for example, by successively illuminating the reference point locations on the display and capturing the corresponding reference IR spot positions on the eye of user 60. Processor 20 thus effectively knows the coordinate location of the reference points on display. (The coordinate positions are effectively known to processor 20 through its internal addressing arrangement with display

12, for example.) Thus, when an IR spot is detected in an image for a current reading position, processor 20 is able to interpolate to a coordinate location on display 12 for the current reading location. The calibration point locations may be the four corners framing the text on the page, as described above.

Also stored in memory 22 for the displayed work is data that provides for immersive effects corresponding to the text, referred to as the "effects data set". In the case of the e-book embodiment of Figs. 1 and 1a, memory 22 stores an electronic version of the text of a work that processor 20 uses to display on display 12. In that case, effects-related data set is preferably included in a software layer or subcode of the electronic version of the text that is not visible to the user 60.

Processor 20 retrieves text in text layer of the electronic version from memory 22 and formats the appropriate portion for a page on display 12. Thus, processor 20 also effectively knows (through internal addressing with display, for example), the text that is currently being displayed at any coordinate location on display. As described above, as the user 60 reads the text being displayed on display 12, processor 20 tracks the user's eye by detecting the position of the IR spot in the images from camera 18, and interpolates to determine the current coordinate location being viewed on display 12. Processor 20 uses the location currently being viewed to identify the text it is currently displaying at that location of the display 12. Processor 20 uses the memory location of the text being displayed to retrieve the corresponding immersive effect data from the sublayer for the identified text. The data is used to provide an immersive effect corresponding to the text at the location currently being viewed.

The sublayer containing effects data (or a general effects-related data set) may directly include immersive effects instructions for the corresponding text. It may alternatively contain scripts that provide for description of the scenes of the literary text, which may be processed (translated) into corresponding immersive effects instructions (e.g., to adjust and control lights, speakers and/or other sensory presentation device outputs in a manner that reflects the described scene). It is noted a data set or layer may qualify as an effects-related data set or layer even if that is not its principle usage. The Philips Physical Markup Language (PML) is particularly suitable for implementation as an effects-related layer. Alternatively, an intelligent translation program may use the current text in

the text layer and translate it using word and phrase recognition to one or more appropriate immersive effects.

Fig. 1d depicts a system 5a that comprises the e-book 10 described in Figs. 1 and 1a situated in an illustrative setting that provides an immersive reading experience to the user 60. Fig. 1d is a view from above of user 60 positioned at the center of a space that is uniformly surrounded by four speakers 30a-30d and four lights 40a-40d as shown. E-book 10 has communication interfaces with speakers 30a-30d and lights 40a-40d through a wired or wireless connection (omitted for clarity in Fig. 1d). Speakers 30a-30d and lights 40a-40d also include appropriate electronics that allow for control via signaling received from e-book 10 via the connections. E-book 10 is set at a predetermined orientation on table T while the reader is reading (e.g., via alignment with a template on table T).

As described above, processor 20 uses the current location being viewed on the display to identify the corresponding electronic text in memory, and to withdraw the corresponding effects data for the text from the effects layer. The effects data is used to generate an immersive experience corresponding to the current text being read. Processor 20 may first format or convert the effects data into control instructions for the speakers and lights. For example, where the text user 60 is currently reading is describing a tranquil sunrise over a lake, the corresponding immersive effects data withdrawn from the sublayer for the display position may be instructions to slowly raise frontal lighting and also generate low surrounding sound of water ripples. Processor 20 is programmed to know the type of immersive effects that speakers 30a-30d and lights 40a-40d can generate and also knows the relative positions of these devices with respect to the predetermined orientation of e-book (for example, through pre-programming or manual input). Speakers 30a-30d and lights 40a-40d face toward the position of e-book or are omnidirectional. Thus, for example, processor 20 outputs a signal to light 40a in front of user 60 to slowly raise the front lighting level, corresponding to the sunrise in the current text, and signals are also output to each of speakers 30a-30d to generate the rippling of water sounds to user 60. The user 60 is consequently provided with an immersive experience corresponding to the text currently being read.

Generally, immersive effect(s) once generated will correspond to a segment of text. For example, the literary text corresponding to the sunrise scene may go on for a paragraph, and the corresponding immersive effect can thus be generated once at the

beginning of the scene and remain until there is a change in the literary text. As a result, the effects-related layer (or, generally, the effects related data set) may reflect a lesser amount of data needed in relation to the text layer. Of course, the effects-related layer may occasionally repeat prior effects data already initiated for a segment and include data to terminate an implemented effect. In addition, where a scene rapidly changes or calls for additional effects as it progresses, the effects-related layer includes relatively more data. As the user 60 reads forward in the text, the reading position on the display continues to be tracked and the corresponding immersive effects data determined from the effects-related layer are updated. When the updated data reflect a change in immersive effect, the updated immersive effect(s) is output. If, for example, the text being read shifts to a thunderstorm at night, the corresponding effects data for that text position on the display may include instructions to generate random thunder sounds and flashes of light. Speakers 30a-30d and lights 40a-40d are controlled to provide these effects.

The e-book embodiment of Figs. 1 and 1a provides a number of advantages in terms of calibration and operation. Processor 20 uses known positions on display for eye tracking calibration, and the resulting determination of current viewing location can be internally correlated with the text currently displayed at that determined location. (The text currently being displayed at the viewing position is the text currently being viewed by the reader.) The text layer is used to identify the corresponding immersive effects in the sublayer for the current viewing position. Thus, processor 20 effectively internally aligns the current reading location as determined through eye tracking with the position parameter used to withdraw the corresponding immersive effect data from memory.

