



(19) **United States**

(12) **Patent Application Publication**

Lapstun et al.

(10) **Pub. No.: US 2006/0028400 A1**

(43) **Pub. Date:**

Feb. 9, 2006

(54) **HEAD MOUNTED DISPLAY WITH WAVE FRONT MODULATOR**

Publication Classification

(75) Inventors: **Paul Lapstun**, Balmain (AU); **Kia Silverbrook**, Balmain (AU)

(51) **Int. Cl.**
G09G 5/00 (2006.01)
(52) **U.S. Cl.** **345/8**

Correspondence Address:
SILVERBROOK RESEARCH PTY LTD
393 DARLING STREET
BALMAIN NSW 2041 (AU)

(57) **ABSTRACT**

An augmented reality device for inserting virtual imagery into a user's view of their physical environment, the device comprising: a display device through which the user can view the physical environment; an optical sensing device for sensing at least one surface in the physical environment; and, a controller for projecting the virtual imagery via the display device; wherein during use, the controller uses wave front modulation to match the curvature of the wave fronts of light reflected from the display device to the user's eyes with the curvature of the wave fronts of light that would be transmitted through the device display if the virtual imagery were situated at a predetermined position relative to the surface, such that the user sees the virtual imagery at the predetermined position regardless of changes in position of the user's eyes with respect to the see-through display.

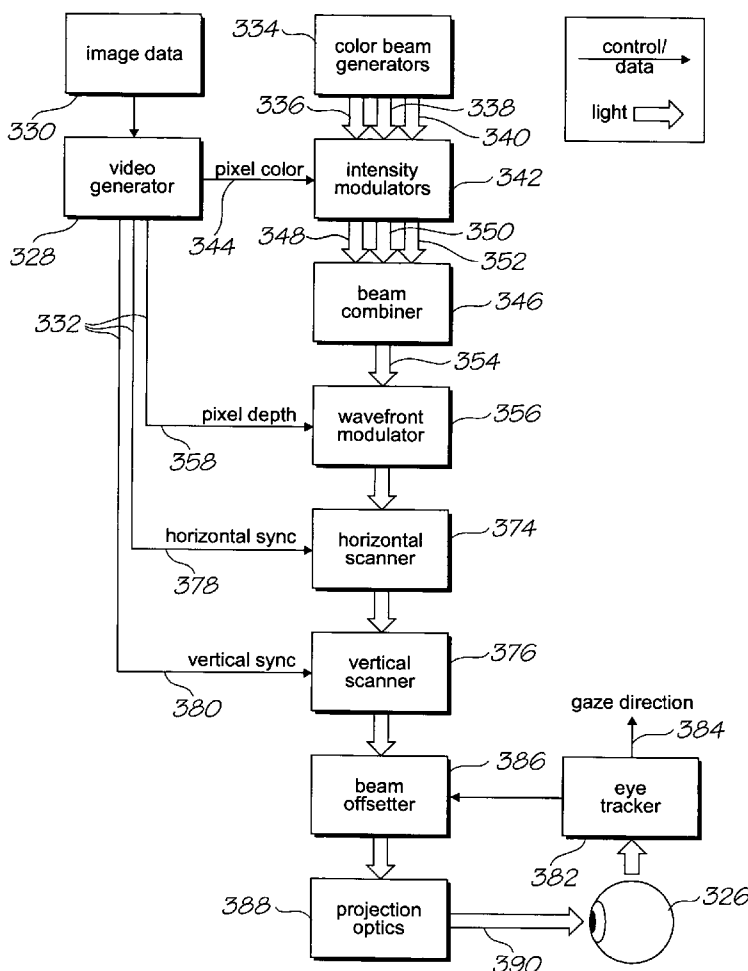
(73) Assignee: **Silverbrook Research Pty Ltd**

(21) Appl. No.: **11/193,481**

(22) Filed: **Aug. 1, 2005**

(30) **Foreign Application Priority Data**

Aug. 3, 2004	(AU)	2004904324
Aug. 3, 2004	(AU)	2004904325
Aug. 20, 2004	(AU)	2004904740
Aug. 24, 2004	(AU)	2004904803
Sep. 21, 2004	(AU)	2004905413
Jan. 5, 2005	(AU)	2005900034



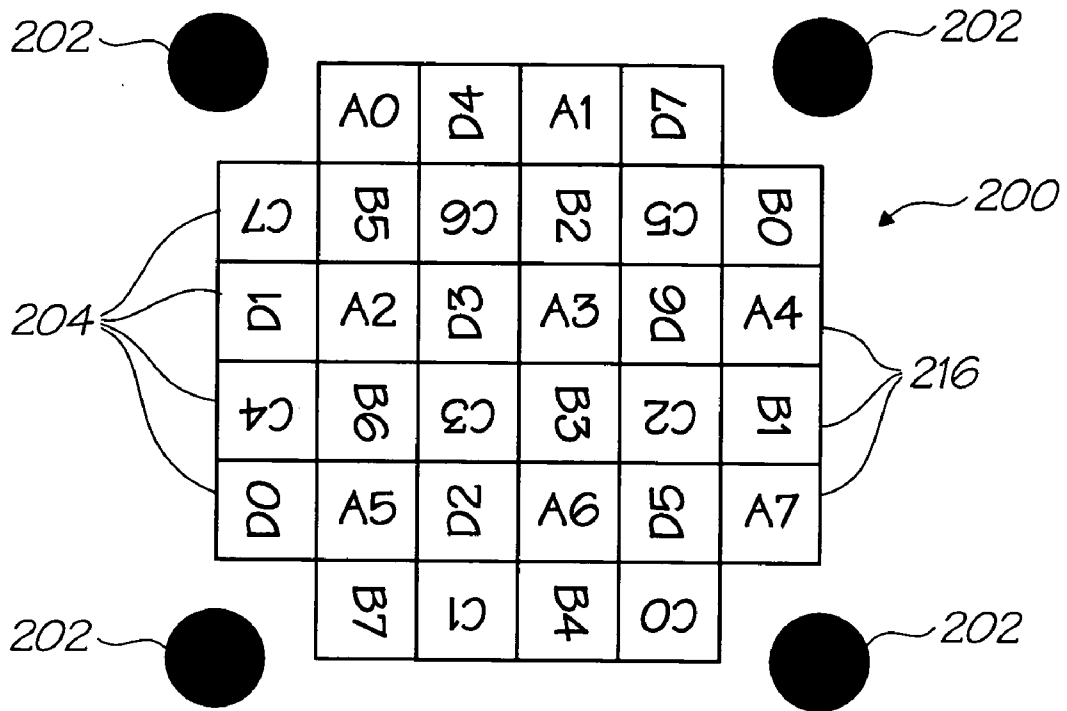


FIG. 1

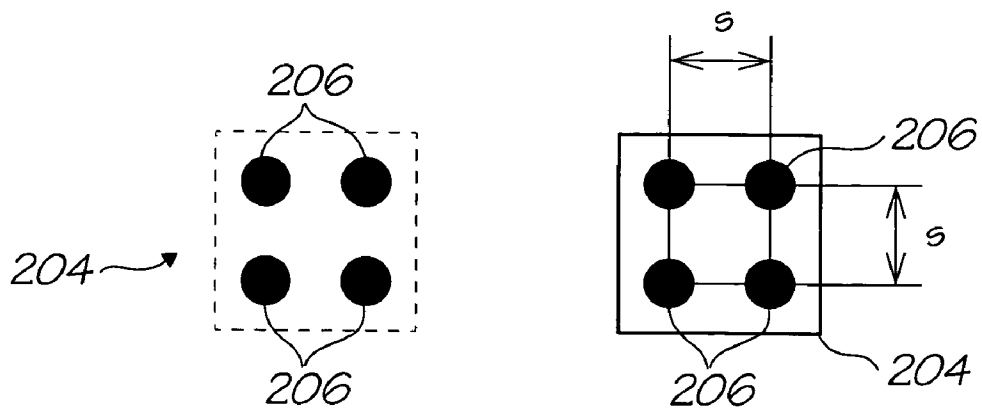


FIG. 2

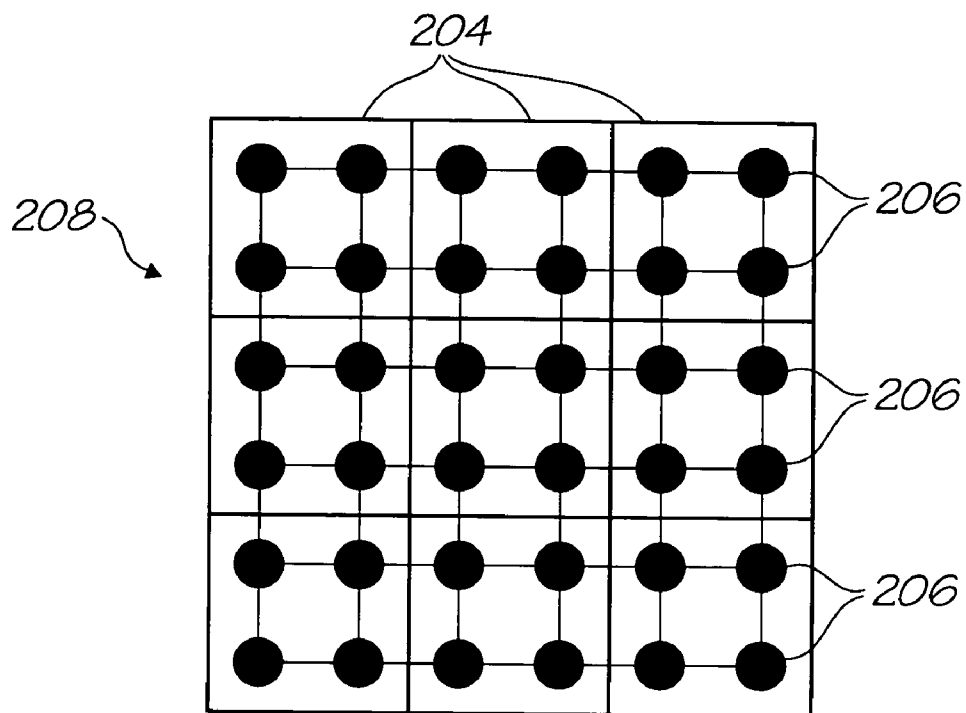


FIG. 3

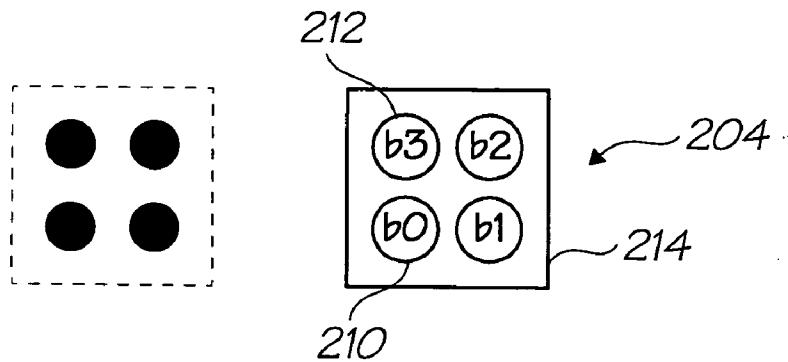


FIG. 4

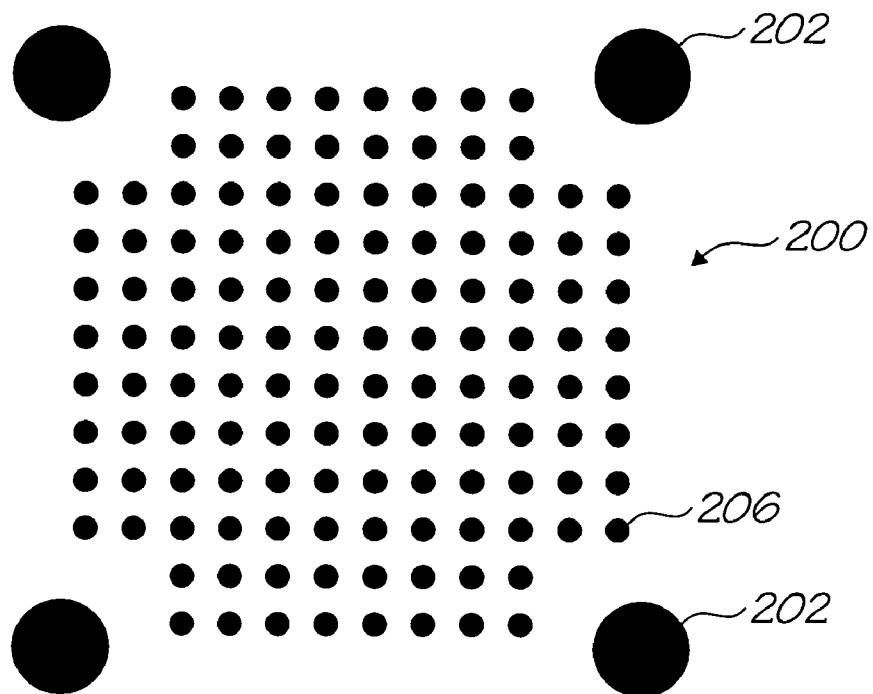


FIG. 5

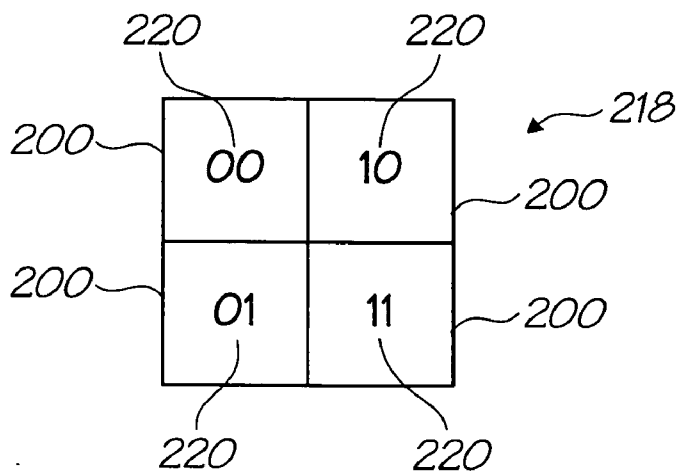


FIG. 6

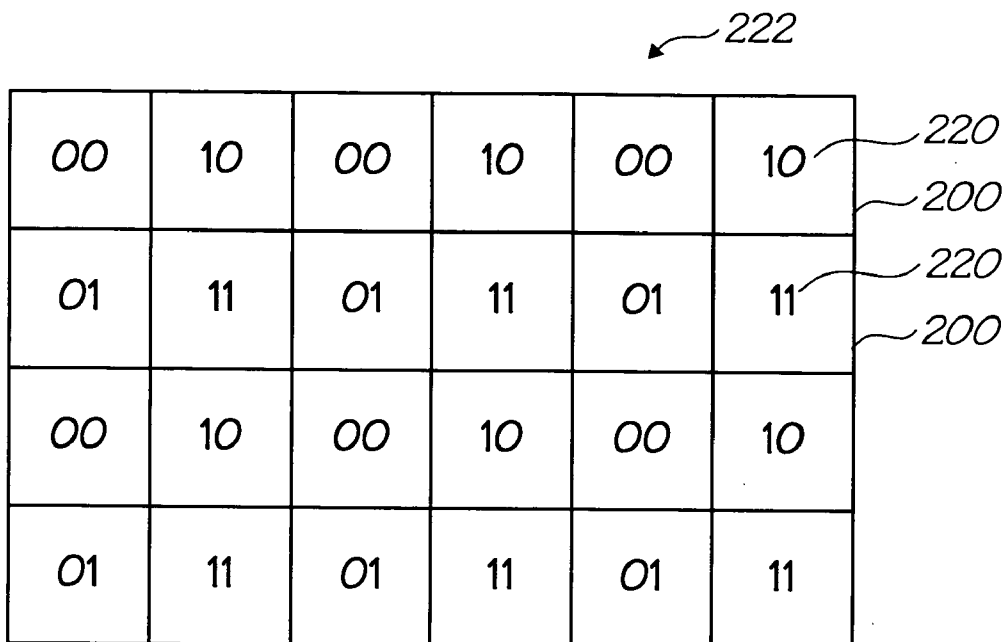


FIG. 7

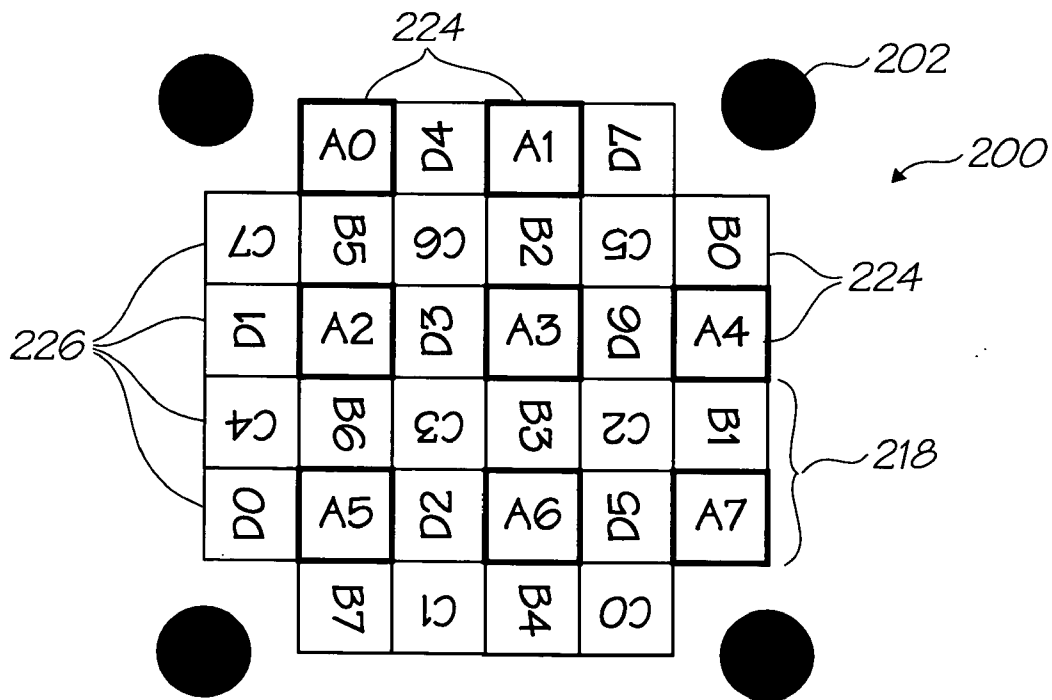


FIG. 8

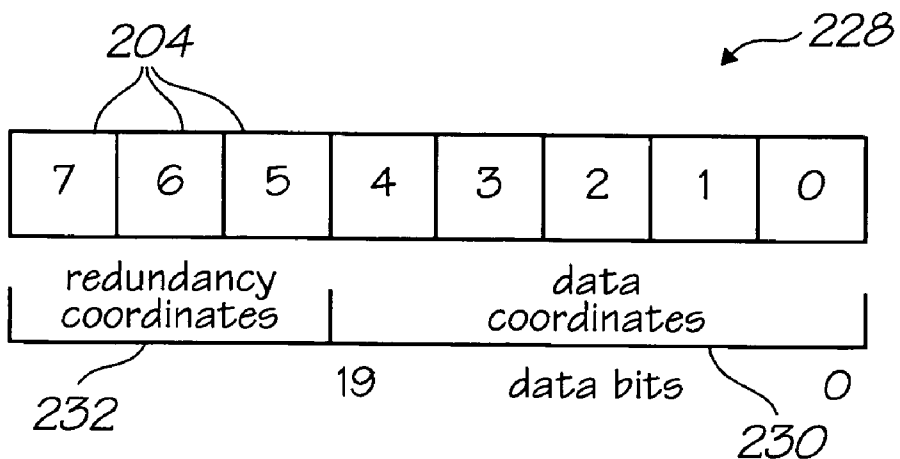


FIG. 9

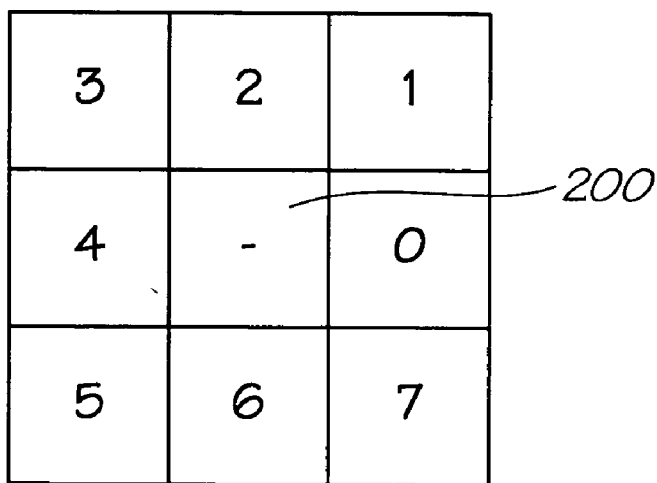


FIG. 10

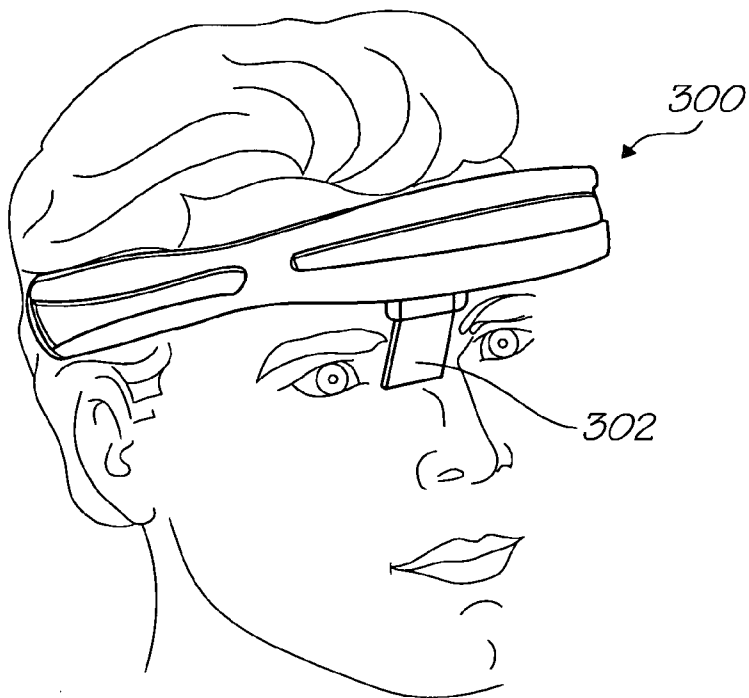


FIG. 11

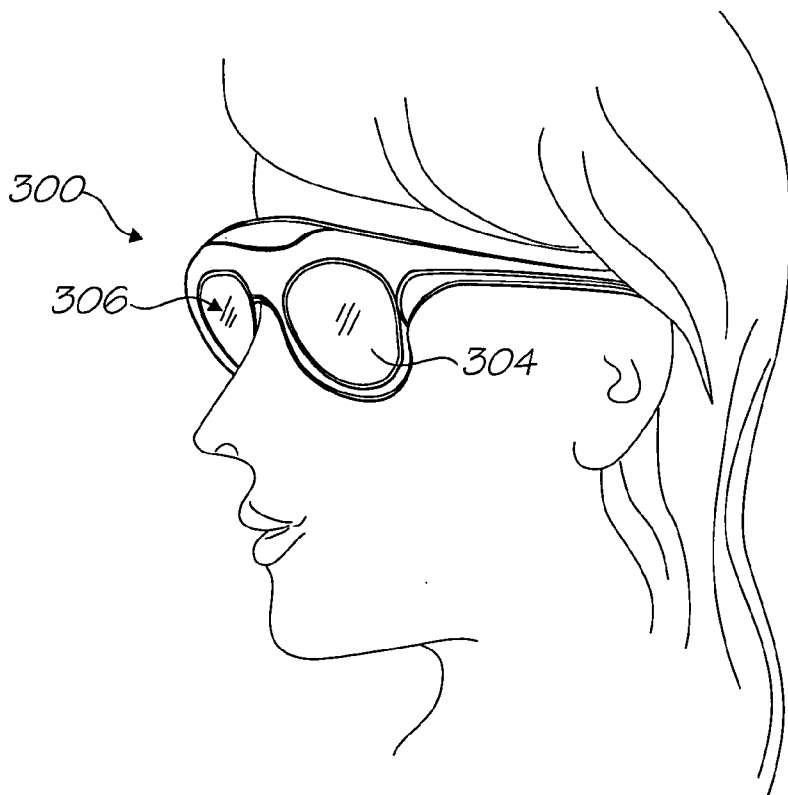


FIG. 12