Fig. 2 provides an alternative embodiment of the invention, where a paper book 10a is being read in lieu of the e-book 10 of Fig. 1. (Analogous to Fig. 1, the components of Fig. 2 described below (together or apart from paper book 10a) may also be viewed as a system 5b in itself, as well as part of a larger system, such as shown in Fig. 1d.) Book 10a is shown resting on table T. (For ease of description book 10a is again shown as being propped up, such as via an underlying book holder, but may alternatively be lying flat on table. The book is currently turned to a particular page being read, as represented by lines 14a of the page. User 60 is currently reading a particular portion of the text, denoted by "X". Camera 18a and IR LED 19a for capturing images for eye tracking clips to the side of book 10a (such as halfway down the left-hand page as shown), and communicates with a

separate computer 100 over a wired or wireless interface 100a. As shown in Fig. 2a, computer 100a comprises processor 20a and memory 22a, which perform eye tracking of user 60. Memory 22a also includes an effects-related data set, described further below.

In this embodiment, memory 22a generally does not need and thus generally does not store an electronic version of the text, since the text exists on the pages of the book. The immersive effects data in this embodiment will thus typically be stored in a separate immersive effects data set instead of a text sublayer. The immersive effects data included will preferably be referenced in memory according to the page and position on the page of the paper book where the corresponding text appears. The position parameter used in the immersive effects data set in this example is given in relative terms, for example, the vertical position can be the percentage from the top margin to the bottom margin, and the horizontal position can be the percentage from the left margin to the right margin.

Camera 18a and IR LED 19a are clipped to the side of the book at a position chosen by user 60, which remains fixed during calibration and subsequent tracking. An analogous calibration procedure is used by processor 20a to correlate reference IR spot positions in the camera images with reference point locations on the page (such as the four corners of the page). The calibration point locations are chosen so that they align with the position parameters in the immersive effects data set. For example, if positions in the immersive effects data set are based on vertical and horizontal percentages from the margins as noted above, then the four corners used in the calibration of the eye tracking are preferably chosen to lie on the intersections of the margins. Such alignment between the calibration of the eye tracking and the immersive effect data set allows the location determined via eye tracking to be used directly as the position parameter to identify the corresponding immersive effect in the data set.

In the case of a paper book embodiment, calibration will often utilize a user interface (not shown) to guide the user through the calibration. As one specific example, such a user interface can make use of speakers that may be used for the immersive effects to give instructions for the calibration, e.g., "now look at top left corner". Where the positions in the immersive effects data set is based on the margins of the text as described above, in order to align the eye tracking with the data the user 60 may be instructed to look at particular words on one or more particular pages. For example, the first word in the first line of page 3 of the text may coincide with both the upper and left hand margins (i.e., not

indented). Thus, the user 60 may be instructed to turn to page 3 and look at the first word in the first line, which is used as the upper left calibration point. Calibration points for the other corners may be chosen to align with the data set positions in like manner. Such alignment data may be included in the immersive effects data set.

The reference IR spot positions and corresponding locations on the page are subsequently used to determine the user's current reading location X on the page. Processor 20a detects the IR spot in the image for a current reading location, and interpolates the current reading location on the page using one or more of the reference spot positions and corresponding reference location(s) on the page. As noted, the reference point locations on the page are chosen to correspond to the intersections of the margins for this embodiment. Thus, the current viewing location on the page determined through eye tracking may be determined through interpolation as a percentage down from the top margin toward the bottom margin, and a percentage from the left margin toward the right margin. Accordingly, the current viewing location (determined via eye tracking) and the position parameter of the immersive effects data set are both given in terms of percentage from the margins, and they are also aligned via the calibration. If the page number is known, the page number and detected viewing location may thus be used to identify the effects data in the data set corresponding to the text currently being read.

An initial page number being viewed may be manually input and processor 20a may keep track of which page is being read by tracking movements of adjacent page edges in the edges of the image captured by camera 18a. Alternatively, turning of pages may be tracked based on a reader's eye moving from the lower right position on a page to the upper left position. For this type of tracking of the page, other user 60 movements in the image (e.g., minor head movements resulting from movement of the arm to turn the page) may also be used to confirm that the user is paging forward in the book, determine that the user 60 is turning back to a prior page, or that the user 60 is simply re-reading material without turning the page. (If a user 60 turns multiple pages using this technique, another manual input of page number will typically be required.)

Thus, computer 100 interpolates the current location on the page the viewer is reading through eye-tracking and also has the page number being viewed. The corresponding effects data for that determined location and page number is retrieved by processor 20a from the effects data set stored in memory 22a. This retrieved effects data

corresponds to the text at that page and location on the page and thus corresponds to the text currently being read by user 60.

The retrieved effects data for the text currently being read is used to generate immersive effects in manner analogous to that described above for the e-book 10 embodiment, including the particular setting described with respect to Fig. 1d. As the user 60 reads further on in the page, images from camera 18a are used to eye track and update the viewing location on the page, the location is correlated by processor 20a to the corresponding effects data in the effects data set, and is used to generate updated immersive effects corresponding to the text currently being read. When the user 60 turns a page, computer 100 receives the updated page as described above, and such immersive effect processing continues for text currently being read by user 60 on the new page.

Alternatives may be used in place of the techniques described above for determining the current page being viewed. For example, a camera (not shown) from above may be used to capture images of the page itself and image processing may be used to recognize the page number the book is opened to. Such processing of the images may be performed in the computer 100 or in a separate processor (not shown), with the page number transmitted to computer 100. Other alternative techniques that provide the page to which the book is turned may be used, for example, one available technique senses circuitry in each page. Of course, an ongoing manual input of the page by the user 60 via a user interface is another alternative.