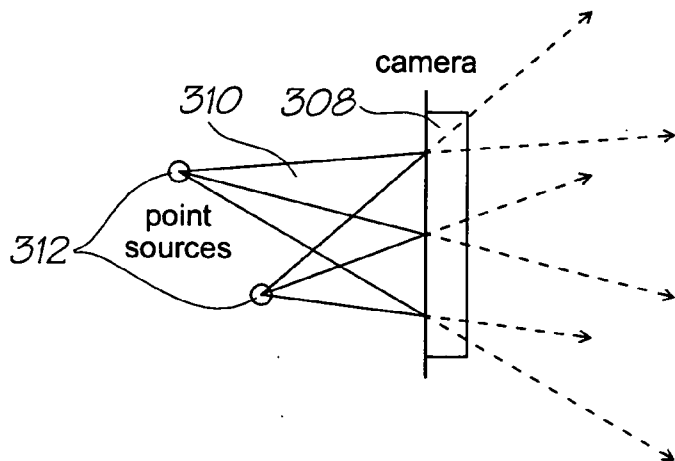


FIG. 13

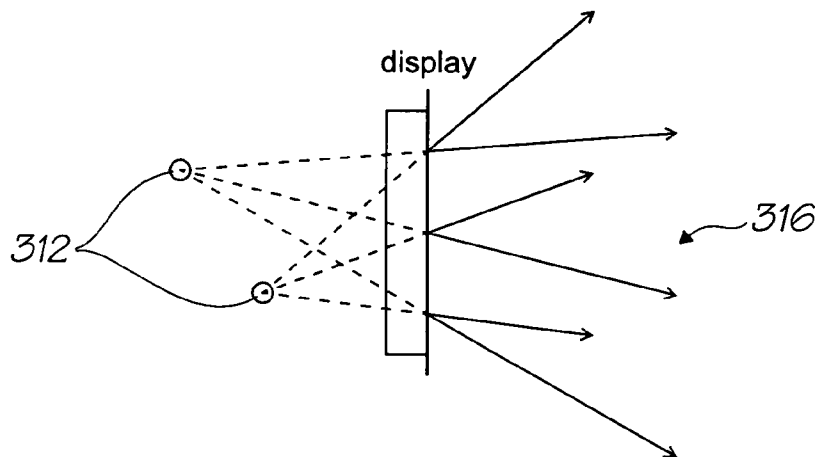


FIG. 14

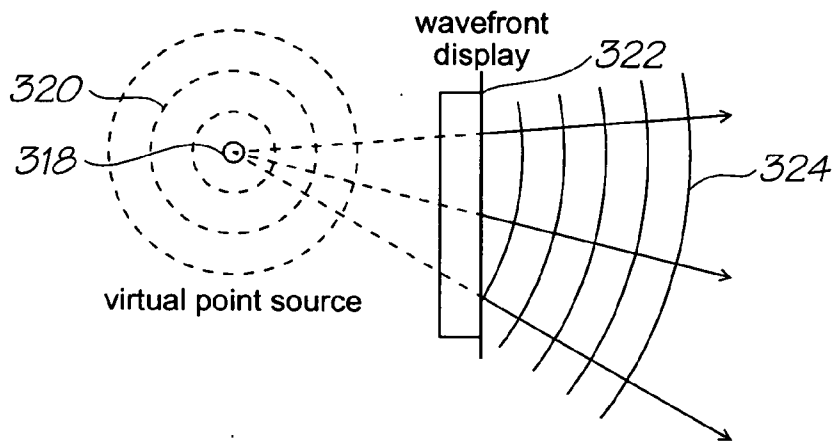


FIG. 15

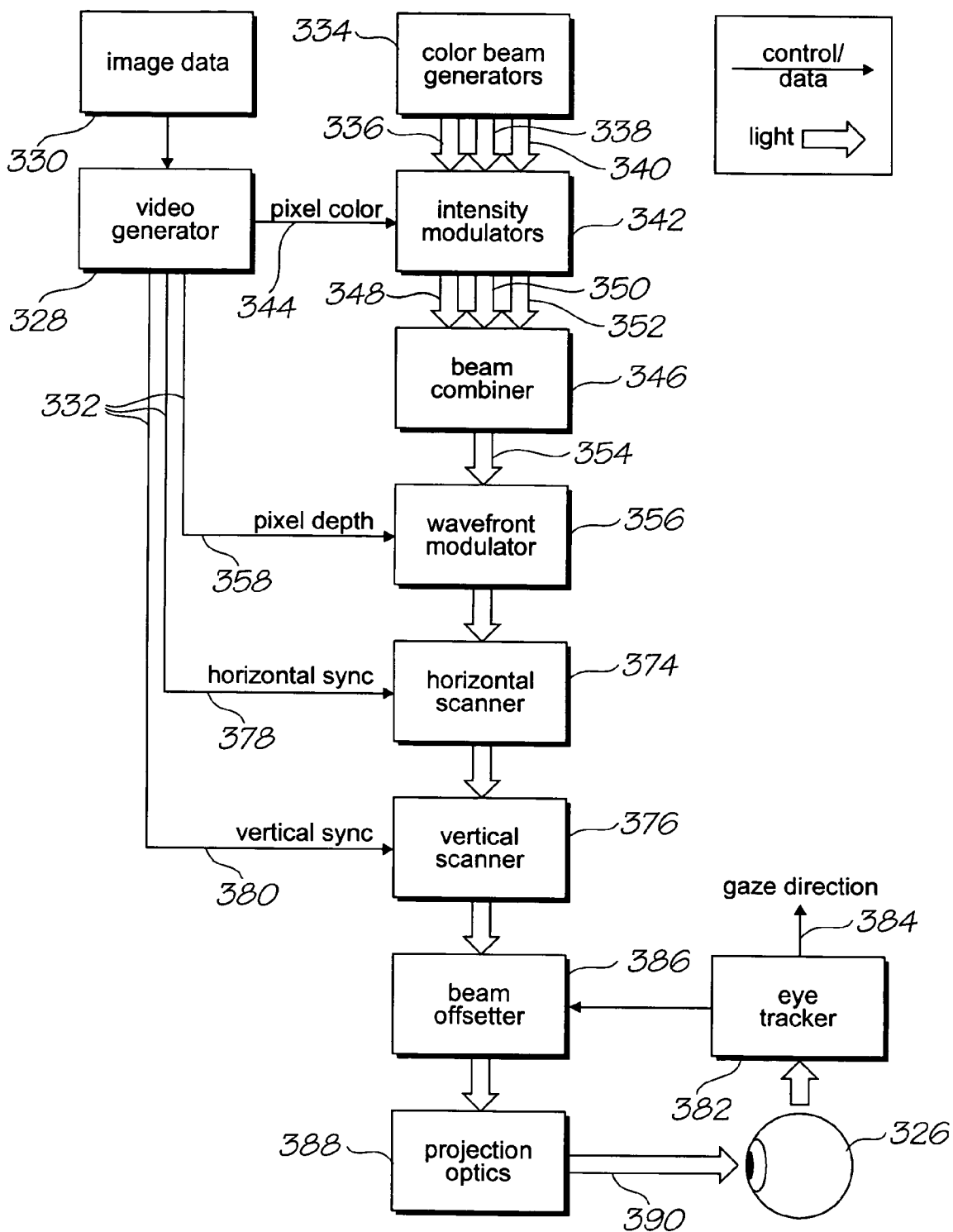


FIG. 16

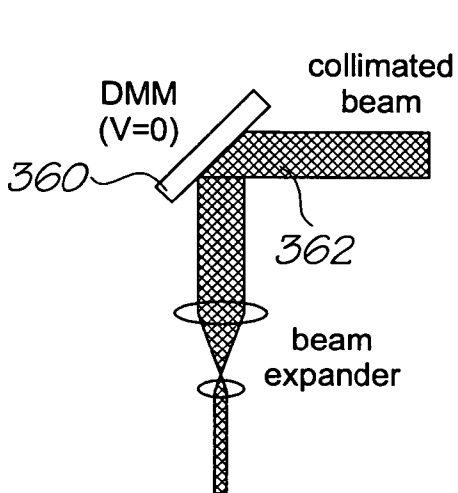


FIG. 17A

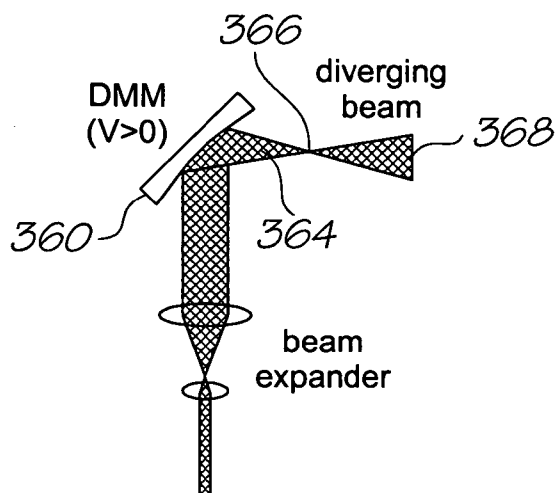


FIG. 17B

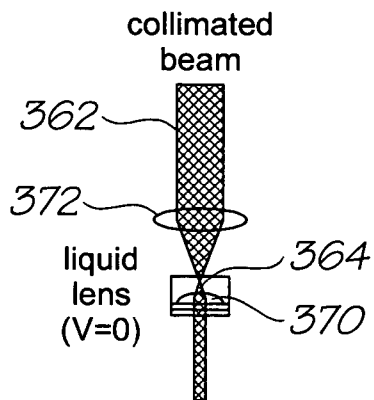


FIG. 18A

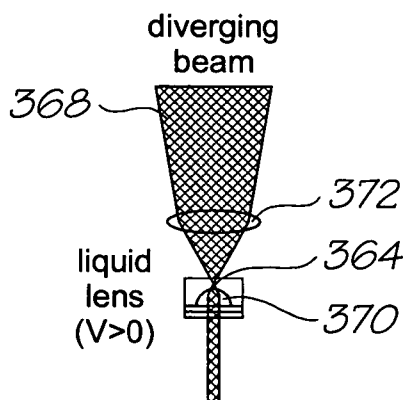


FIG. 18B

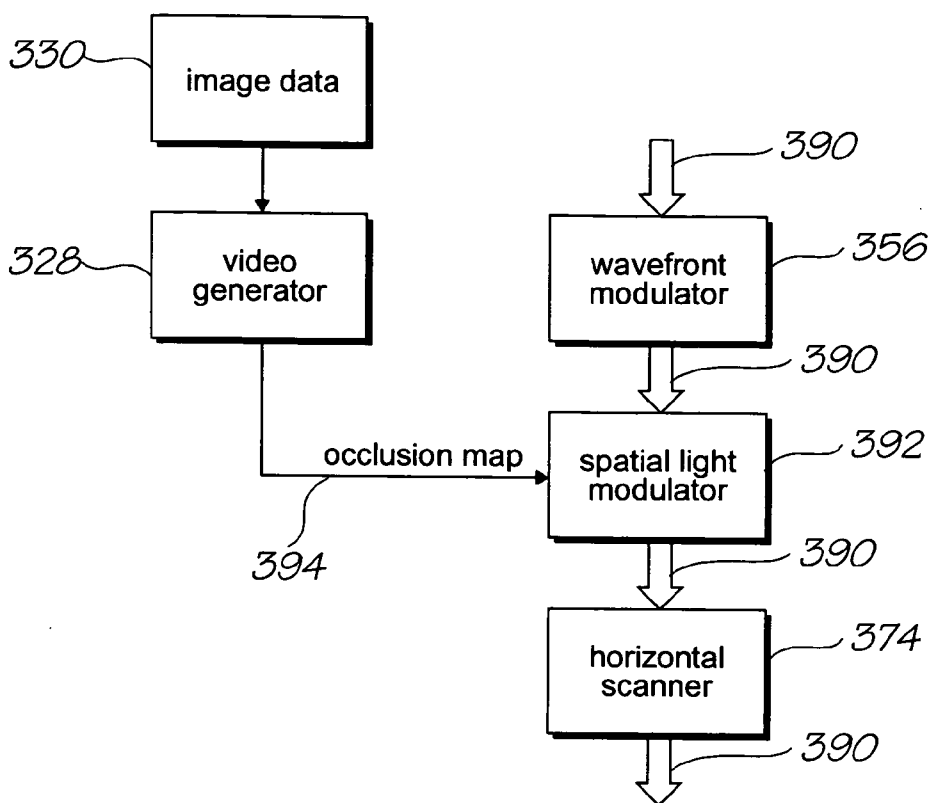


FIG. 19

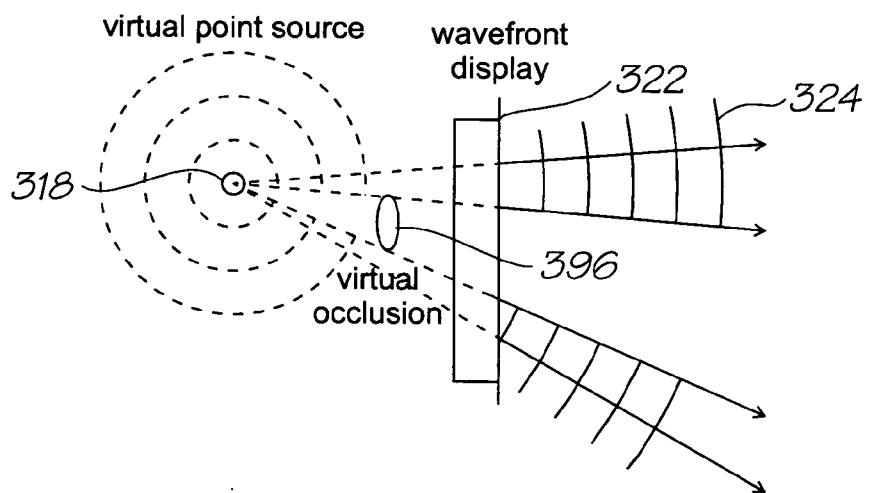


FIG. 20

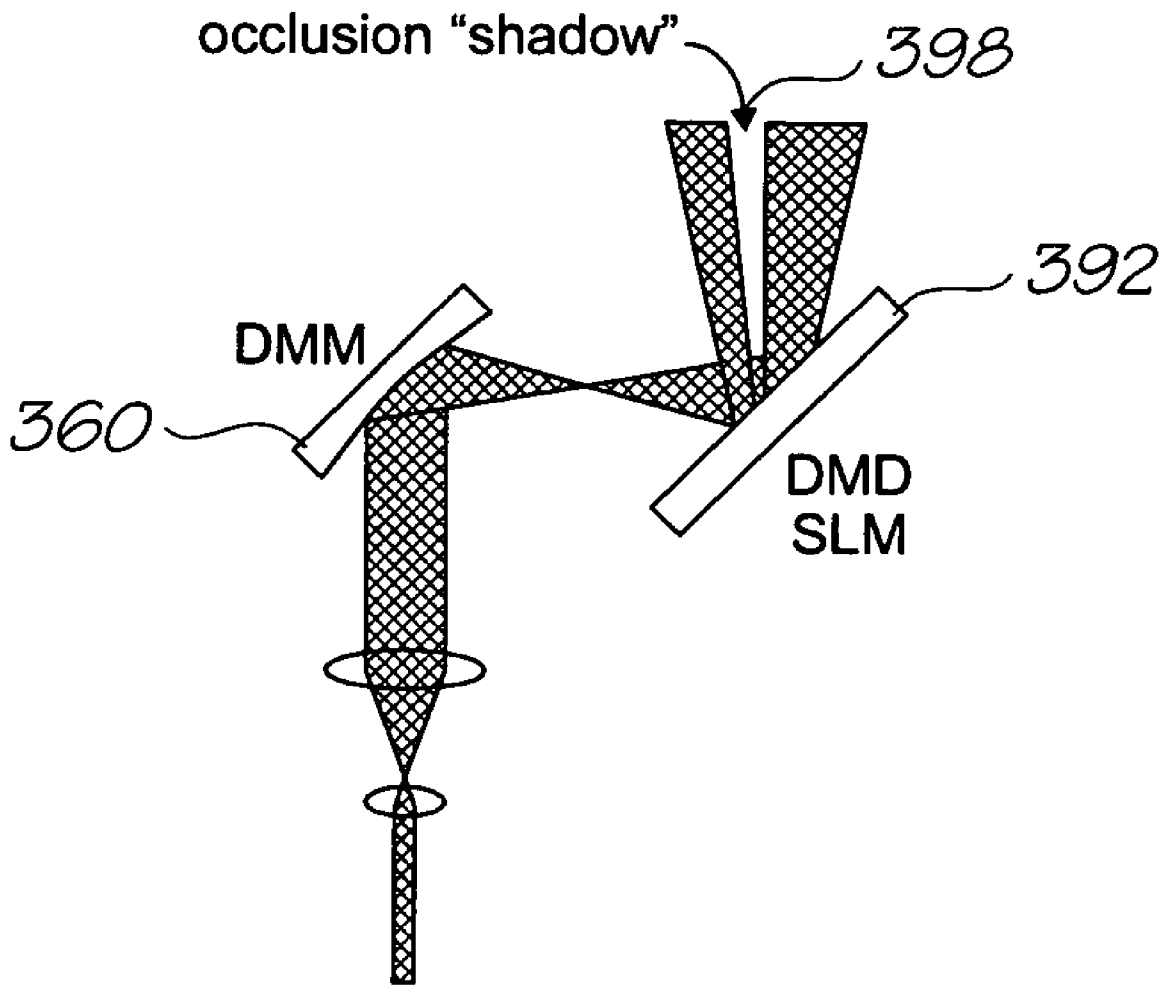


FIG. 21

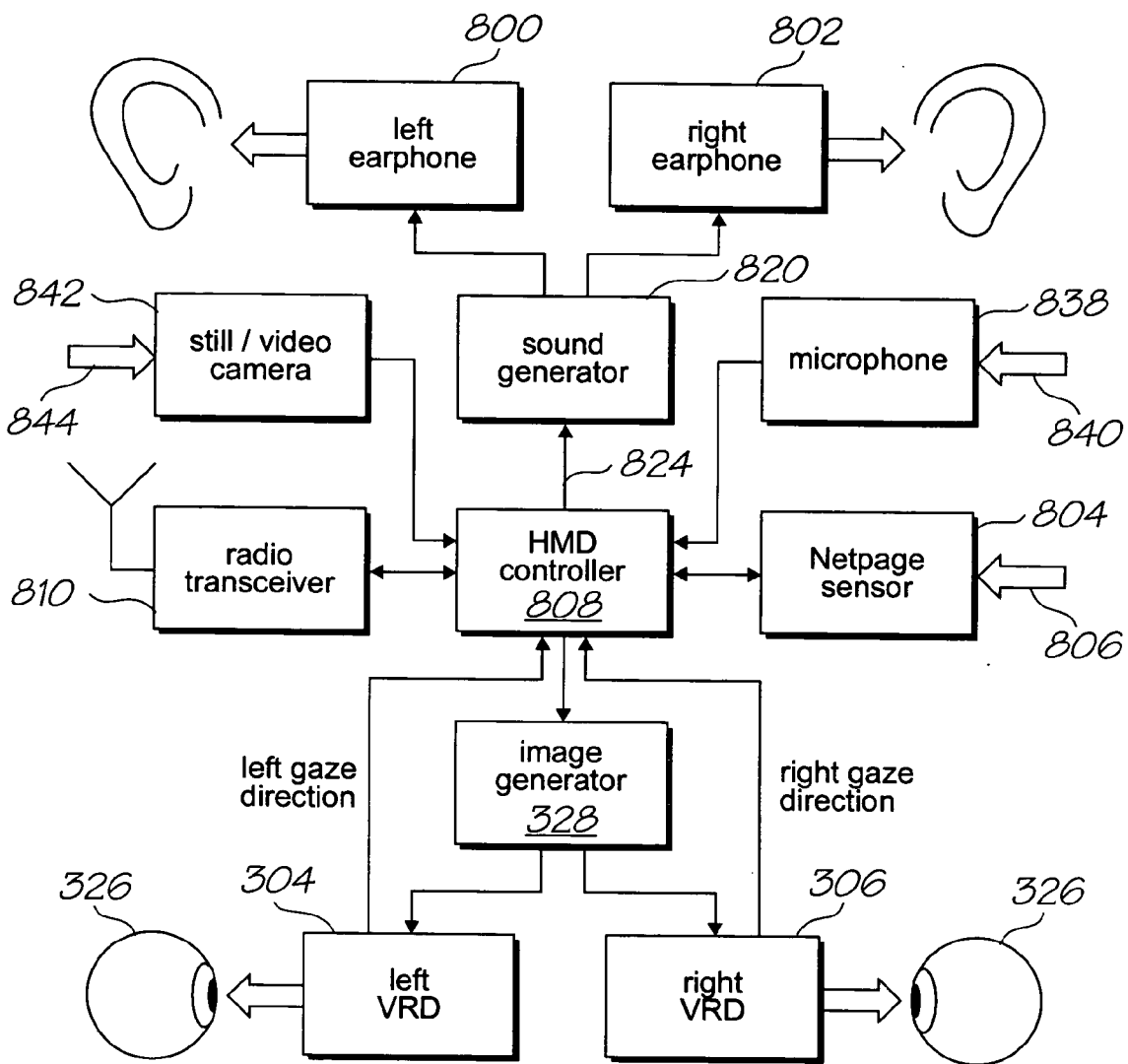


FIG. 22