Although the book of Figs. 2 and 2a is referred to as a "paper" book above, it is noted that the pages of the book of this embodiment can alternatively be generated electronically. Although an electronic book corresponding to this embodiment may supply the pages electronically, the immersive effects data may (for example) be in a separate effects data set and not a sublayer of the electronic text. (Thus, the electronic book may not have or be able to exploit the calibration and processing attendant to the sublayer described in the embodiment of Figs. 1 and 1a.) The techniques for calibration and generating immersive effects for this electronic book may generally correspond to the paper book of this embodiment. However, the displayed page may be known internally to processor 20a.

Updating and re-calibration may also be done automatically by tracking eye movement during reading. For example, just after turning a page, the eye will move along

the top line, which generally corresponds to the top margin. Right to left eye movements on a page generally signify the beginning of a new line, which for most lines on a page begins at the left margin. For any one particular line, the location just prior to a right to left movement may not exactly coincide with the right margin, so a sampling of lines may be used to adequately determine the right margin. The movement of the eye just prior to turning a page generally corresponds to the bottom margin, although a sampling of pages may also be used. Such eye movement correlated to such known viewing positions on the page may be used to re-calibrate (or to initially calibrate) the eye tracking.

In the particular embodiments given above, speakers and lights are used as examples of devices that provide immersive effects to user. However, any controllable device that presents a sensation to a user may be used to provide immersive effects (and will generically be referred to as "presentation devices"). Other visual devices apart from lights that are readily recognized as presentation devices include, for example, video displays and video projectors. Many devices other than speakers can produce audio effects, for example, alarms, a clock ticking, and even pyrotechnic devices. Presentation devices also include many devices that invoke a tactile sensation, for example, fans, misting and fog generators, seat vibrators (rumblers), ultrasonic generators, and the like. Tactile devices may also include temperature control devices, for example, air conditioning and heaters. Presentation devices also include devices that generate an olfactory sensation, for example, a misting or smoke device that is controllable to present odors. Similarly, presentation devices also include devices that invoke a taste sensation. For example, misting or smoke devices may also be used to invoke tastes. Other specific presentation devices are readily identifiable. Where the current effects data provides an immersive effect for another type of available device, the effect may be output in creating the immersive experience.

Where a presentation device is not available (including not properly located) for creating the required immersive effect, the pertinent processor will generally determine to forego the effect. If an alternative presentation device exists that may provide a suitable corresponding immersive effect, the processor may adapt the effect to the available device. In a very simple case, the presentation device is simply a set of headphones worn by the user that is controlled to provide aural immersive effects corresponding to the text. As

known in the art, intelligent processing can be used to convincingly imitate 5.1 Dolby Surround sound on a set of headphones.

In the above embodiments, the described processing is carried out in one processor. The various processing and storage tasks may reside in different components. For example, for the e-book embodiment of Figs. 1-1a in the setting of Fig. 1d, the e-book 10 may perform all the noted processing tasks up to and including the generation of the immersive effects data for the text being read using the corresponding data from the immersive effects layer. A separate computer may receive the data and control presentation devices to provide the appropriate immersive effect. Analogous divisions of tasks may be used for the paper book embodiment of Figs. 2-2a.

The eye tracking locations generally needs to correspond (align) with the position parameter used with the immersive effects data set. In the above embodiments, the calibration of the eye tracking is done in a manner so that the determined viewing location can be used directly to identify the corresponding effects data (in the effects data set or the sublayer). Other alternative techniques may be used. For example, the eye tracking calibration may be done independently, and the effects data set may include a position for the first word on each page to align the eye tracking locations for that page with the position parameter for the immersive effects data set.

In addition, as the embodiment of the paper book above demonstrates, the eye tracking locations on the page may be given in relative terms and may be used with an immersive effects data set referenced by relative positions. Thus, the actual dimensions of the book do not have to be known and the coordinate locations do not have to be determined. Other embodiments may use the dimensions of the page and/or coordinate or other actual locations in the eye tracking and/or in the position parameter of the immersive effect data set. For example, the immersive effects data set may be referenced by coordinate position on the page, or line number and horizontal line coordinate. In that case, relative locations in an eye tracking determination may be translated to absolute positions. (Such translation may use, for example, the dimensions of the page, numbers of lines per page, and/or other like parameter.)

It is generally preferable that the immersive experience be suspended when the user is not actually reading text. For example, if the user is simply glancing over the page (or pages (e.g., to locate a particular passage), generation of corresponding immersive effects

will be incoherent. The eye tracking processing may include additional features to confirm that the user is actually reading text on the page. For example, the processing may additionally monitor the user's eye movement (for example, via the IR spot in the sequence of current images in relation to the reference points) and generate immersive effects when the user's eyes are moving in a left to right and top to bottom pattern typical of reading text. Of course, if the user's eyes stray outside the dimensions of the page (for example, as detected by interpolation of the IR spot in the current images), the current effects may be temporarily frozen, faded over time, or treated according to another setting or user preference. Similarly, if the eye tracking detects a break in the normal reading pattern, the immersive effect currently being output may, for example, be terminated, faded, or continued. If the eye tracking detects other discernible activity, such as the user re-reading text on a page (signaled, for example, by a break in the normal reading sequence such that the eyes move back to previous text and restart a typical reading pattern without a turning of the page), the effects for the material being re-read may be re-generated, the effects may be suspended while re-reading, or other action may be taken. An input user preference may be desirable for this and like settings.

Detection of the right-to-left movements of the reading position can also be used to determine the vertical position of the reader on the page, since this will typically signify the user beginning a new line of text on the page. Counting the left-to-right movement after turning a page may be used to determine the vertical line currently being read on a particular page. The immersive effects data set may also include the number of lines of text for the pages. (Alternatively, the number of lines can be detected using image processing, for example.) Thus, the current vertical position may be determined as a percentage of lines on the page. This can supplement (or replace) determination of the vertical component of the current reading location using interpolation based on calibration of reference point locations described above. Monitoring eye movement may also be used to learn a user's reading speed. For example, the time taken between successive right-to-left movements (signifying the beginning of new lines) may be used to determine the average reading time per line of text. Once the average time is known, the time elapsed since the last right-to-left movement can then be used to estimate the current reading position horizontally along a line. This may be used, for example, to adjust or supplement (or replace) the horizontal component of the current reading location as otherwise

determined using eye tracking techniques. In addition, the determined reading speed can be used to transition smoothly to immersive effects corresponding to upcoming text, as well as to also fade or end currently generated immersive effects in a manner that corresponds to the reader's reading pace.

In addition, monitoring eye movement for reading patterns through eye tracking may also be used to provide supplemental information to the reader in the case of an e-book. When it is detected that a reader has departed from the current text and moved back to prior text (e.g., when a user re-reads a section), the e-book may mark the text where the user was reading before moving back. This provides the user with a handy reference for determining where he/she left off. Such a marker corresponding to the present position in the text being read can also be used to set-up and calibrate the eye tracking.

The above-described embodiments may be readily adapted to a comic book and other reading material having different types of layouts. As applied to such a comic book, the effects data set may include effects data for illustrations in the panel (or the panel itself), as well as text in the text bubbles. Eye tracking may be used to determine the current viewing position on the page, and use the position to obtain the corresponding immersive effects data for the panel, illustration and/or text being viewed. The corresponding immersive effects may be generated and output analogous to the above descriptions.

In the case of a comic book or like work, a user's pattern of reading can also be used to supplement the determination of the immersive effect, or enhance the output. For example, a reader of a comic book may typically first look at the illustration in a panel, and then read the text bubbles (or text groupings) in a certain order (e.g., left to right, top to bottom). Within each bubble, the text is read as normal. A generic pattern of eye movement is thus a large movement on the page (to a subsequent panel), relatively large movements around a certain sub-region of the page (glancing through the illustrations of a panel), followed by one or more periods of left-to-right eye movement at different locations within the sub-region (corresponding to successive reading of text bubbles). Thus, the user's pattern of viewing may also be monitored using a succession of images and compared with a generic pattern of viewing to determine whether the user is currently reading the material or otherwise "glancing" through the material, reviewing previously read material, or otherwise viewing it in a non-standard sequence. Analogous treatment

may be applied to viewing patterns for magazines (which generally include photos, captions and text for an article). Where the user is not viewing the material according to the standard pattern, the immersive effect output may be suspended, faded, or like adjustment to prevent generation of haphazard effects. Similarly, the system may monitor and learn the user's particular viewing pattern for the comic book or other type of literaray work and generate the corresponding immersive effects when viewed according to that pattern.

In addition, the immersive effect data set for a comic book or other like work may include additional data that allows such patterns to be further exploited. For example, the data set may not only use location to reference immersive effects data, but also identify the layout of the page (e.g., locations of panels, illustrations and text bubbles). Successive location determinations on the page via eye-tracking thus also give a sequence of viewing through the panels, illustrations and text. This can be compared with the generic viewing pattern and used to confirm or supplement the determination of immersive effect based on the eye tracking location. Also, the viewer's particular viewing habits for comics may be learned and used as the underlying pattern for comparison. By effectively including a layout of the locations of panels, illustrations and text in the immersive effects data set, the generic viewing pattern (or the learned pattern) can be used to predict where the user will look next in the sequence. This may be used, for example, to transition between immersive effects.

The systems and components supporting the immersive reading experience may also be readily adapted such that a book (such as an e-book or a paper book) or other work comporting with the invention is automatically recognized and integrates with presentation devices and supporting electronics found in a room (or other space) when a user carrying the book enters the room and begins reading. For example, an electronic book 10 such as described in Fig. 1 may be automatically detected in a room and establish a communication link (e.g., a wireless link) with another computer (which may be a server) associated the room that interfaces with presentation devices located in the room. The room-related computer receives the immersive effect data related to the current material being viewed by the user 60. This is determined by the e-book processor 20 via eye tracking and detecting the location of the current text being read, along with the corresponding immersive effects data set or layer, as previously described. The immersive effects data for the current text is

transmitted by e-book processor to room-related computer via the established communication link.

In this particular example, the room-related computer also receives data that gives the orientation and location of the e-book 10 and/or user 60 in the room and also has stored in memory the locations and types of presentation devices. Room-related computer uses the relative positions of e-book 10 and presentation devices (referenced to the orientation of e-book), the types of presentation devices and the immersive effect data requirements received from e-book to select and control presentation devices to generate the appropriate immersive effects for user 60. The output provides the immersive experience properly oriented and located for the text currently being read. Analogous treatment of a paper book such as described with respect to Fig. 2, as well as other works, may also be readily adapted.

The use of presentation devices resident to a room (or other space) as described above and integration of an e-book entering the room is adapted from description found in commonly owned co-filed U.S. Provisional Patent Application entitled "Orientation and Position Adaptation For Immersive Experiences" by Cortenraad and Bergman, Attorney Docket US05110, filed concurrently herewith, assigned U.S. Provisional Patent Application Ser. No. 60/***,*** the entire contents of which are hereby incorporated by reference herein. Other configurations are readily adapted: The e-book, book or other literary work of the present invention may be regarded as a portable electronic device as described throughout the above-cited application of Cortenraad and Bergman, with configurations as described therein.