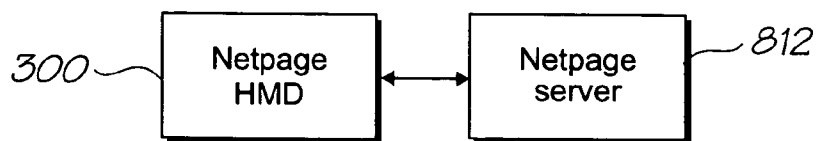


FIG. 23

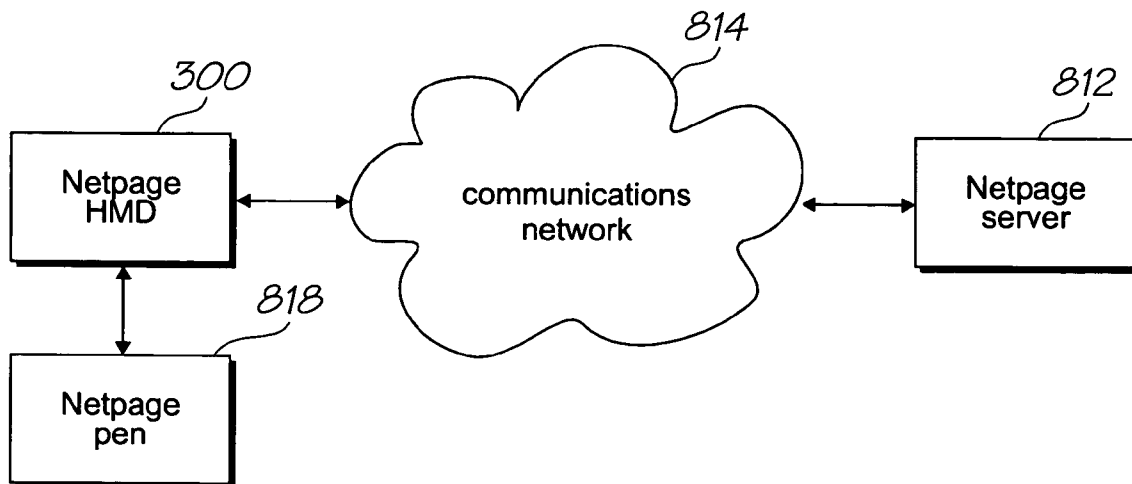


FIG. 24

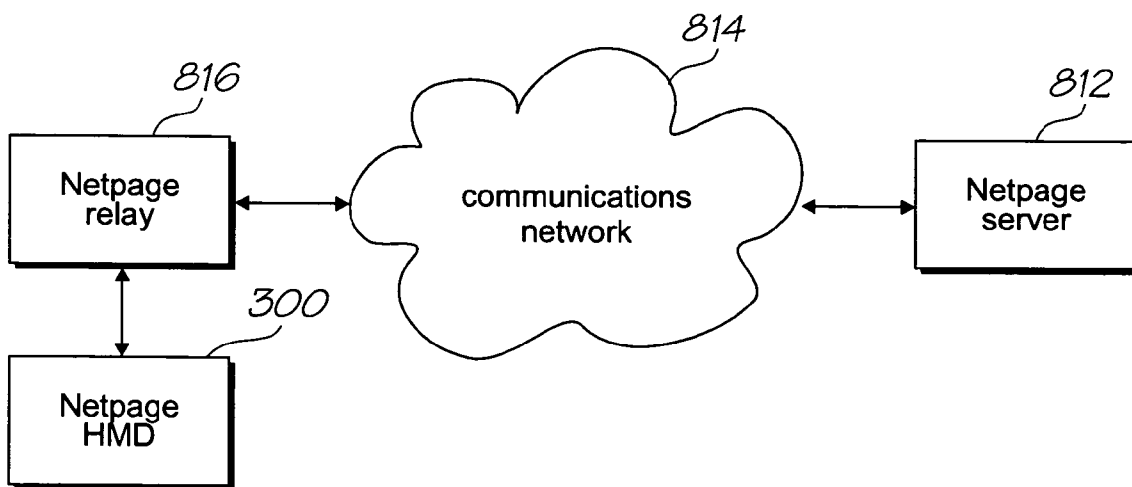


FIG. 25

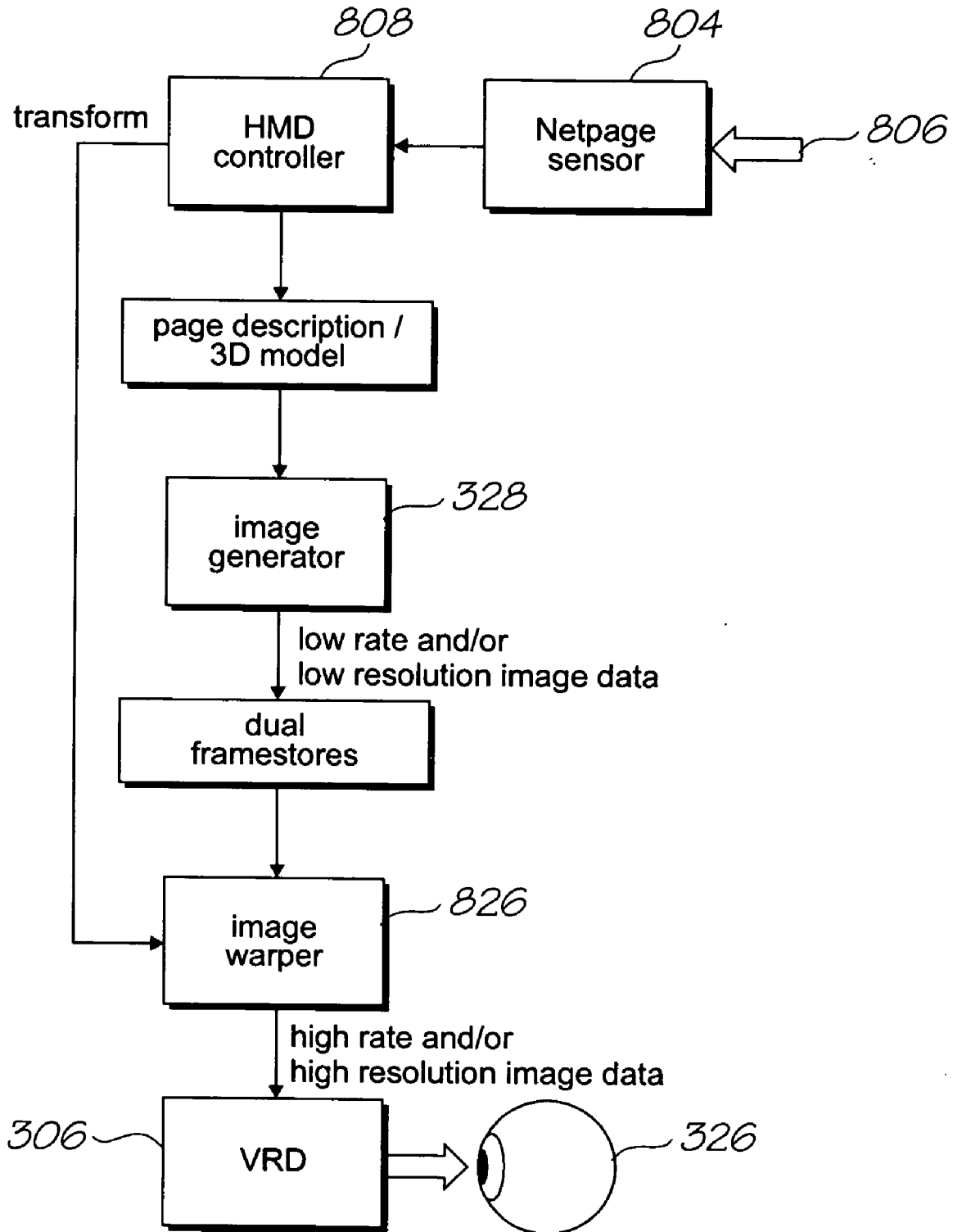


FIG. 26

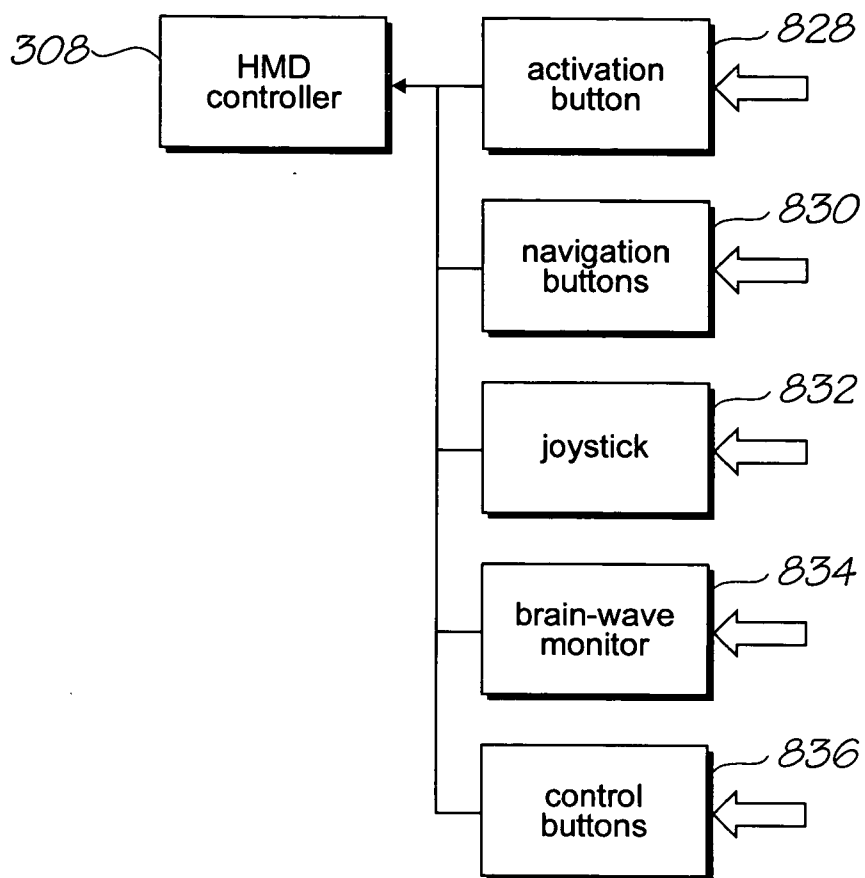


FIG. 27

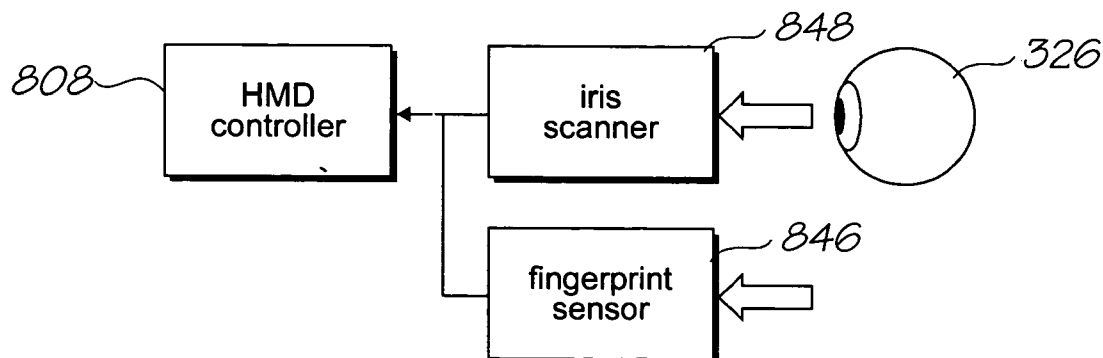


FIG. 28

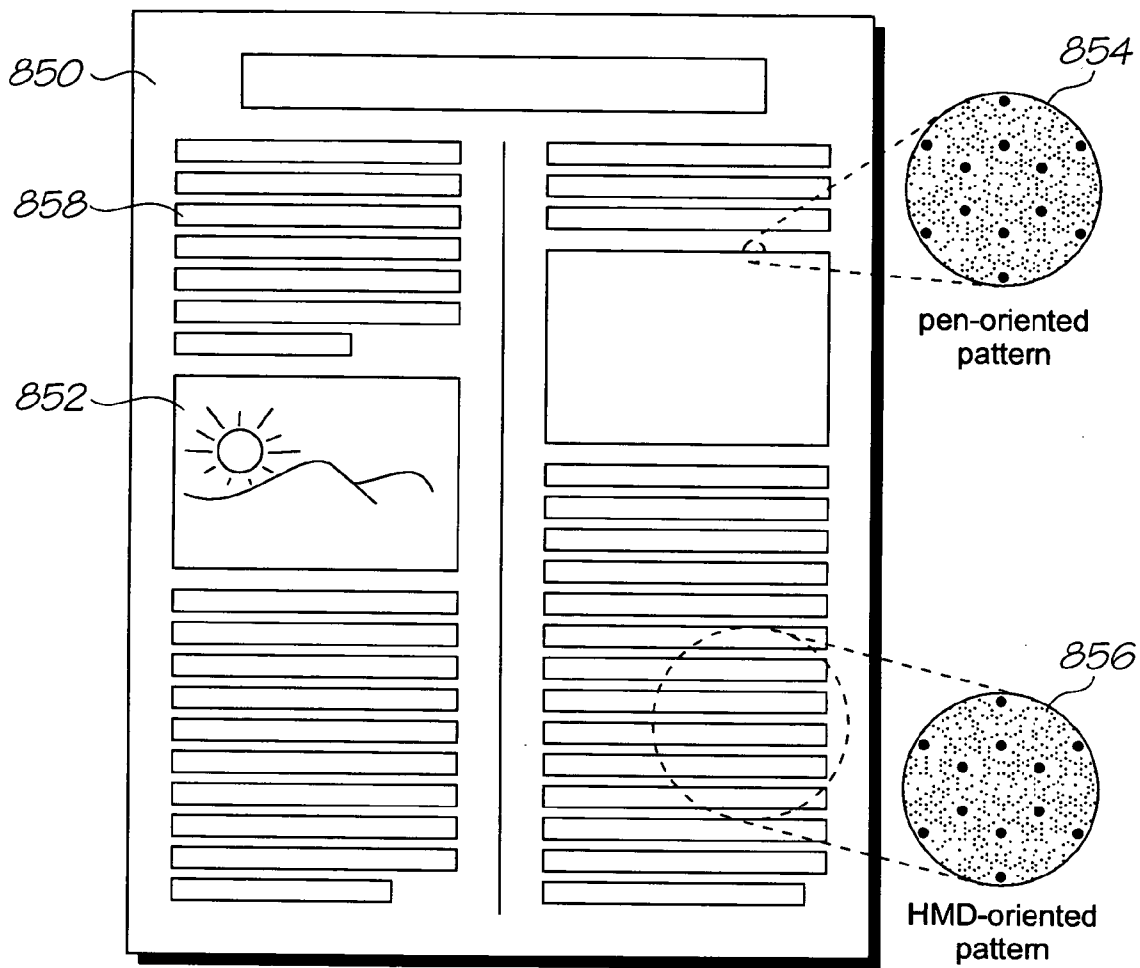


FIG. 29

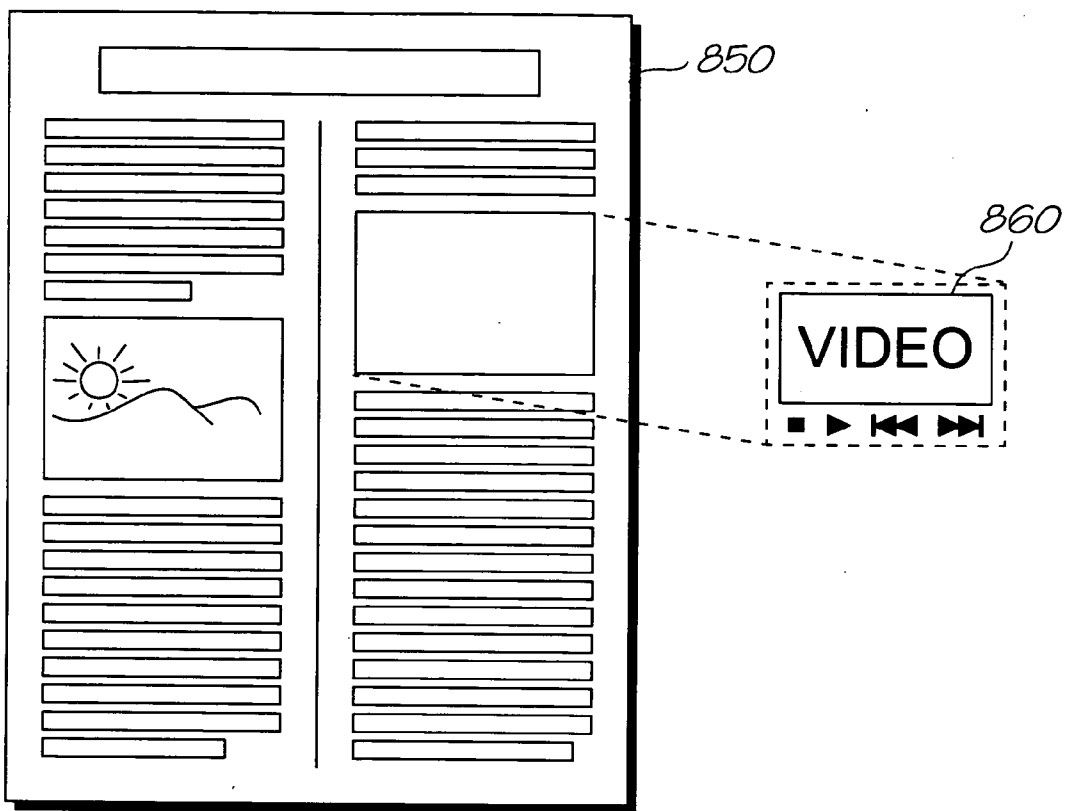


FIG. 30

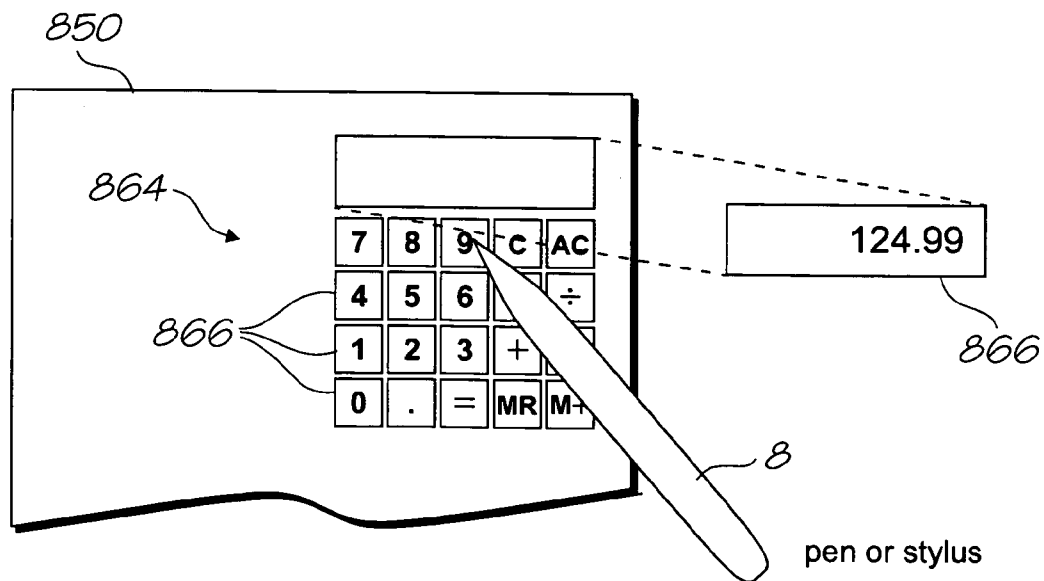


FIG. 31

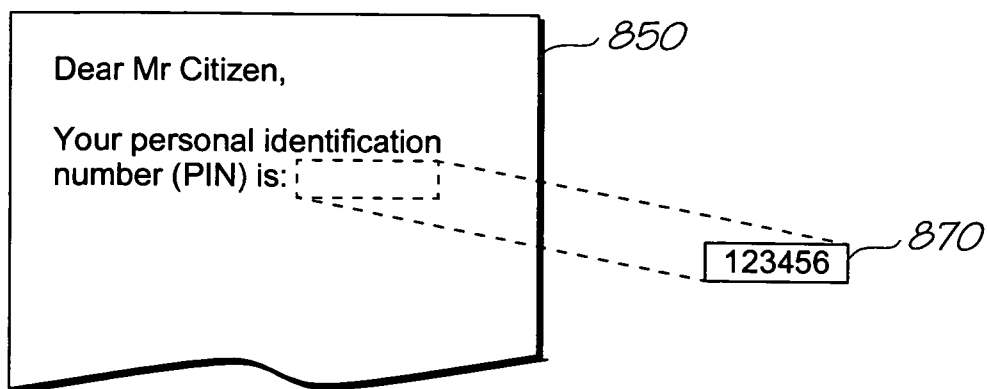


FIG. 32

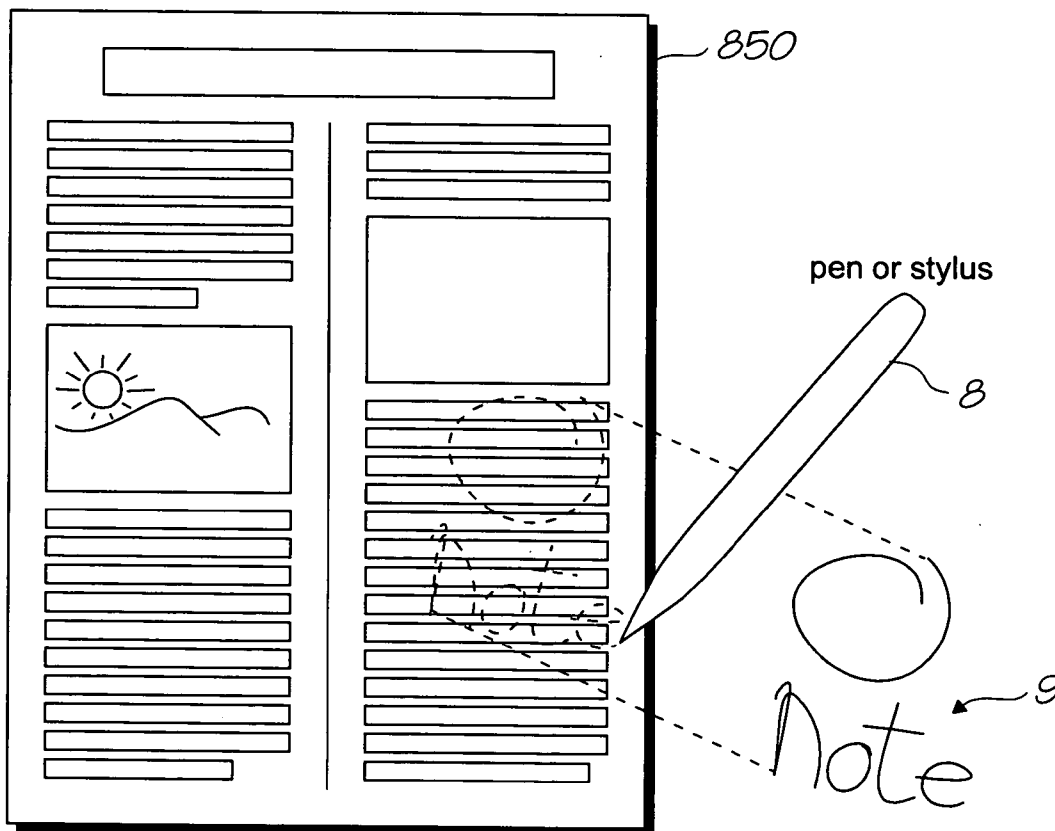


FIG. 33

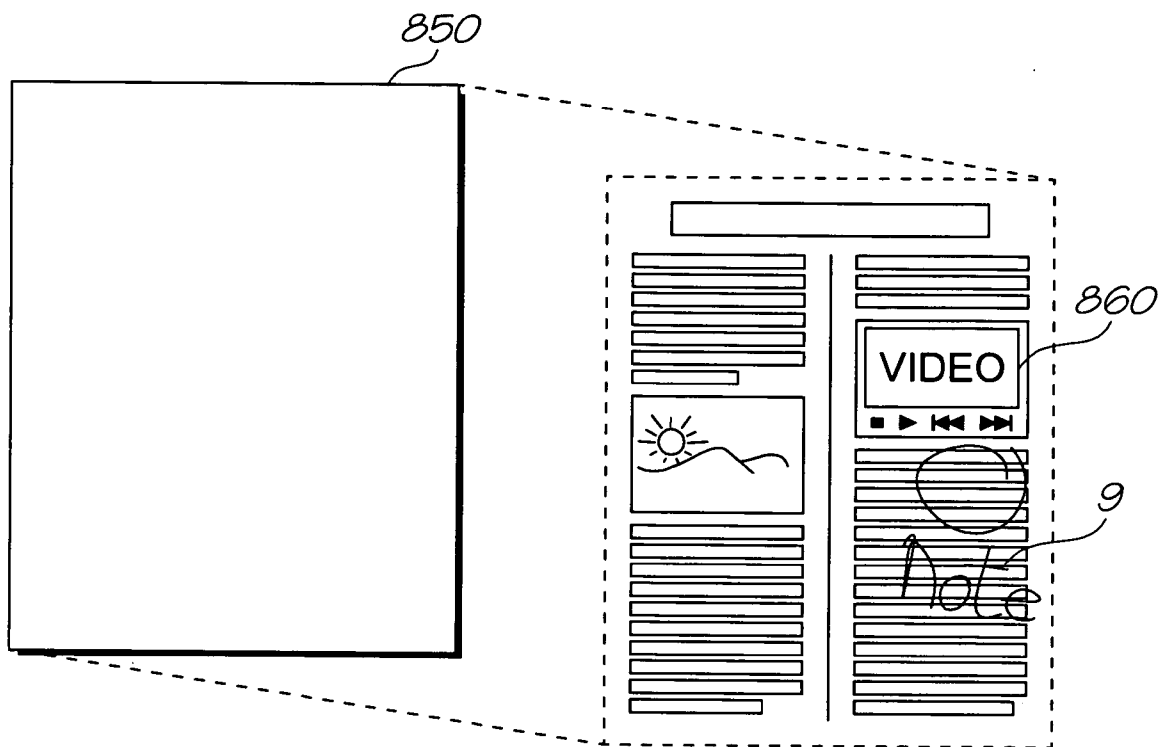


FIG. 34