Automatic detection and locating of the e-book, paper book, and/or user in a room or other space may use an available image recognition system and/or technique. Systems and techniques of image recognition are available that are adaptable to recognize the contours of a book or other literary work. Location may be determined, for example, by using multiple cameras to capture images in the room and applying standard stereo techniques of two or three dimensional computer vision. By applying these techniques, the position of a book and/or person in a room or other space in two or three dimensions may be readily calculated from its positions in the captured images. A particular system that is adaptable to use available image recognition techniques to recognize the contours of a book or other literary work (or the user) and describes processing which may be used with

the images to determine the recognized work's coordinate position in two or three dimensions is described in PCT Published International Application having International Publication No. WO 02/41664 A2, entitled "Automatically Adjusting Audio System" by M. Trajkovic et al., having International Publication Date 23 May 2002, the entire contents of which are hereby incorporated by reference herein. Alternatively, the e-book may be detected via a radio beacon, and the location may be monitored using internal accelerometers.

As noted above, user preferences may be utilized for aspects of the immersive experience. Among other things that may be subjected to user preferences is the level of immersion desired. For example, the user can choose to activate only background sounds, but not lighting. Such preferences may be manually input and may also be learned by the system for particular users over time. It is also noted that, like in films, music as an immersive effect may be used to enhance the reading experience. For example, music may be added to create an emotion in anticipation of an upcoming event about to be read by the user. For example, foreboding music may begin just prior to the beginning of a suspenseful passage, thus giving the user a tense immersive experience. The music may build and climax as the scene being read unfolds.

Fig. 3 is a flowchart of a method of providing an immersive reading experience. Referring to Fig. 3, the reader's eyes are tracked in block 200. In block 210, the location of the portion of text currently being read by the reader is determined from the tracking of block 200. In block 220 output is produced for at least one immersive effect corresponding to the portion of text currently being read by the reader.

While the invention has been described with reference to several embodiments, it will be understood by those skilled in the art that the invention is not limited to the specific forms shown and described. Thus, various changes in form and details may be made therein without departing from the spirit and scope of the invention as defined by the appended claims. For example, as noted there are many alternative configurations and locations for the various components that support the invention. For example, processing performed in support of the invention (such as processing performed by processor 20 and 20a) may be carried out by multiple processors in multiple locations. Interfaces shown for the various embodiments between various components can have multiple paths and types. Where a computer is referred to, it is utilized broadly and may include, for example,

servers, PCs, microcomputers and microcontrollers. It may also generally include a processor, microprocessor and CPU with its associated memory. (Any other computing device that is possibly excluded from the broad scope of "computer" may nonetheless be substituted if it may be configured to carry out the required processing.) Although the specific presentation devices shown for the examples of Fig. 1d are speakers and lights, any other type of presentation device may serve the room or space. Any other light source in addition to an IR LED that provides the requisite reflection for eye tracking may be used, as may other techniques of eye tracking. Depending on the selected calibration points used for eye tracking, other mathematical processes, such as extrapolation, may also be used to determine a current viewing location. The effects (and when applicable, the electronic text for a book) may be garnered in other ways than described above, for example, by accessing a commercial or other remote server in real-time and downloading the effects data and/or text data. The effects may also be obtained from a local server, DVD, memory card, or like source. Thus, the particular techniques described above are by way of example only and not to limit the scope of the invention.

What Is Claimed Is:

1) A method of providing an immersive reading experience, the method comprising:

a) tracking a reader's (60) eyes (200);

b) determining the location (X) in material (10a, 14) currently being viewed by the reader (60) from the tracking of the reader's eyes (210); and

c) utilizing the location (X) in the material (10a, 14) to produce output for at least one immersive effect corresponding to the portion of material (10a, 14) currently being viewed by the reader (60) (220).

2) The method as in Claim 1, wherein the material (10a, 14) comprises text (14, 14a) that the reader (60) is reading.

3) The method of Claim 1, wherein the tracking a reader's (60) eyes uses images of the reader (60).

4) The method of Claim 3, wherein the location (X) in the material (10a, 14) currently being viewed by the reader (60) is determined by generating a reflection in the user's eye that is detected in the images.

5) The method of Claim 1, wherein the determined location (X) in the material (10a, 14) is used to identify immersive effects data corresponding to the material (10a, 14) currently being read, the identified immersive effect data used to generate instructions for the at least one immersive effect corresponding to the location (X) of material (10a, 14) currently being read by the reader (60).

6) The method of Claim 5, wherein the identified immersive effects data is identified from a set of immersive effects data for the material (10a, 14), wherein the immersive effects data in the set is correlated to locations throughout the material (10a, 14).

7) The method of Claim 5, wherein the immersive effects data for the material (14) currently being read is obtained from a sublayer of electronic text (14) of the material (14).

8) The method of Claim 1, wherein the immersive effect evokes a sensory perception of the reader (60), the sensory perception comprising at least one of hearing, visual, olfactory, tactile and taste.

9) A system (5, 5a, 5b) that generates an immersive reading experience for literary material (10a, 14) currently being viewed by a user (60), the system (5, 5a, 5b) comprising:

a) at least one camera (18, 18a) that captures images of the user (60) viewing the material (10a, 14); and

b) at least one processor (20, 20a) operatively coupled to the camera (18, 18a) that receives the images from the camera (18, 18a) , the processor (20, 20a) configured to process the images and use the images to determine the location (X) in the material (10a, 14) currently being viewed by the user (60) and further configured to process the location (X) in the material (10a, 14) to generate at least one immersive effect related output corresponding to the current material being viewed by the user (60).

10) The system (5, 5a, 5b) as in Claim 9, wherein the material (10a, 14) being viewed comprises text (14, 14a) that is read by the user (60).

11) The system (5, 5a, 5b) as in Claim 9, wherein the determination of location (X) in the material (10a, 14) currently being viewed by the user using the images by the processor (20, 20a) comprises detecting a location (x) of a reflection generated in the eyes of the user (60) that is correlated to the location (X) in the material (10a, 14).

12) The system (5, 5a, 5b) as in Claim 9, wherein the processor (20, 20a) is further configured to access a set of immersive effects data, the set of immersive effects

data comprising immersive effects data correlated to locations throughout the material (10a, 14), the processor (20, 20a) using the determined location (X) in the material currently being read by the user (60) to identify the immersive effect data in the data set corresponding to the material (10a, 14) currently being read.

13) The system (5, 5a, 5b) as in Claim 12, wherein the identified immersive effects data is utilized to generate the at least one immersive effect related output corresponding to the material (10a, 14) currently being read.

14) The system (5, 5a) as in Claim 9, wherein the determined location (X) is used to identify corresponding immersive effects data from a sublayer of electronic text (14) of the literary material (14).

15) The system (5, 5a) as in Claim 14, wherein the sublayer comprises a description of the currently viewed material (14) that is translated by processor (20) to generate the corresponding at least one immersive effect related output.

16) The system (5, 5a, 5b) as in Claim 9, wherein generation of the at least one immersive effect related output by the processor (20, 20a) is also based on the type and location of one or more presentation devices (30, 40) available to the system (5, 5a, 5b) for presenting immersive effects to the user (60).

17) The system (5, 5a, 5b) as in Claim 16, wherein the type of at least one presentation device (30, 40) available to the system (5, 5a, 5b) provides at least one of an audio sensation, visual sensation, tactile sensation, olfactory sensation and taste sensation.

18) The system (5, 5a, 5b) as in Claim 16, wherein the at least one processor (20, 20a) further determines the relative positions of the user (60) with respect to the presentation devices (30, 40), referenced to the current orientation of the literary material (10a, 14), and uses the relative positions in generation of the at least one immersive effect related output.

19) The system (5, 5a) as in Claim 9, wherein the literary material (14) viewed by the user (60) is presented via an electronic book (10), the processor (20) further configured to output to the user (60) a marker in a display (12) of the electronic book (10) corresponding to the location (X) currently being viewed by the user (60).

20) The system (5, 5a, 5b) as in Claim 9, wherein the images received are processed by processor (20, 20a) to determine the user's (60) current sequence of viewing the material (10a, 14), the generating of at least one immersive effect related output being modified when the user's (60) current viewing sequence differs from a pattern of viewing.

21) The system (5, 5a, 5b) as in Claim 9, wherein the images received are processed by processor (20, 20a) to track the user's (60) eye movement, the movement being used to determine the location (X) of the material (10a, 14) currently being viewed by the user (60).

22) A processor (20, 20a) associated with generation of an immersive reading experience for a literary work (10a, 14) being read by a user (60), the processor (20, 20a) configured to receive images of the user (60) reading the work (10a, 14) and further configured to process the images of the user (60) to determine the location (X) of the text (14, 14a) currently being read by the user (60), the processor (20, 20a) further configured to use the determined location (X) to generate an immersive effect related output corresponding to the text (14, 14a) currently being read by the user (60).

23) The processor (20, 20a) as in Claim 22, wherein the processor (20, 20a) uses the determined location (X) to identify immersive effect data in an immersive effect data set, the immersive effect data identified corresponding to the text (14, 14a) currently being read.

24) The processor (20, 20a) as in Claim 22, wherein the processor (20, 20a) processes images of the user (60) using eye tracking to determine the location (X) of the text (14, 14a) currently being read by the user (60).

25) The processor (20, 20a) as in Claim 22, wherein the immersive effect related output generated is further a function of a type and location of at least one presentation device (30, 40) known by processor (20, 20a).