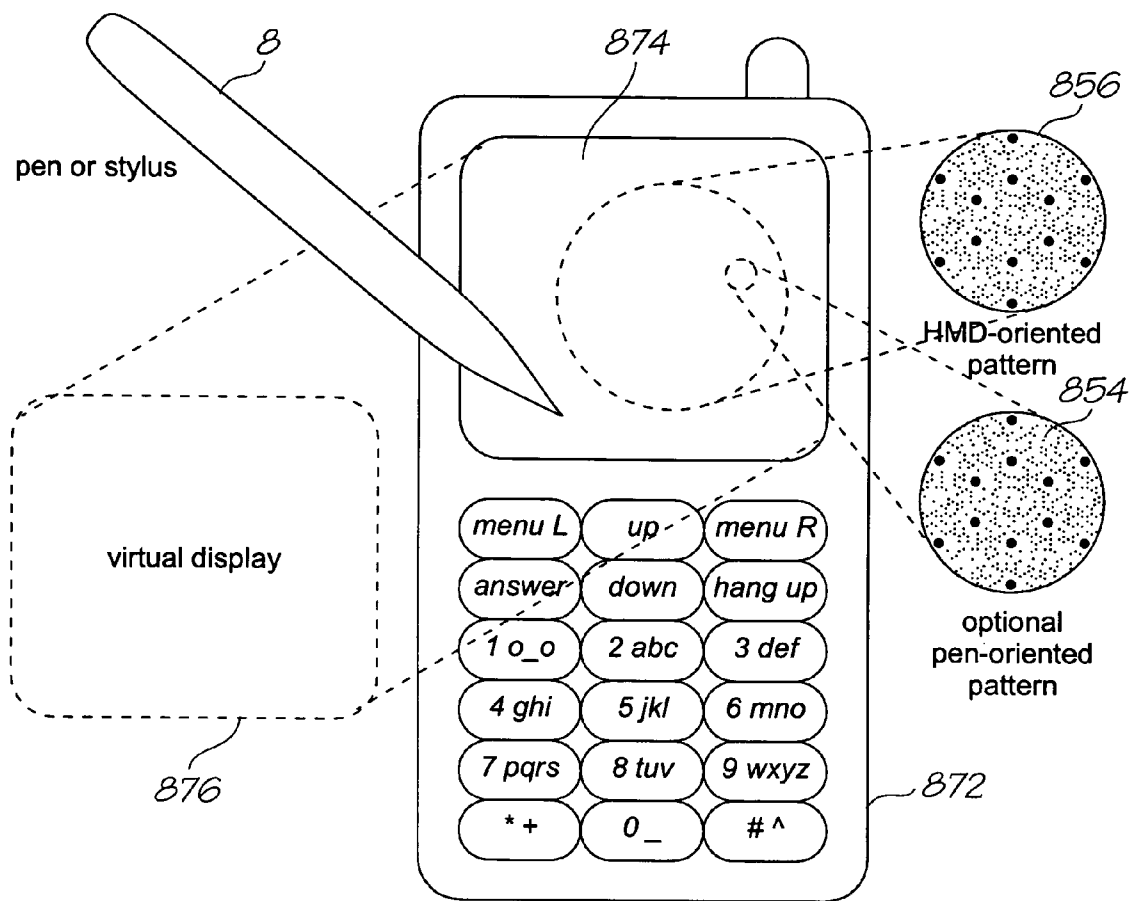


FIG. 35

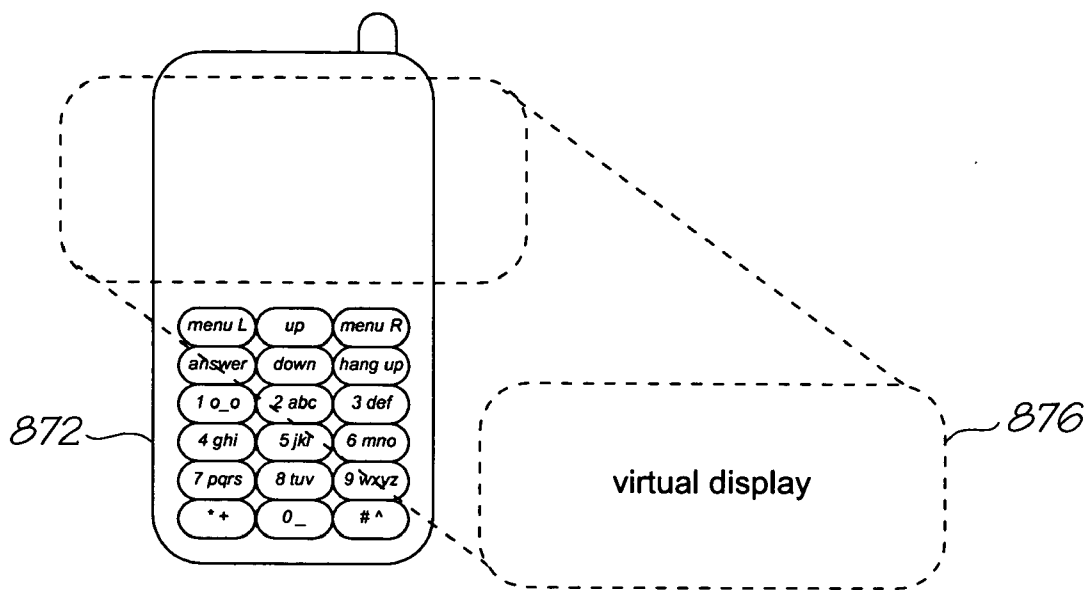


FIG. 36

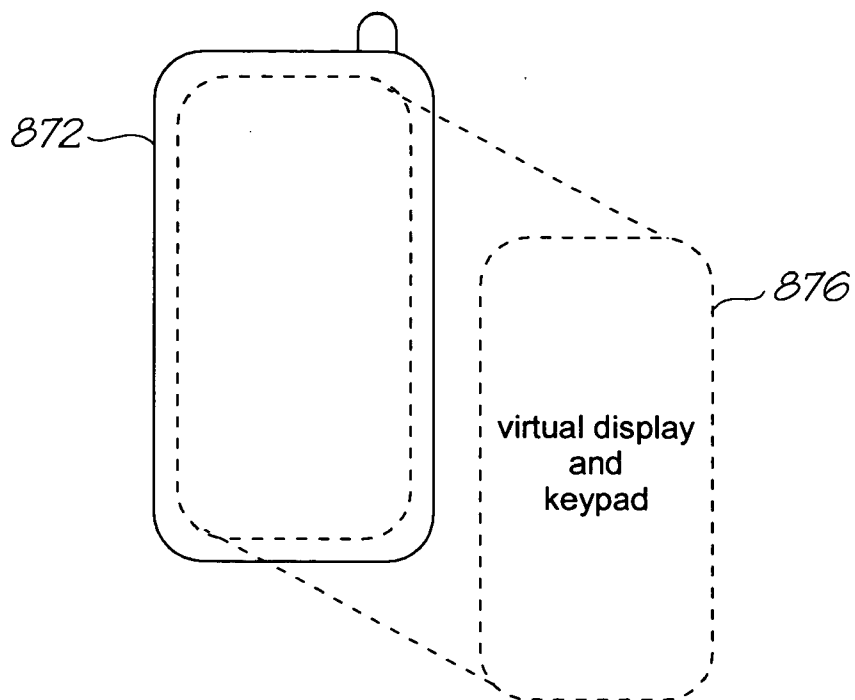


FIG. 37

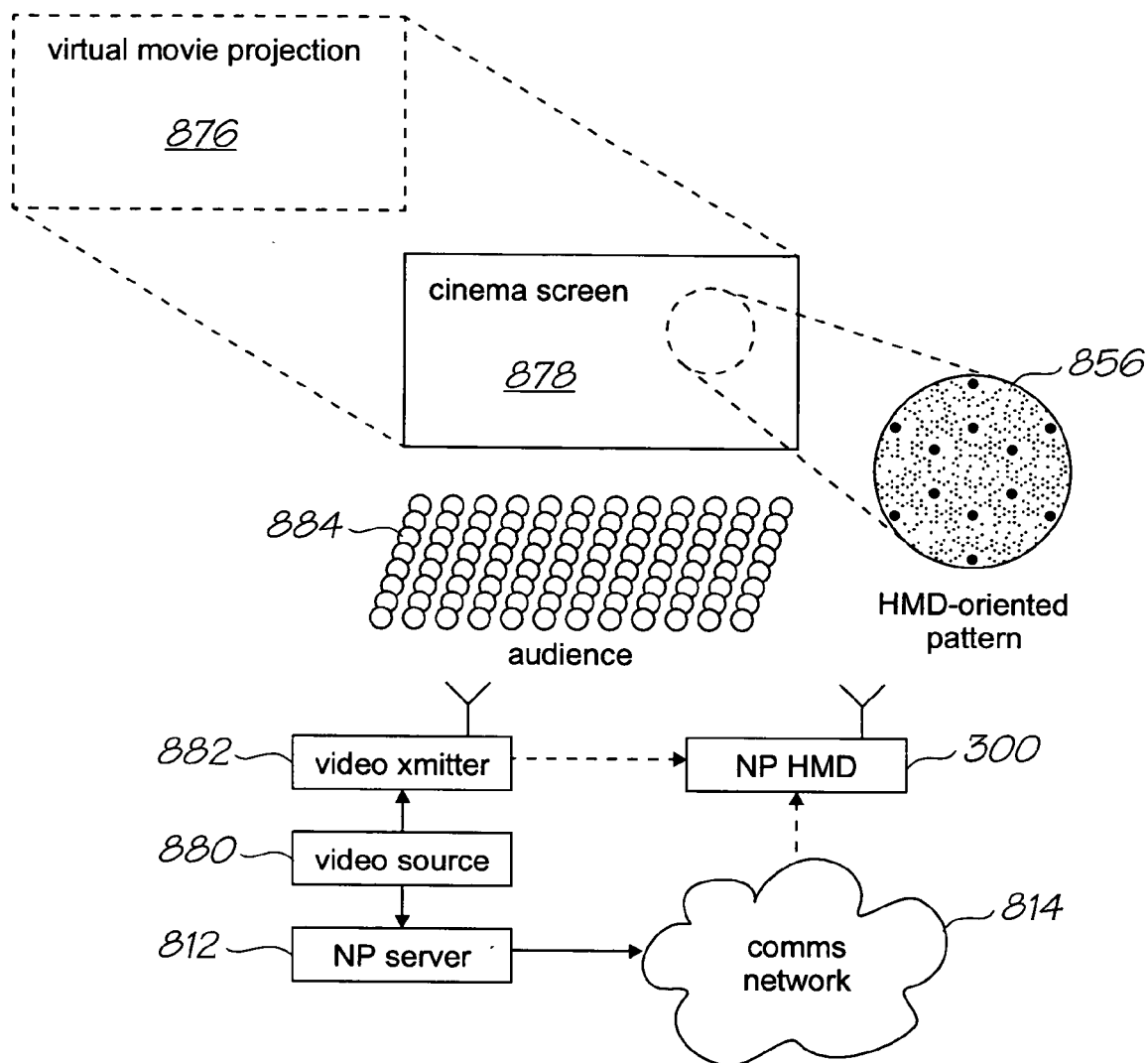


FIG. 38

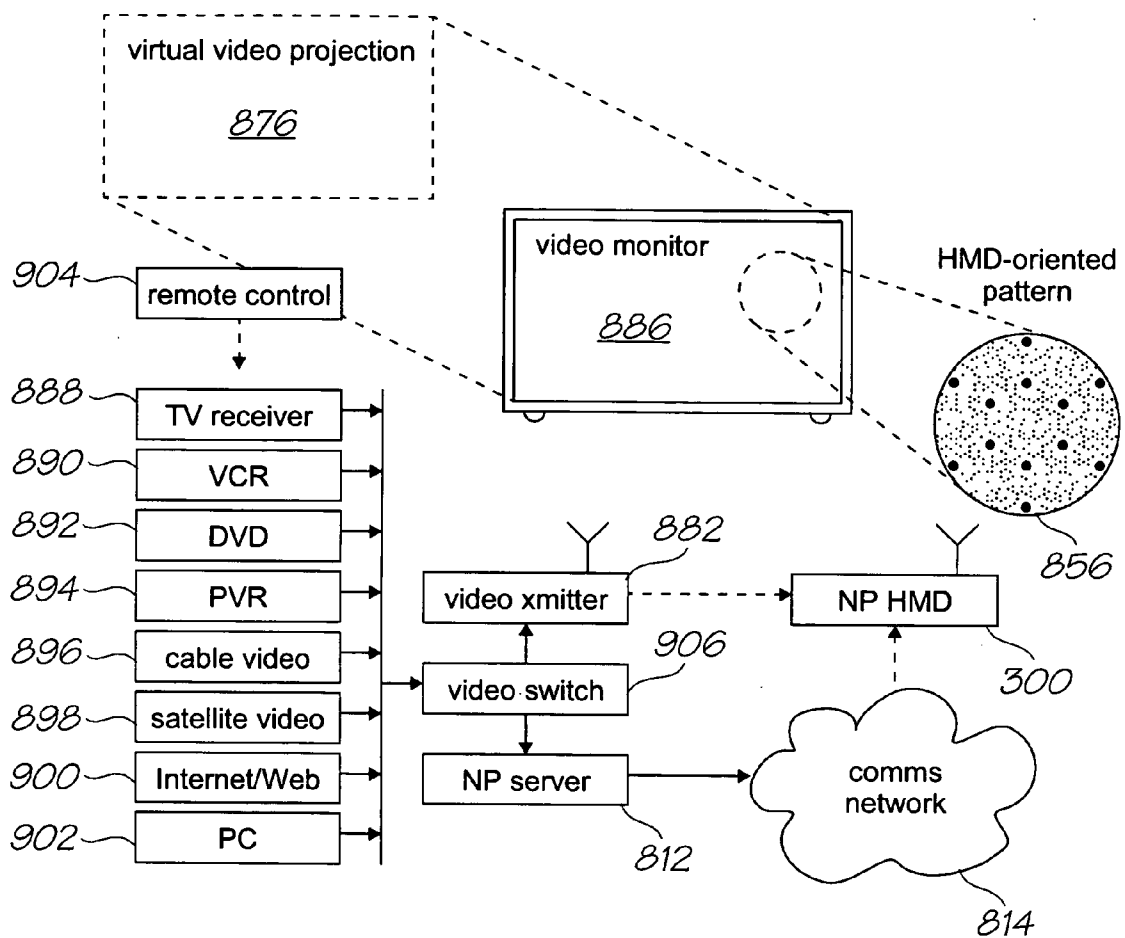


FIG. 39

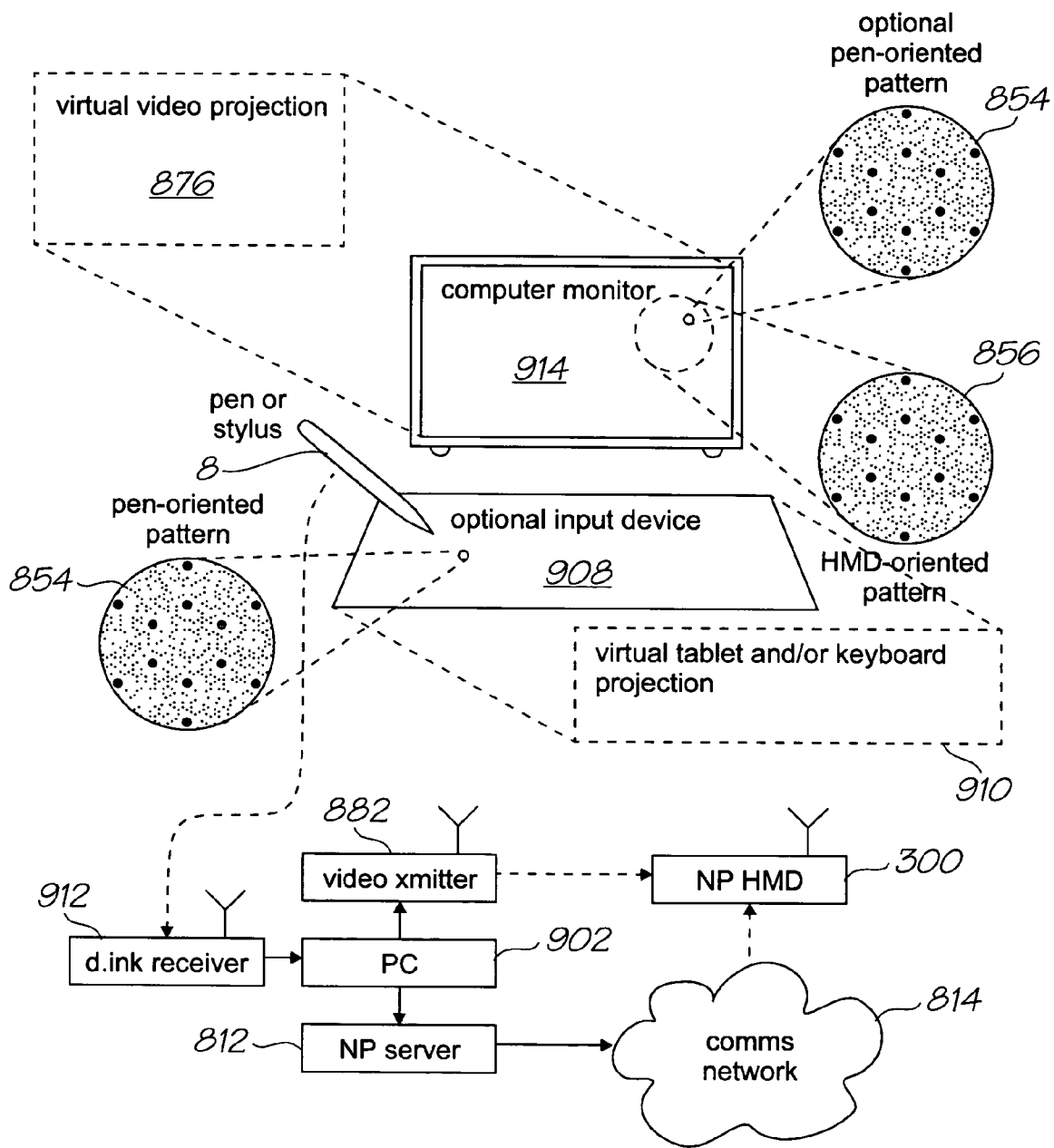


FIG. 40

HEAD MOUNTED DISPLAY WITH WAVE FRONT MODULATOR

FIELD OF THE INVENTION

[0001] The present invention relates to the fields of interactive paper, printing systems, computer publishing, computer applications, human-computer interfaces, information appliances, augmented reality, and head-mounted displays.

CO-PENDING REFERENCES

NPS108US	NPS109US	NPS110US
----------	----------	----------

[0002]

CROSS-REFERENCES

10/815621	10/815612	10/815630	10/815637	10/815638	10/815640	10/815642
10/815643	10/815644	10/815618	10/815639	10/815635	10/815647	10/815634
10/815632	10/815631	10/815648	10/815641	10/815645	10/815646	10/815617
10/815620	10/815615	10/815613	10/815633	10/815619	10/815616	10/815614
10/815636	10/815649	11/041650	11/041651	11/041652	11/041649	11/041610
11/041609	11/041626	11/041627	11/041624	11/041625	11/041556	11/041580
11/041723	11/041698	11/041648	10/815609	10/815627	10/815626	10/815610
10/815611	10/815623	10/815622	10/815629	10/815625	10/815624	10/815628
10/913375	10/913373	10/913374	10/913372	10/913377	10/913378	10/913380
10/913379	10/913376	10/913381	10/986402	IRB013US	11/172815	11/172814
10/409876	10/409848	10/409845	11/084769	11/084742	11/084806	09/575197
09/575195	09/575159	09/575132	09/575123	6825945	09/575130	09/575165
6813039	09/693415	09/575118	6824044	09/608970	09/575131	09/575116
6816274	09/575139	09/575186	6681045	6678499	6679420	09/663599
09/607852	6728000	09/693219	09/575145	09/607656	6813558	6766942
09/693515	09/663701	09/575192	6720985	09/609303	6922779	09/609596
6847883	09/693647	09/721895	09/721894	09/607843	09/693690	09/607605
09/608178	09/609553	09/609233	09/609149	09/608022	09/575181	09/722174
09/721896	10/291522	6718061	10/291523	10/291471	10/291470	6825956
10/291481	10/291509	10/291825	10/291519	10/291575	10/291557	6862105
10/291558	10/291587	10/291818	10/291576	6829387	6714678	6644545
6609653	6651879	10/291555	10/291510	10/291592	10/291542	10/291820
10/291516	6867880	10/291487	10/291520	10/291521	10/291556	10/291821
10/291525	10/291586	10/291822	10/291524	10/291553	6850931	6865570
6847961	10/685523	10/685583	10/685455	10/685584	10/757600	10/804034
10/793933	6889896	10/831232	10/884882	10/943875	10/943938	10/943874
10/943872	10/944044	10/943942	10/944043	10/949293	10/943877	10/965913
10/954170	10/981773	10/981626	10/981616	10/981627	10/974730	10/986337
10/992713	11/006536	11/020256	11/020106	11/020260	11/020321	11/020319
11/026045	11/059696	11/051032	11/059674	NPA19NUS	11/107944	11/107941
11/082940	11/082815	11/082827	11/082829	11/082956	11/083012	11/124256
11/026045	11/059696	11/051032	11/059674	NPA19NUS	11/107944	11/107941
11/082940	11/082815	11/082827	11/082829	11/082956	11/083012	11/124256
11/123136	11/154676	11/159196	NPA225US	09/575193	09/575156	09/609232
09/607844	6457883	09/693593	10/743671	11/033379	09/928055	09/927684
09/928108	09/927685	09/927809	09/575183	6789194	09/575150	6789191
10/900129	10/900127	10/913328	10/913350	10/982975	10/983029	6644642
6502614	6622999	6669385	6827116	10/933285	10/949307	6549935
NPN004US	09/575187	6727996	6591884	6439706	6760119	09/575198
09/722148	09/722146	6826547	6290349	6428155	6785016	6831682
6741871	09/722171	09/721858	09/722142	6840606	10/202021	10/291724
10/291512	10/291554	10/659027	10/659026	10/831242	10/884885	10/884883
10/901154	10/932044	10/962412	10/962510	10/962552	10/965733	10/965933
10/974742	10/982974	10/983018	10/986375	11/107817	11/148238	11/149160
09/693301	6870966	6822639	6474888	6627870	6724374	6788982
09/722141	6788293	09/722147	6737591	09/722172	09/693514	6792165
09/722088	6795593	10/291823	6768821	10/291366	10/291503	6797895
10/274817	10/782894	10/782895	10/778056	10/778058	10/778060	10/778059
10/778063	10/778062	10/778061	10/778057	10/846895	10/917468	10/917467
10/917466	10/917465	10/917356	10/948169	10/948253	10/948157	10/917436
10/943856	10/919379	10/943843	10/943878	10/943849	10/965751	11/071267

-continued

CROSS-REFERENCES

11/144840	11/155556	11/155557	09/575154	09/575129	6830196	6832717
09/721862	10/473747	10/120441	6843420	10/291718	6,789,731	10/291543
6766944	6766945	10/291715	10/291559	10/291660	10/409864	NPT019USNP
10/537159	NPT022US	10/410484	10/884884	10/853379	10/786631	10/853782
10/893372	10/893381	10/893382	10/893383	10/893384	10/971051	10/971145
10/971146	10/986403	10/986404	10/990459	11/059684	11/074802	10/492169
10/492152	10/492168	10/492161	10/492154	10/502575	10/683151	10/531229
10/683040	NPW009USNP	10/510391	10/919260	10/510392	10/919261	10/778090
09/575189	09/575162	09/575172	09/575170	09/575171	09/575161	10/291716
10/291547	10/291538	6786397	10/291827	10/291548	10/291714	10/291544
10/291541	6839053	10/291579	10/291824	10/291713	6914593	10/291546
10/917355	10/913340	10/940668	11/020160	11/039897	11/074800	NPX044US
11/075917	11/102698	11/102843	6593166	10/428823	10/849931	11/144807
6454482	6808330	6527365	6474773	6550997	10/181496	10/274119
10/309185	10/309066	10/949288	10/962400	10/969121	UP21US	UP23US
09/517539	6566858	09/112762	6331946	6246970	6442525	09/517384
09/505951	6374354	09/517608	6816968	6757832	6334190	6745331
09/517541	10/203559	10/203560	10/203564	10/636263	10/636283	10/866608
10/902889	10/902833	10/940653	10/942858	10/727181	10/727162	10/727163
10/727245	10/727204	10/727233	10/727280	10/727157	10/727178	10/727210
10/727257	10/727238	10/727251	10/727159	10/727180	10/727179	10/727192
10/727274	10/727164	10/727161	10/727198	10/727158	10/754536	10/754938
6921144	10/884881	10/943941	10/949294	11/039866	11/123011	11/123010
11/144769	11/148237	10/922846	10/922845	10/854521	10/854522	10/854488
10/854487	10/854503	10/854504	10/854509	10/854510	10/854496	10/854497
10/854495	10/854498	10/854511	10/854512	10/854525	10/854526	10/854516
10/854508	10/854507	10/854515	10/854506	10/854505	10/854493	10/854494
10/854489	10/854490	10/854492	10/854491	10/854528	10/854523	10/854527
10/854524	10/854520	10/854514	10/854519	10/854513	10/854499	10/854501
10/854500	10/854502	10/854518	10/854517	10/934628	11/003786	11/003354
11/003616	11/003418	11/003334	11/003600	11/003404	11/003419	11/003700
11/003601	11/003618	11/003615	11/003337	11/003698	11/003420	11/003682
11/003699	11/071473	11/003463	11/003701	11/003683	11/003614	11/003702
11/003684	11/003619	11/003617	10/760254	10/760210	10/760202	10/760197
10/760198	10/760249	10/760263	10/760196	10/760247	10/760223	10/760264
10/760244	10/760245	10/760222	10/760248	10/760236	10/760192	10/760203
10/760204	10/760205	10/760206	10/760267	10/760270	10/760259	10/760271
10/760275	10/760274	10/760268	10/760184	10/760195	10/760186	10/760261
10/760258	11/014764	11/014763	11/014748	11/014747	11/014761	11/014760
11/014757	11/014714	11/014713	11/014762	11/014724	11/014723	11/014756
11/014736	11/014759	11/014758	11/014725	11/014739	11/014738	11/014737
11/014726	11/014745	11/014712	11/014715	11/014751	11/014735	11/014734
11/014719	11/014750	11/014749	11/014746	11/014769	11/014729	11/014743
11/014733	11/014754	11/014755	11/014765	11/014766	11/014740	11/014720
11/014753	11/014752	11/014744	11/014741	11/014768	11/014767	11/014718
11/014717	11/014716	11/014732	11/014742	11/097268	11/097185	11/097184
10/728804	10/728952	10/728806	10/728834	10/729790	10/728884	10/728970
10/728784	10/728783	10/728925	10/728842	10/728803	10/728780	10/728779
10/773189	10/773204	10/773198	10/773199	6830318	10/773201	10/773191
10/773183	10/773195	10/773196	10/773186	10/773200	10/773185	10/773192
10/773197	10/773203	10/773187	10/773202	10/773188	10/773194	10/773193
10/773184	11/008118	11/060751	11/060805	MTB40US	11/097308	11/097309
11/097335	11/097299	11/097310	11/097213	11/097212	10/760272	10/760273
10/760187	10/760182	10/760188	10/760218	10/760217	10/760216	10/760233
10/760246	10/760212	10/760243	10/760201	10/760185	10/760253	10/760255
10/760209	10/760208	10/760194	10/760238	10/760234	10/760235	10/760183
10/760189	10/760262	10/760232	10/760231	10/760200	10/760190	10/760191
10/760227	10/760207	10/760181	10/407212	10/407207	10/683064	10/683041
6750901	6476863	6788336	6623101	6406129	6505916	6457809
6550895	6457812	10/296434	6428133	6746105		

[0003] The disclosures of these co-pending applications are incorporated herein by cross-reference. Some applications are temporarily identified by their docket number. This will be replaced by the corresponding USSN when available.

BACKGROUND OF THE INVENTION

[0004] Virtual reality completely occludes a person's view of their physical reality (usually with goggles or a helmet) and substitutes an artificial, or virtual view projected on to

the inside of an opaque visor. Augmented reality changes a user's view of the physical environment by adding virtual imagery to the user's field of view (FOV).

[0005] Augmented reality typically relies on either a see-through Head Mounted Display (HMD) or a video-based HMD. A video-based HMD captures video of the user's field of view, augments it with virtual imagery, and redisplay it for the user's eyes to see. A see-through HMD, as discussed above, optically combines virtual imagery with the user's

actual field of view. A video-based HMD has the advantage that registration between the real world and the virtual imagery is relatively easy to achieve, since parallax due to eye position relative to the HMD does not occur. It has the disadvantage that it is typically bulky and has a narrow field of view, and typically provides poor depth cues (i.e. a sense of depth or the distance from the eye to an object).

[0006] A see-through HMD has the advantage that it can be relatively less bulky with a wider field of view, and can provide good depth cues. It has the disadvantage that registration between the real world and the virtual imagery is difficult to achieve without intrusive calibration procedures and sophisticated eye tracking.

[0007] Registration between the real world and the virtual imagery can be provided by inertial sensors to track head movement, or by tracking fiducial markers positioned in the physical environment. The HMD uses the fiducials as reference points for the virtual imagery. A HMD often relies on inertial tracking to maintain registration during head movement, but this is a somewhat inaccurate approach.

[0008] The use of fiducials in the real world is less popular because fiducial tracking is usually not fast enough for typical user head movements, fiducials are typically sparsely placed making fiducial detection complex, and the fiducial encoding capacity is typically small which limits the number of individual fiducials that can uniquely identify themselves. This can lead to fiducial ambiguity in large installations.

SUMMARY OF THE INVENTION

[0009] According to a first aspect, the present invention provides an augmented reality device for inserting virtual imagery into a user's view of their physical environment, the device comprising:

[0010] a display device through which the user can view the physical environment;

[0011] an optical sensing device for sensing at least one surface in the physical environment; and, a controller for projecting the virtual imagery via the display device; wherein during use, the controller uses wave front modulation to match the curvature of the wave fronts of light reflected from the display device to the user's eyes with the curvature of the wave fronts of light that would be transmitted through the device display if the virtual imagery were situated at a predetermined position relative to the surface, such that the user sees the virtual imagery at the predetermined position regardless of changes in position of the user's eyes with respect to the see-through display.

[0012] The human visual system's ability to locate a point in space is determined by the center and radius of curvature of the wavefronts emitted by the point as they impinge on the eyes. A three dimensional object can be thought of as an infinite number of point sources in space.

[0013] The present invention puts each pixel of the virtual image projected by the display device at a predetermined point relative to the sensed surface with a wavefront display that adjusts the curvature of the waves to correspond to the position of the point. This keeps the virtual image in registration with the user's field of view without first establishing (and maintaining) registration between the eye and the see-through display.

[0014] Optionally, the display device has a see-through display for one of the user's eyes. Alternatively, the display device has two see-through displays, one for each of the user's eyes respectively.

[0015] Optionally, the surface has a pattern of coded data disposed on it, such that the controller uses information from the coded data to identify the virtual imagery to be displayed.

[0016] Optionally, the display device, the optical sensing device and the controller are adapted to be worn on the user's head.

[0017] Optionally, the optical sensing device is a camera-based and during use, provides identity and position data related to the coded surface to the controller for determining the virtual imagery displayed.

[0018] Optionally, display device has a virtual retinal display (VRD) for each of the user's eyes, each of the VRD's scans at least one beam of light into a raster pattern and modulates the or each beam to produce spatial variations in the virtual imagery. Optionally, the VRD scans red, green and blue beams of light to produce color pixels in the raster pattern.

[0019] Optionally, the VRD's present a slightly different image to each of the user's eyes, the slight differences being based on eye separation, and the distance to the predetermined position of the virtual imagery to create a perception of depth via stereopsis.

[0020] Optionally, the wavefront modulator uses a deformable membrane mirror, liquid crystal phase corrector, a variable focus liquid lens or a variable focus liquid mirror.

[0021] Optionally, the wave front modulator uses a deformable membrane mirror, liquid crystal phase corrector, a variable focus liquid lens or a variable focus liquid mirror.

[0022] Optionally, the virtual imagery is a movie, a computer application interface, computer application output, hand drawn strokes, text, images or graphics.

[0023] Optionally, the display device has pupil trackers to detect an approximate point of fixation of the user's gaze such that a virtual cursor can be projected into the virtual imagery and navigated using gaze direction.

Additional Aspects

[0024] Related aspects of the invention are set out below together with the a discussion of their backgrounds to provide suitable context for the broad descriptions of these aspects.

Head Mounted Display with Coded Surface Sensor

Background

[0025] As discussed above, the use of fiducials in the real world is less popular because fiducial tracking is usually not fast enough for typical user head movements, fiducials are typically sparsely placed making fiducial detection complex, and the fiducial encoding capacity is typically small which limits the number of individual fiducials that can uniquely identify themselves. This can lead to fiducial ambiguity in large installations.

Summary

[0026] Accordingly, this aspect provides an augmented reality device for a user in a physical environment with a coded surface, the device comprising:

[0027] a display device through which the user can view the physical environment;

[0028] an optical sensing device for sensing the coded surface; and,

[0029] a controller for determining an identity, position and orientation of the coded surface; wherein,

[0030] the controller projects virtual imagery via the display device such that the virtual imagery is viewed by the user in a predetermined position with respect to the coded surface.

[0031] By providing a coded surface instead of sparse fiducials, the invention avoids tracking and ambiguity problems. The relatively dense coding allows the surface to be accurately positioned and oriented to maintain registration with the virtual imagery.

[0032] Optionally, the display device has a see-through display for one of the user's eyes. Alternatively, the display device has two see-through displays, one for each of the user's eyes respectively.

[0033] Optionally, the augmented reality device further comprises a hand-held sensor for sensing and decoding information from the coded surface.

[0034] Optionally, the coded surface has first and second coded data disposed on it in first and second two dimensional patterns respectively, the first pattern having a scale sized such that the optical sensing device can capture images with a resolution suitable for the display device to decode the first coded data, and the second pattern having a scale sized such that the hand-held sensor can capture images with a resolution suitable for it to decode the second coded data.

[0035] Optionally, the hand-held sensor is an electronic stylus with a writing nib wherein during use, the stylus captures images of the second pattern when the nib is in contact with, or proximate to, the coded surface.

[0036] Optionally, the display device, the optical sensing device and the controller are adapted to be worn on the user's head.

[0037] Optionally, the optical sensing device is camera-based and during use, provides identity and position data related to the coded surface to the controller for determining the virtual imagery displayed.

[0038] Optionally, the display device has a virtual retinal display (VRD) for each of the user's eyes, each of the VRD's scans at least one beam of light into a raster pattern and modulates the or each beam to produce spatial variations in the virtual imagery. Optionally, the VRD scans red, green and blue beams of light to produce color pixels in the raster pattern.

[0039] Optionally, each of the virtual retinal displays have a wavefront modulator to match the curvature of the wavefronts of light reflected from the see-through display to the user's eyes with the curvature of the wave fronts of light that would be transmitted through the see-through display for

that eye if the virtual imagery were actual imagery at a predetermined position relative to the coded surface, such that the user views the virtual imagery at the predetermined position regardless of changes in position of the user's eyes with respect to the see-through display.

[0040] Optionally, each of the virtual retinal displays present a slightly different image to each of the user's eyes, the slight differences being based on eye separation, and the distance to the predetermined position of the virtual imagery to create a perception of depth via stereopsis.

[0041] Optionally, the wavefront modulator uses a deformable membrane mirror, liquid crystal phase corrector, a variable focus liquid lens or a variable focus liquid mirror.

[0042] Optionally, the virtual imagery is a movie, a computer application interface, computer application output, hand drawn strokes, text, images or graphics.

[0043] Optionally, the display device has pupil trackers to detect an approximate point of fixation of the user's gaze such that a virtual cursor can be projected into the virtual imagery and navigated using gaze direction.

Virtual Retinal Display with Occlusion Support

Background

[0044] A virtual retinal display (VRD) projects a beam of light onto the eye, and scans the beam rapidly across the eye in a two-dimensional raster pattern. It modulates the intensity of the beam during the scan, based on a source video signal, to produce a spatially-varying image. The combination of human persistence of vision and a sufficiently fast and bright scan creates the perception of an object in the user's field of view.