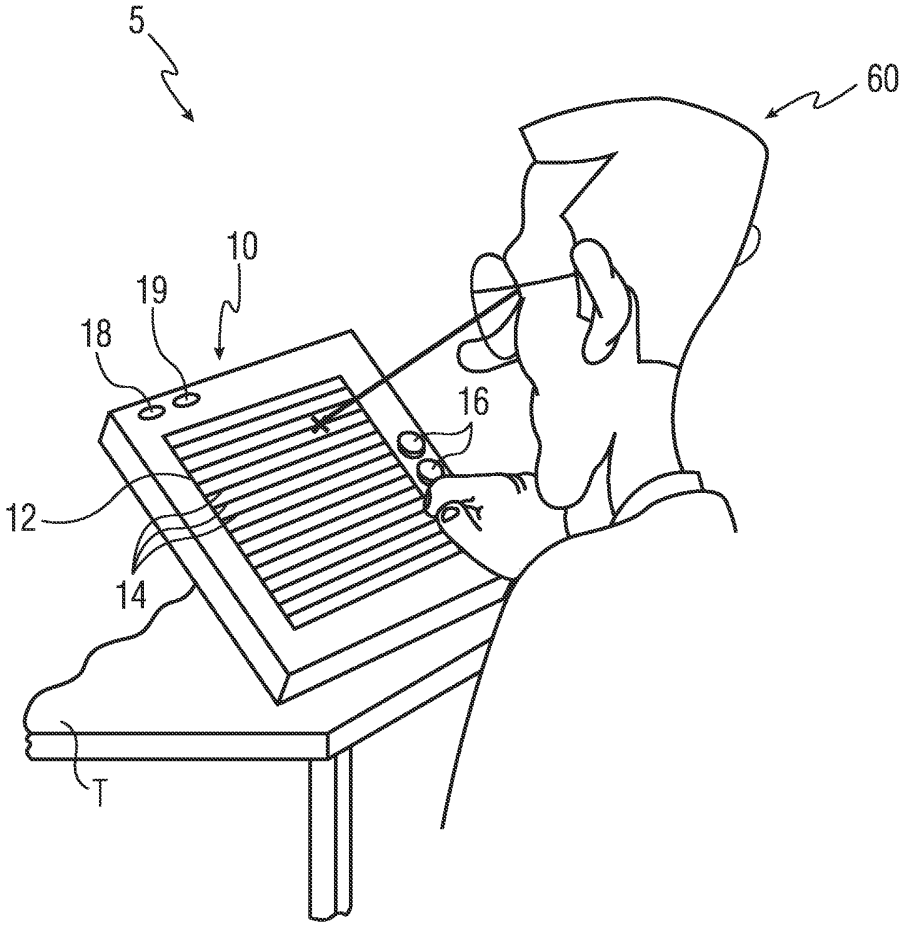


FIG. 1

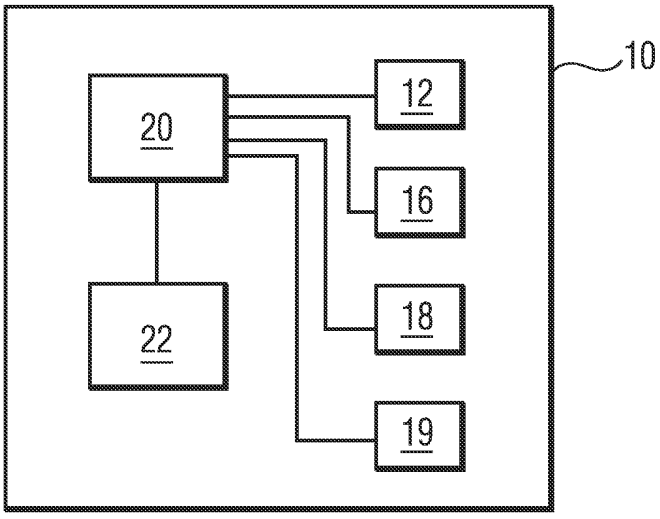


FIG. 1A

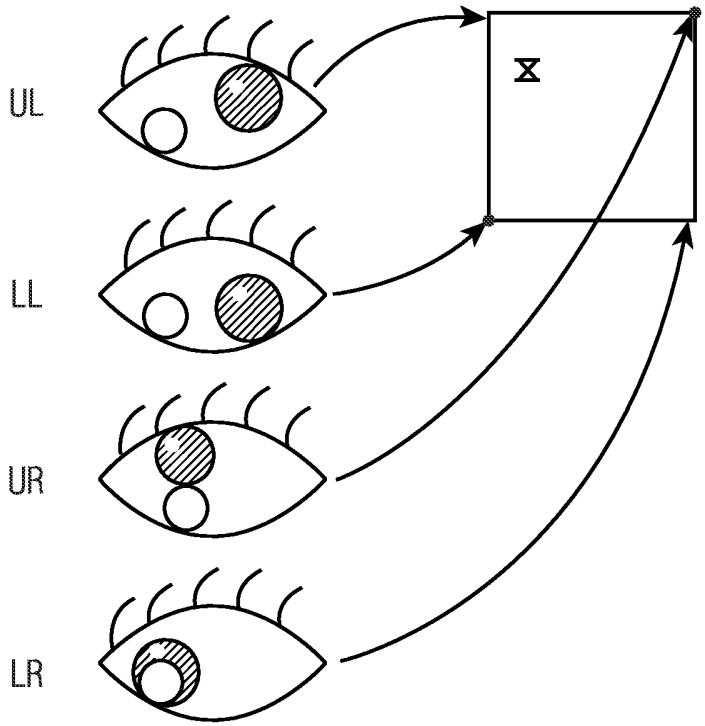


FIG. 1B

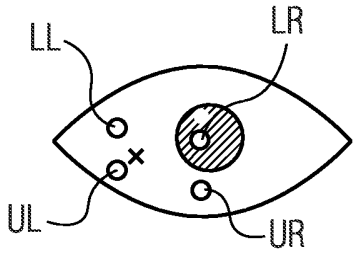


FIG. 1C

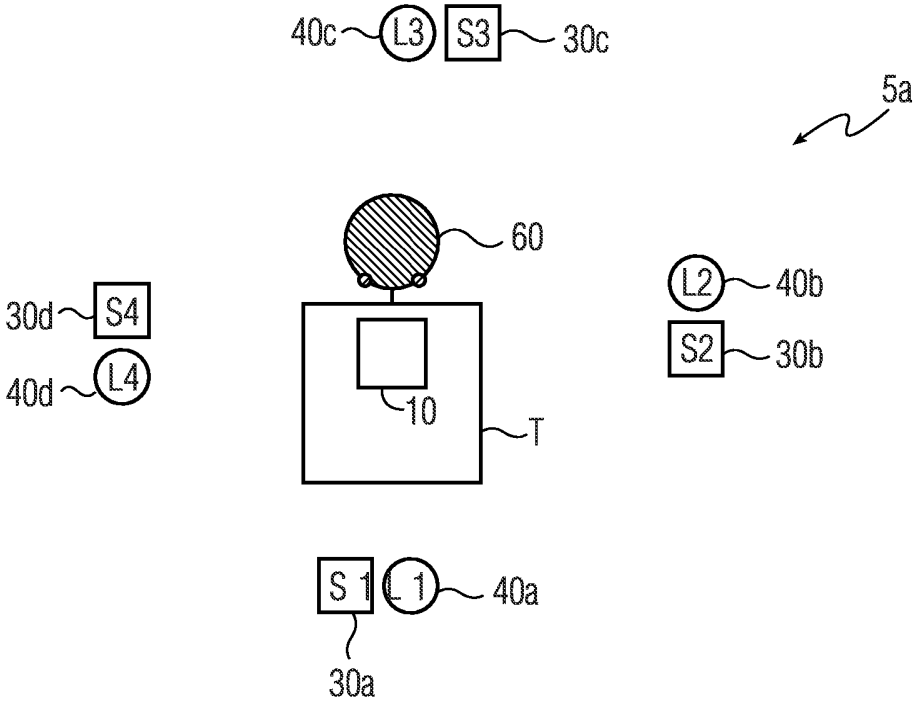