[0045] The VRD renders occlusions as part of any displayed virtual imagery, according to the user's current viewpoint relative to their physical environment. It does not, however, intrinsically support occlusion parallax according to the position of the user's eye relative to the HMD unless it uses eye tracking for this purpose. In the absence of eye tracking, the HMD renders each VRD view according to a nominal eye position. If the actual eye position deviates from the assumed eye position, then the wavefront display nature of the VRD prevents misregistration between the real world and the virtual imagery, but in the presence of occlusions due to real or virtual objects, it may lead to object overlap or holes.

SUMMARY

[0046] Accordingly, this aspect provides an augmented reality device for inserting virtual imagery into a user's view, the device comprising:

[0047] an optical sensing device for optically sensing the user's physical environment; and,

[0048] a display device with a virtual retinal display for projecting a beam of light as a raster pattern of pixels, each pixel having a wavefront of light with a curvature that provides the user with spatial cues as to the perceived origin of the pixel such that the user perceives the virtual imagery to be at a predetermined location in the physical environment; wherein during use,

- [0049] the virtual retinal display accounts for any occlusions that at least partially obscure the user's view of the perceived location of the virtual imagery by using a spatial light modulator that blocks occluded parts of the wavefront and allows non-occluded parts of the wavefront to pass.
- [0050] To support occlusion parallax, the VRD can be augmented with a spatial light (amplitude) modulator (SLM) such as a digital micromirror device (DMD). The SLM can be introduced immediately after the wavefront modulator and before the raster scanner. The video generator provides the SLM with an occlusion map associated with each pixel in the raster pattern. The SLM passes non-occluded parts of the wavefront but blocks occluded parts. The amplitude-modulation capability of the SLM may be multi-level, and each map entry in the occlusion map may be correspondingly multi-level. However, in the limit case the SLM is a binary device, i.e. either passing light or blocking light, and the occlusion map is similarly binary.
- [0051] Optionally, the VRD projects red, green and blue beams of light, the intensity of each beam being modulated to color each pixel of the raster pattern.
- [0052] Optionally, the VRD has a video generator for providing the spatial light modulator with an occlusion map for each pixel of the raster pattern.
- [0053] Optionally, the display device has a controller connected to the optical sensing device and an image generator for providing image data to the video generator in response to the controller, such that the virtual imagery is selected and positioned by the controller. Optionally, the controller has a data connection to an external source for receiving data related to the virtual imagery.
- [0054] Optionally, the display device has a see-through display such that the VRD projects the raster pattern via the see-through display.
- [0055] In a particularly preferred form the display device has two of the VRDs and two of the see-through displays, one VRD and see-through display for each eye.
- [0056] Optionally, the occlusion is a physical occlusion or a virtual occlusion generated by the controller to at least partially obscure the virtual imagery.
- [0057] Optionally, the display device and the optical sensing device are adapted to be worn on the user's head.
- [0058] Optionally, the optical sensing device senses a surface in the physical environment, the surface having a pattern of coded data disposed on it, such that the display device uses information from the coded data to select and position the virtual imagery to be displayed.
- [0059] Optionally, the optical sensing device is camera-based and during use, provides identity and position data related to the coded surface to the controller for determining the virtual imagery displayed.
- [0060] Optionally, the VRD has a wavefront modulator to match the curvature of the wavefronts of light projected for each pixel in the raster pattern, with the curvature of the wavefronts of light that would be transmitted through the see-through display if the virtual imagery were actual imagery at a predetermined position relative to the coded surface, such that the user views the virtual imagery at the predetermined position regardless of changes in position of the user's eyes with respect to the see-through display.
- [0061] Optionally, the spatial light modulator uses a digital micromirror device to create an occlusion shadow in the scanned raster pattern.
- [0062] Optionally, the camera generates an occlusion map for the scanned raster patterns in the source video signal, and the spatial light modulator uses the occlusion map to control the digital micromirror device.
- [0063] Optionally, each of the VRDs presents a slightly different image to each of the user's eyes, the slight differences being based on eye separation, and the distance to the predetermined position of the virtual imagery to create a perception of depth via stereopsis.
- [0064] Optionally, the wave front modulator has a deformable membrane mirror, liquid crystal phase corrector, a variable focus liquid lens or a variable focus liquid mirror.
- [0065] Optionally, the virtual imagery is a movie, a computer application interface, computer application output, hand drawn strokes, text, images or graphics.
- [0066] Optionally, the display device has pupil trackers to detect an approximate point of fixation of the user's gaze such that a virtual cursor can be projected into the virtual imagery and navigated using gaze direction.

BRIEF DESCRIPTION OF THE DRAWINGS

[0067] Preferred embodiments of the invention will now be described by way of example only with reference to the accompanying drawings, in which:

- [0068] FIG. 1 shows the structure of a complete tag;
- [0069] FIG. 2 shows a symbol unit cell;
- [0070] FIG. 3 shows nine symbol unit cells;
- [0071] FIG. 4 shows the bit ordering in a symbol;
- [0072] FIG. 5 shows a tag with all bits set;
- [0073] FIG. 6 shows a tag group made up of four tag types;
- [0074] FIG. 7 shows the continuous tiling of tag groups;
- [0075] FIG. 8 shows the interleaving of codewords A, B, C & D with a tag;
- [0076] FIG. 9 shows a codeword layout;
- [0077] FIG. 10 shows a tag and its eight immediate neighbours labelled with its corresponding bit index;
- [0078] FIG. 11 shows a user wearing a HMD with single eye display;
- [0079] FIG. 12 shows a user wearing a HMD with respective displays for each eye;
- [0080] FIG. 13 is a schematic representation of a camera capturing light rays from two point sources;
- [0081] FIG. 14 is a schematic representation of a display of the image of the two points sources captured by the camera of FIG. 13;
- [0082] FIG. 15 is a schematic representation of a wavefront display of a virtual point source of light;

[0083] FIG. 16 is a diagrammatic representation of a HMD with a single eye display;

[0084] FIG. 17a schematically shows a wavefront display using a DMM;

[0085] FIG. 17b schematically shows the wavefront display of FIG. 17a with the DMM deformed to diverge the project beam;

[0086] FIG. 18a schematically shows a wavefront display using a deformable liquid lens;

[0087] FIG. 18b schematically shows the wavefront display of FIG. 18a with the liquid lens deformed to diverge the projected beam;

[0088] FIG. 19 diagrammatically shows the modification to the HMD of FIG. 16 in order to support occlusions;

[0089] FIG. 20 schematically shows the wavefront display of FIG. 15 with occlusion support;

[0090] FIG. 21 schematically shows the wavefront display of FIG. 18b modified for occlusion support;

[0091] FIG. 22 is a diagrammatic representation of a HMD with a binocular display;

[0092] FIG. 23 shows a HMD directly linked to the Netpage server;

[0093] FIG. 24 shows the HMD linked to a Netpage Pen and a Netpage server via a communications network FIG. 25 shows a HMD linked to a Netpage relay which is in turn linked to a Netpage server via a communications network;

[0094] FIG. 26 schematically shows a HMD with image warper;

[0095] FIG. 27 shows a HMD linked to a cursor navigation and selection devices;

[0096] FIG. 28 shows a HMD with biometric sensors;

[0097] FIG. 29 shows a physical Netpage with pen-scale and HMD-scale tag patterns;

[0098] FIG. 30 shows the SVD on a printed Netpage;

[0099] FIG. 31 shows printed calculator with a SVD for the display and Netpage pen;

[0100] FIG. 32 shows a printed form with a SVD for a text field displaying confidential information;

[0101] FIG. 33 shows the page of FIG. 29 with handwritten annotations captured as digital ink and shown as a SVD;

[0102] FIG. 34 shows a Netpage with static and dynamic page elements incorporated into the SVD;

[0103] FIG. 35 shows a mobile phone with display screen printed with pen-scale and HMD-scale tag patterns;

[0104] FIG. 36 shows a mobile phone with SVD that extends beyond the display screen;

[0105] FIG. 37 shows a mobile phone with display screen and keypad provided by the SVD;

[0106] FIG. 38 shows a cinema screen with HMD-scale tag pattern for screening movies as SVD's;

[0107] FIG. 39 shows a video monitor with HMD-scale tag pattern for a SVD of a video signal from a range of sources; and

[0108] FIG. 40 shows a computer screen with pen-scale and HMD-scale tag patterns, and a tablet with a pen-scale tag pattern for an SVD of a keyboard.

DETAILED DESCRIPTION

[0109] As discussed above, the invention is well suited for incorporation in the Assignee's Netpage system. In light of this, the invention has been described as a component of a broader Netpage architecture. However, it will be readily appreciated that augmented reality devices have much broader application in many different fields. Accordingly, the present invention is not restricted to a Netpage context.

[0110] Additional cross referenced documents are listed at the end of the Detailed Description. These documents are predominantly non-patent literature and have been numbered for identification at the relevant part of the description. The disclosures of these documents are incorporated by cross reference.

Netpage Surface Coding

Introduction

[0111] This section defines a surface coding used by the Netpage system (described in co-pending application Docket No.

[0112] NPS110US as well as many of the other cross referenced documents listed above) to imbue otherwise passive surfaces with interactivity in conjunction with Netpage sensing devices (described below).

[0113] When interacting with a Netpage coded surface, a Netpage sensing device generates a digital ink stream which indicates both the identity of the surface region relative to which the sensing device is moving, and the absolute path of the sensing device within the region.

Surface Coding

[0114] The Netpage surface coding consists of a dense planar tiling of tags. Each tag encodes its own location in the plane. Each tag also encodes, in conjunction with adjacent tags, an identifier of the region containing the tag. In the Netpage system, the region typically corresponds to the entire extent of the tagged surface, such as one side of a sheet of paper.

[0115] Each tag is represented by a pattern which contains two kinds of elements. The first kind of element is a target. Targets allow a tag to be located in an image of a coded surface, and allow the perspective distortion of the tag to be inferred. The second kind of element is a macrodot. Each macrodot encodes the value of a bit by its presence or absence.

[0116] The pattern is represented on the coded surface in such a way as to allow it to be acquired by an optical imaging system, and in particular by an optical system with a narrowband response in the near-infrared. The pattern is typically printed onto the surface using a narrowband near-infrared ink.

Tag Structure

[0117] FIG. 1 shows the structure of a complete tag 200. Each of the four black circles 202 is a target. The tag 200, and the overall pattern, has four-fold rotational symmetry at the physical level.

[0118] Each square region represents a symbol 204, and each symbol represents four bits of information. Each symbol 204 shown in the tag structure has a unique label 216. Each label 216 has an alphabetic prefix and a numeric suffix.

[0119] FIG. 2 shows the structure of a symbol 204. It contains four macrodots 206, each of which represents the value of one bit by its presence (one) or absence (zero).

[0120] The macrodot 206 spacing is specified by the parameters throughout this specification. It has a nominal value of 143 μm , based on 9 dots printed at a pitch of 1600 dots per inch. However, it is allowed to vary within defined bounds according to the capabilities of the device used to produce the pattern.

[0121] FIG. 3 shows an array 208 of nine adjacent symbols 204. The macrodot 206 spacing is uniform both within and between symbols 208.

[0122] FIG. 4 shows the ordering of the bits within a symbol 204.

[0123] Bit zero 210 is the least significant within a symbol 204; bit three 212 is the most significant. Note that this ordering is relative to the orientation of the symbol 204. The orientation of a particular symbol 204 within the tag 200 is indicated by the orientation of the label 216 of the symbol in the tag diagrams (see for example FIG. 1). In general, the orientation of all symbols 204 within a particular segment of the tag 200 is the same, consistent with the bottom of the symbol being closest to the centre of the tag.

[0124] Only the macrodots 206 are part of the representation of a symbol 204 in the pattern. The square outline 214 of a symbol 204 is used in this specification to more clearly elucidate the structure of a tag 204. FIG. 5, by way of illustration, shows the actual pattern of a tag 200 with every bit 206 set. Note that, in practice, every bit 206 of a tag 200 can never be set.

[0125] A macrodot 206 is nominally circular with a nominal diameter of $(5/9)s$. However, it is allowed to vary in size by $\pm 10\%$ according to the capabilities of the device used to produce the pattern.

[0126] A target 202 is nominally circular with a nominal diameter of $(17/9)s$. However, it is allowed to vary in size by $\pm 10\%$ according to the capabilities of the device used to produce the pattern.

[0127] The tag pattern is allowed to vary in scale by up to 10% according to the capabilities of the device used to produce the pattern. Any deviation from the nominal scale is recorded in the tag data to allow accurate generation of position samples.

Tag Groups

[0128] Tags 200 are arranged into tag groups 218. Each tag group contains four tags arranged in a square. Each tag 200 has one of four possible tag types, each of which is

labelled according to its location within the tag group 218. The tag type labels 220 are 00, 10, 01 and 11, as shown in FIG. 6.

[0129] FIG. 7 shows how tag groups are repeated in a continuous tiling of tags, or tag pattern 222. The tiling guarantees the any set of four adjacent tags 200 contains one tag of each type 220.

Codewords

[0130] The tag contains four complete codewords. The layout of the four codewords is shown in FIG. 8. Each codeword is of a punctured 2^4 -ary (8, 5) Reed-Solomon code. The codewords are labelled A, B, C and D. Fragments of each codeword are distributed throughout the tag 200.

[0131] Two of the codewords are unique to the tag 200. These are referred to as local codewords 224 and are labelled A and B. The tag 200 therefore encodes up to 40 bits of information unique to the tag.

[0132] The remaining two codewords are unique to a tag type, but common to all tags of the same type within a contiguous tiling of tags 222. These are referred to as global codewords 226 and are labelled C and D, subscripted by tag type. A tag group 218 therefore encodes up to 160 bits of information common to all tag groups within a contiguous tiling of tags.

Reed-Solomon Encoding

[0133] Codewords are encoded using a punctured 2^4 -ary (8, 5) Reed-Solomon code. A 2^4 -ary (8, 5) Reed-Solomon code encodes 20 data bits (i.e. five 4-bit symbols) and 12 redundancy bits (i.e. three 4-bit symbols) in each codeword. Its error-detecting capacity is three symbols. Its error-correcting capacity is one symbol.

[0134] FIG. 9 shows a codeword 228 of eight symbols 204, with five symbols encoding data coordinates 230 and three symbols encoding redundancy coordinates 232. The codeword coordinates are indexed in coefficient order, and the data bit ordering follows the codeword bit ordering.

[0135] A punctured 2^4 -ary (8, 5) Reed-Solomon code is a 2^4 -ary (15, 5) Reed-Solomon code with seven redundancy coordinates removed. The removed coordinates are the most significant redundancy coordinates.

[0136] The code has the following primitive polynomial:

$$p(x)=x^4+x+1 \quad (\text{EQ } 1)$$

[0137] The code has the following generator polynomial:

$$g(x)=(x+\alpha)(x+\alpha^2) \dots (x+\alpha^6) \quad (\text{EQ } 2)$$

[0138] For a detailed description of Reed-Solomon codes, refer to Wicker, S. B. and V. K. Bhargava, eds., Reed-Solomon Codes and Their Applications, IEEE Press, 1994, the contents of which are incorporated herein by reference.

The Tag Coordinate Space

[0139] The tag coordinate space has two orthogonal axes labelled x and y respectively. When the positive x axis points to the right, then the positive y axis points down.

[0140] The surface coding does not specify the location of the tag coordinate space origin on a particular tagged

surface, nor the orientation of the tag coordinate space with respect to the surface. This information is application-specific.

[0141] For example, if the tagged surface is a sheet of paper, then the application which prints the tags onto the paper may record the actual offset and orientation, and these can be used to normalise any digital ink subsequently captured in conjunction with the surface.

[0142] The position encoded in a tag is defined in units of tags. By convention, the position is taken to be the position of the centre of the target closest to the origin.

Tag Information Content

[0143] Table 1 defines the information fields embedded in the surface coding. Table 2 defines how these fields map to codewords.

TABLE 1

Field definitions	
field	width description
<u>per codeword</u>	
codeword type	2 The type of the codeword, i.e. one of A (b'00'), B (b'01'), C (b'10') and D (b'11').
<u>per tag</u>	
tag type	2 The type ¹ of the tag, i.e. one of 00 (b'00'), 01 (b'01'), 10 (b'10') and 11 (b'11').
x coordinate	13 The unsigned x coordinate of the tag ² .
y coordinate	13 The unsigned y coordinate of the tag ² .
active area flag	1 A flag indicating whether the tag is a member of an active area. b'1' indicates membership.
active area map flag	1 A flag indicating whether an active area map is present. b'1' indicates the presence of a map (see next field). If the map is absent then the value of each map entry is derived from the active area flag (see previous field).
active area map	8 A map ³ of which of the tag's immediate eight neighbours are members of an active area. b'1' indicates membership.
data fragment	8 A fragment of an embedded data stream. Only present if the active area map is absent.
<u>per tag group</u>	
encoding format	8 The format of the encoding. 0: the present encoding Other values are TBA.
region flags	8 Flags controlling the interpretation and routing of region-related information. 0: region ID is an EPC 1: region is linked 2: region is interactive 3: region is signed 4: region includes data 5: region relates to mobile application Other bits are reserved and must be zero.
tag size adjustment	16 The difference between the actual tag size and the nominal tag size ⁴ , in 10 nm units, in sign-magnitude format.
region ID	96 The ID of the region containing the tags.
CRC	16 A CRC ⁵ of tag group data.
total	320

¹corresponds to the bottom two bits of the x and y coordinates of the tag
²allows a maximum coordinate value of approximately 14 m
³FIG. 29 indicates the bit ordering of the map
⁴the nominal tag size is 1.7145 mm (based on 1600 dpi, 9 dots per macrodot, and 12 macrodots per tag)
⁵CCITT CRC-16 [7]

[0144] FIG. 10 shows a tag 200 and its eight immediate neighbours, each labelled with its corresponding bit index in the active area map. An active area map indicates whether the corresponding tags are members of an active area. An active area is an area within which any captured input should be immediately forwarded to the corresponding Netpage server for interpretation. It also allows the Netpage sensing device to signal to the user that the input will have an immediate effect.