40c
L3
S3
30c
5a
60
30d
S4
40d
L4
L2
40b
S2
30b
10
T
S 1
L 1
40a
30a

FIG. 1D

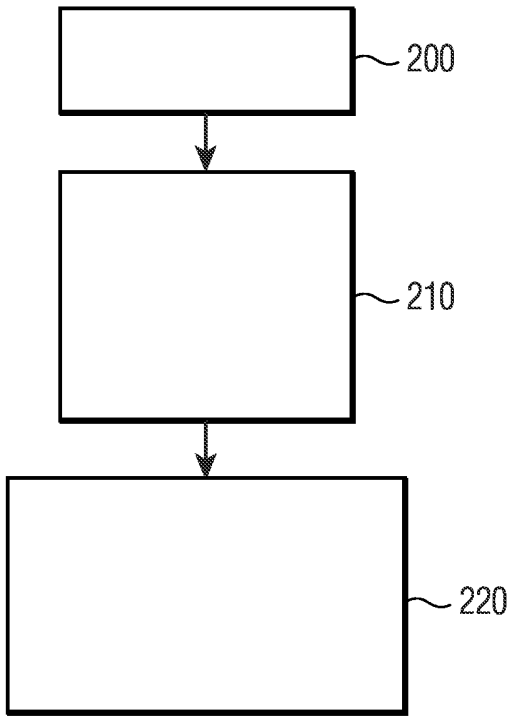

200
210
220

FIG. 3

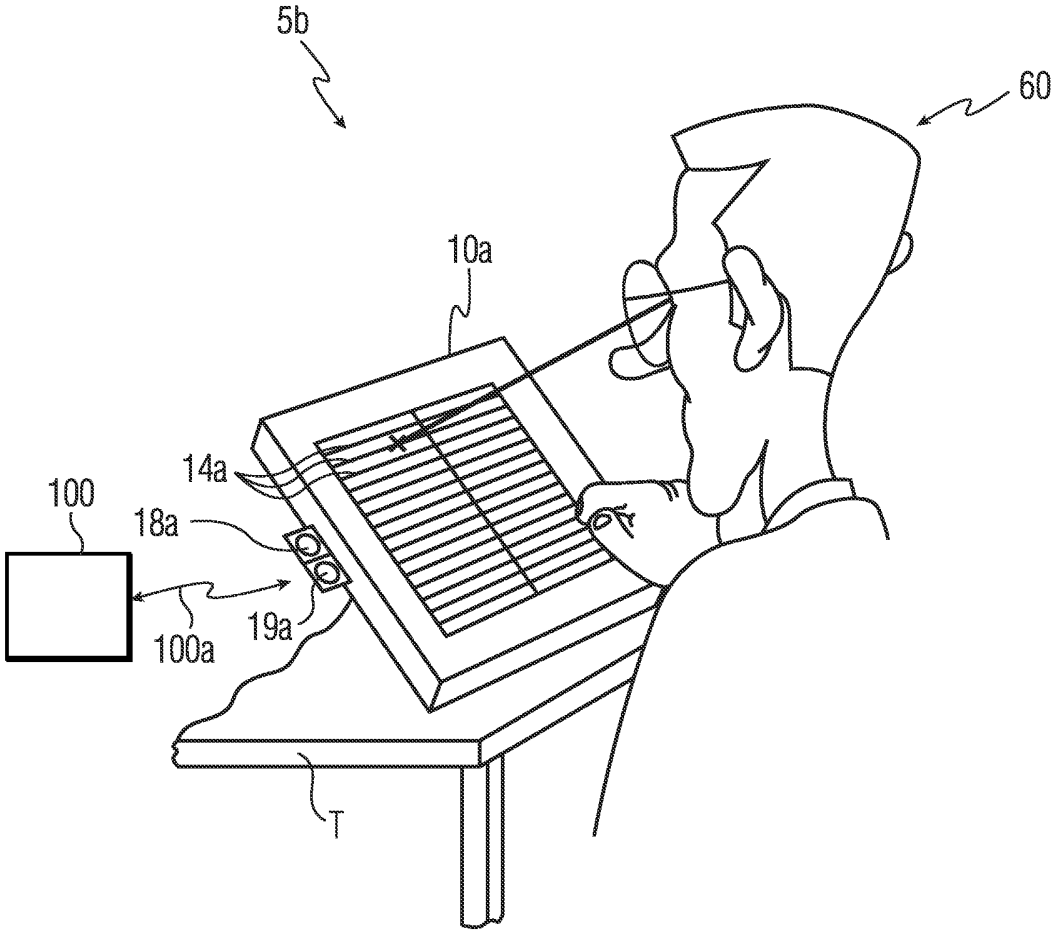


FIG. 2

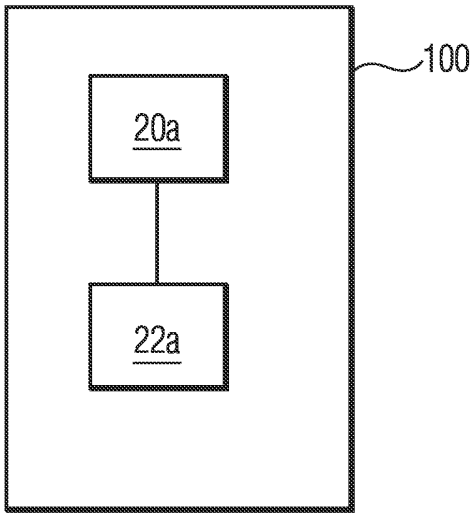


FIG. 2A