TABLE 2

Mapping of fields to codewords					
codeword	codeword bits	field	width	field bits	
A	1:0	codeword type (b'00')	2	all	
	10:2	x coordinate	9	12:4	
	19:11	y coordinate	9	12:4	
	B	1:0	codeword type (b'01')	2	all
2		tag type	1	0	
5:2		x coordinate	4	3:0	
6		tag type	1	1	
9:6		y coordinate	4	3:0	
10		active area flag	1	all	
11		active area map flag	1	all	
19:12		active area map	8	all	
19:12		data fragment	8	all	
C ₀₀		1:0	codeword type (b'10')	2	all
		9:2	encoding format	8	all
		17:10	region flags	8	all
	19:18	tag size adjustment	2	1:0	
C ₀₁	1:0	codeword type (b'10')	2	all	
	15:2	tag size adjustment	14	15:2	
C ₁₀	19:16	region ID	4	3:0	
	1:0	codeword type (b'10')	2	all	
C ₁₁	19:2	region ID	18	21:4	
	1:0	codeword type (b'10')	2	all	
D ₀₀	19:2	region ID	18	39:22	
	1:0	codeword type (b'11')	2	all	
D ₀₁	19:2	region ID	18	57:40	
	1:0	codeword type (b'11')	2	all	
D ₁₀	19:2	region ID	18	75:58	
	1:0	codeword type (b'11')	2	all	
D ₁₁	19:2	region ID	18	93:76	
	1:0	codeword type (b'11')	2	all	
	3:2	region ID	2	95:94	
	19:4	CRC	16	all	

[0145] Note that the tag type can be moved into a global codeword to maximise local codeword utilization. This in turn can allow larger coordinates and/or 16-bit data fragments (potentially configurably in conjunction with coordinate precision). However, this reduces the independence of position decoding from region ID decoding and has not been included in the specification at this time.

Embedded Data

[0146] If the "region includes data" flag in the region flags is set then the surface coding contains embedded data. The data is encoded in multiple contiguous tags' data fragments, and is replicated in the surface coding as many times as it will fit.

[0147] The embedded data is encoded in such a way that a random and partial scan of the surface coding containing the embedded data can be sufficient to retrieve the entire data. The scanning system reassembles the data from retrieved fragments, and reports to the user when sufficient fragments have been retrieved without error.

[0148] As shown in Table 3, a 200-bit data block encodes 160 bits of data. The block data is encoded in the data fragments of A contiguous group of 25 tags arranged in a 5x5 square. A tag belongs to a block whose integer coordinate is the tag's coordinate divided by 5: Within each block the data is arranged into tags with increasing x coordinate within increasing y coordinate.

[0149] A data fragment may be missing from a block where an active area map is present. However, the missing data fragment is likely to be recoverable from another copy of the block.

[0150] Data of arbitrary size is encoded into a superblock consisting of a contiguous set of blocks arranged in a rectangle. The size of the superblock is encoded in each block. A block belongs to a superblock whose integer coordinate is the block's coordinate divided by the superblock size. Within each superblock the data is arranged into blocks with increasing x coordinate within increasing y coordinate.

[0151] The superblock is replicated in the surface coding as many times as it will fit, including partially along the edges of the surface coding.

[0152] The data encoded in the superblock may include more precise type information, more precise size information, and more extensive error detection and/or correction data.

TABLE 3

Embedded data block		
field	width	description
data type	8	The type of the data in the superblock. Values include: 0: type is controlled by region flags 1: MIME Other values are TBA.
superblock width	8	The width of the superblock, in blocks.
superblock height	8	The height of the superblock, in blocks.
data	160	The block data.
CRC	16	A CRC ⁶ of the block data.
total	200	

⁶CCITT CRC-16 [7]

Cryptographic Signature of Region ID

[0153] If the "region is signed" flag in the region flags is set then the surface coding contains a 160-bit cryptographic signature of the region ID. The signature is encoded in a one-block superblock.

[0154] In an online environment any signature fragment can be used, in conjunction with the region ID, to validate the signature. In an offline environment the entire signature can be recovered by reading multiple tags, and can then be validated using the corresponding public signature key. This is discussed in more detail in Netpage Surface Coding

Security section of the cross reference co-pending application Docket No. NPS100US the content of which is incorporated within the present specification.

MIME Data

[0155] If the embedded data type is "MIME" then the superblock contains Multipurpose Internet Mail Extensions (MIME) data according to RFC 2045 (see Freed, N., and N. Borenstein, "Multipurpose Internet Mail Extensions (MIME)—Part One: Format of Internet Message Bodies", RFC 2045, November 1996), RFC 2046 (see Freed, N., and N. Borenstein, "Multipurpose Internet Mail Extensions (MIME)—Part Two: Media Types", RFC 2046, November 1996) and related RFCs. The MIME data consists of a header followed by a body. The header is encoded as a variable-length text string preceded by an 8-bit string length. The body is encoded as a variable-length type-specific octet stream preceded by a 16-bit size in big-endian format.

[0156] The basic top-level media types described in RFC 2046 include text, image, audio, video and application.

[0157] RFC 2425 (see Howes, T., M. Smith and F. Dawson, "A MIME Content-Type for Directory Information", RFC 2045, September 1998) and RFC 2426 (see Dawson, F., and T. Howes, "vCard MIME Directory Profile", RFC 2046, September 1998) describe a text subtype for directory information suitable, for example, for encoding contact information which might appear on a business card.

Encoding and Printing Considerations

[0158] The Print Engine Controller (PEC) supports the encoding of two fixed (per-page) 2⁴-ary (15, 5) Reed-Solomon codewords and six variable (per-tag) 2⁴ (15, 5) Reed-Solomon codewords. Furthermore, PEC supports the rendering of tags via a rectangular unit cell whose layout is constant (per page) but whose variable codeword data may vary from one unit cell to the next. PEC does not allow unit cells to overlap in the direction of page movement.

[0159] A unit cell compatible with PEC contains a single tag group consisting of four tags. The tag group contains a single A codeword unique to the tag group but replicated four times within the tag group, and four unique B codewords. These can be encoded using five of PEC's six supported variable codewords. The tag group also contains eight fixed C and D codewords. One of these can be encoded using the remaining one of PEC's variable codewords, two more can be encoded using PEC's two fixed codewords, and the remaining five can be encoded and pre-rendered into the Tag Format Structure (TFS) supplied to PEC.

[0160] PEC imposes a limit of 32 unique bit addresses per TFS row. The contents of the unit cell respect this limit. PEC also imposes a limit of 384 on the width of the TFS. The contents of the unit cell respect this limit.

[0161] Note that for a reasonable page size, the number of variable coordinate bits in the A codeword is modest, making encoding via a lookup table tractable. Encoding of the B codeword via a lookup table may also be possible. Note that since a Reed-Solomon code is systematic, only the redundancy data needs to appear in the lookup table.

Imaging and Decoding Considerations

[0162] The minimum imaging field of view required to guarantee acquisition of an entire tag has a diameter of 39.6

s (i.e. $(2 \times (12+2))\sqrt{2}$ s), allowing for arbitrary alignment between the surface coding and the field of view. Given a macrodot spacing of 143 μ m, this gives a required field of view of 5.7 mm.

[0163] Table 4 gives pitch ranges achievable for the present surface coding for different sampling rates, assuming an image sensor size of 128 pixels.

TABLE 4

Pitch ranges achievable for present surface coding for different sampling rates; dot pitch = 1600 dpi, macrodot pitch = 9 dots, viewing distance = 30 mm, nib-to-FOV separation = 1 mm, image sensor size = 128 pixels	
sampling rate	pitch range
2	-40 to +49
2.5	-27 to +36
3	-10 to +18

[0164] Given the present surface coding, the corresponding decoding sequence is as follows:

- [0165] locate targets of complete tag
- [0166] infer perspective transform from targets
- [0167] sample and decode any one of tag's four codewords
- [0168] determine codeword type and hence tag orientation
- [0169] sample and decode required local (A and B) codewords
- [0170] codeword redundancy is only 12 bits, so only detect errors
- [0171] on decode error flag bad position sample
- [0172] determine tag x-y location, with reference to tag orientation
- [0173] infer 3D tag transform from oriented targets
- [0174] determine nib x-y location from tag x-y location and 3D transform
- [0175] determine active area status of nib location with reference to active area map
- [0176] generate local feedback based on nib active area status
- [0177] determine tag type from A codeword
- [0178] sample and decode required global (C and D) codewords (modulo window alignment, with reference to tag type)
- [0179] although codeword redundancy is only 12 bits, correct errors; subsequent CRC verification will detect erroneous error correction
- [0180] verify tag group data CRC
- [0181] on decode error flag bad region ID sample
- [0182] determine encoding type, and reject unknown encoding
- [0183] determine region flags

- [0184] determine region ID
- [0185] encode region ID, nib x-y location, nib active area status in digital ink
- [0186] route digital ink based on region flags

[0187] Note that region ID decoding need not occur at the same rate as position decoding.

[0188] Note that decoding of a codeword can be avoided if the codeword is found to be identical to an already-known good codeword.

Head Mounted Display

[0189] The Netpage system provides a paper- and pen-based interface to computer-based and typically network-based information and applications. The Netpage coding is discussed in detail above and the Netpage pen is described in the above cross referenced documents and in particular, a co-filed US application, temporarily identified here by its docket NPS109US.

[0190] The Netpage Head Mounted Display is an augmented reality device that can use surfaces coded with Netpage tag patterns to situate a virtual image in a user's field of view. The virtual imagery need not be in precise registration with the tagged surface, but can be 'anchored' to the tag pattern so that it appears to be part of the user's physical environment regardless of whether they change their direction of gaze.

Overview

[0191] A printed Netpage, when presented in a user's field of view (FOV), can be augmented with dynamic imagery virtually projected onto the page via a see-through head-mounted display (HMD) worn by the user. The imagery is selected according to the unique identity of the Netpage, and is virtually projected to match the three-dimensional position and orientation of the page with respect to the user. The imagery therefore appears locked to the surface of the page, even as the position and orientation of the page changes due to head or page movement. The HMD provides the correct stereopsis, vergence and accommodation cues to allow fatigue-free perception of the imagery "on" the surface. "Stereopsis", "vergence" and "accommodation" relate to depth cues that the brain uses for three dimensional spatial awareness of objects in the FOV. These terms are explained below in the description of the Human Visual System.

[0192] Although the imagery is "attached" to the surface, it can still be three-dimensional and extend "out of" the surface. The page is coded with identity- and position-indicating tags in the usual way, but at a larger scale to allow longer-range acquisition. The HMD uses a Netpage sensor to image the tags and thereby identify the page and determine its position and orientation. If the page also supports pen interaction, then it may be coded with two sets of tags at different scales and utilising different infrared inks; or it may be coded with a multi-resolution tags which can be imaged and decoded at multiple scales; or the HMD tag sensor can be adapted to image and decode pen-scale tags. In any case the whole page surface is ideally tagged so that it remains identifiable even when partially obscured, such as by another page or by the user's hand. The Netpage HMD is lightweight and portable. It uses a radio interface to query a Netpage system and obtain static and dynamic page data.

It uses an on-board processor to determine page position and orientation, and to project imagery in real time to minimise display latency.

[0193] The Netpage HMD, in conjunction with a suitable Netpage, therefore provides a situated virtual display (SVD) capability. The display is situated in that its location and content are page-driven. It is virtual in that it is only virtually projected on the page and is therefore only seen by the user. Note that the Netpage Viewer [8] and the Netpage Explorer [3] both provide Netpage SVD capabilities, but in more constrained forms.

[0194] An SVD can be used to display a video clip embedded in a printed news article; it can be used to show an object virtually associated with a page, such as a “pasted” photo; it can be used to show “secret” information associated with a page; and it can be used to show the page itself, for example in the absence of ambient light. More generally, an SVD can transform a page (or any surface) into a general-purpose display device, and more generally still, into a general-purpose computer system interface. SVDs can augment or subsume all current “display” applications, whether they be static or dynamic, passive or interactive, personal or shared, including such applications as commercial print publications, on-demand printed documents, product packaging, posters and billboards, television, cinema, personal computers, personal digital assistants (PDAs), mobile phones, smartphones and other personal devices. As well as augmenting the planar surfaces of essentially two-dimensional objects such as paper pages, SVDs can equally augment the multi-faceted or non-planar surfaces of three-dimensional objects.

[0195] Augmented reality in general typically relies on either a see-through HMD or a video-based HMD [15]. A video-based HMD captures video of the user’s field of view, augments it with virtual imagery, and redisplay it for the user’s eyes to see. A see-through HMD, as discussed above, optically combines virtual imagery with the user’s actual field of view. A video-based HMD has the advantage that registration between the real world and the virtual imagery is relatively easy to achieve, since parallax due to eye position relative to the HMD doesn’t occur. It has the disadvantage that it is typically bulky and has a narrow field of view, and typically provides poor depth cues.

[0196] As shown in FIGS. 11 and 12, a see-through HMD has the advantage that it can be relatively less bulky with a wider field of view, and can provide good depth cues. It has the disadvantage that registration between the real world and the virtual imagery is difficult to achieve without intrusive calibration procedures and sophisticated eye tracking. A HMD often relies on inertial tracking to maintain registration during head movement, since fiducial tracking is usually insufficiently fast, but this is a somewhat inaccurate approach.

[0197] In a basic form, the HMD 300 may have a single display 302 for one eye only. However, as shown in FIG. 12 by using a wave front display 304, 306 for each eye respectively, the Netpage HMD 300 achieves perfect registration in a see-through display without calibration or tracking.

[0198] The use of fiducials in the real world to provide a basis for registration is well-established in augmented reality

applications [15, 44]. However, fiducials are typically sparsely placed, making fiducial detection complex, and the fiducial encoding capacity is typically small, leading to a small fiducial identity space and fiducial ambiguity in large installations.

[0199] The surface coding used by the Netpage system is dense, overcoming sparseness issues encountered with fiducials. The Netpage system guarantees global identifier uniqueness, overcoming ambiguity issues encountered with fiducials. More broadly, the Netpage system provides the first systematic and practical mechanism for coding a significant proportion of the surfaces with which people interact on a day-to-day basis, providing an unprecedented opportunity to deploy augmented reality technology in a consumer setting. The scope of Netpage applications, and the universality of the devices used to interact with Netpage coded surfaces, makes the acquisition and assimilation of Netpage devices extremely attractive to consumers.

[0200] The tag image processing and decoding system developed for Netpage operates in real time at high-quality display frame rates (e.g. 100 HZ or higher). It therefore obviates the need for inaccurate inertial tracking.

The Human Visual System

[0201] The human eye consists of a converging lens system, made up of the cornea and crystalline lens, and a light-sensitive array of photoreceptors, the retina, onto which the lens system projects a real image of the eye’s field of view. The cornea provides a fixed amount of focus which constitutes over two thirds of the eye’s focusing power, while the crystalline lens provides variable focus under the control of the ciliary muscles which surround it. When the muscles are relaxed the lens is almost flat and the eye is focused at infinity. As the muscles contract the lens bulges, allowing the eye to focus more closely. The point of closest achievable focus, the near point, recedes with age. It may be less than 10 cm in a teenager, but usually exceeds 25 cm by middle age.

[0202] A diaphragm known as the iris controls the amount of light entering the eye and defines its entrance pupil. It can expand to as much as 8 mm in darkness and contract to as little as 2 mm in bright light.

[0203] The limits of the visual field of the eye are about 60 degrees upwards, 75 degrees downwards, 60 degrees inwards (in the nasal direction), and about 90 degrees outwards (in the temporal direction). The visual fields of the two eyes overlap by about 120 degrees centrally. This defines the region of binocular vision.

[0204] The retina consists of an uneven distribution of about 130 million photoreceptor cells. Most of these, the so-called rods, exhibit broad spectral sensitivity in the visible spectrum. A much smaller number (about 7 million), the so-called cones, variously exhibit three kinds of relatively narrower spectral sensitivity, corresponding to short, medium and long wavelength parts of the visible spectrum. The rods confer monochrome sensitivity in low lighting conditions, while the cones confer color sensitivity in relatively brighter lighting conditions. The human visual system effectively interpolates short, medium and long-wavelength cone stimuli in order to perceive spectral color.

[0205] The highest density of cones occurs in a small central region of the retina known as the macula. The macula

contains the fovea, which in turn contains a tiny rod-free central region known as the foveola. The retina subtends about 3.3 degrees of visual angle per mm. The macula, at about 5 mm, subtends about 17 degrees; the fovea, at about 1.5 mm, about 5 degrees; and the foveola, at about 0.4 mm, about 1.3 degrees. The density of photoreceptors in the retina falls off gradually with eccentricity, in line with increasing photoreceptor size. A line through the center of the foveola and the center of the pupil defines the eye's visual axis. The visual axis is tilted inwards (in the nasal direction) by about 5 degrees with respect to the eye's optical axis.

[0206] The photoreceptors in the retina connect to about a million retinal ganglion cells which convey visual information to the brain via the optic nerve. The density of ganglion cells falls off linearly with eccentricity, and much more rapidly than the density of photoreceptors. This linear fall-off confers scale-invariant imaging. In the foveola, each ganglion cell connects to an individual cone. Elsewhere in the retina a single ganglion cell may connect to many tens of rods and cones. Foveal visual acuity peaks at around 4 cycles per degree, is a couple of orders of magnitude less at 30 cycles per degree, and is immeasurable beyond about 60 cycles per degree [33]. This upper limit is consistent with the maximum cone density in the foveola of around twice this number, and the corresponding ganglion cell density. Visual acuity drops rapidly with eccentricity. For a 5-degree visual field, it drops to 50% of peak acuity at the edges. For a 30-degree visual field, it drops to 5%.

[0207] The human visual system provides two distinct modes of visual perception, operating in parallel. The first supports global analysis of the visual field, allowing a object of interest to be detected, for example due to movement. The second supports detailed analysis of the object of interest.

[0208] In order to perceive and analyse an object of interest in detail, the head and/or the eyes are rapidly moved to align the eyes' visual axes with the object of interest. This is referred to as fixation, and allows high-resolution foveal imaging of the object of interest. Fixational movements, or saccades, and fixational pauses, during which foveal imaging takes place, are interleaved to allow the brain to perceive and analyse an extended object in detail. An initial gross saccade of arbitrary magnitude provides initial fixation. This is followed by a series of finer saccades, each of at most a few degrees, which scan the object onto the foveola. Microsaccades, a fraction of a degree in extent, are implicated in the perception of very fine detail, such as individual text characters. An ocular tremor, known as nystagmus, ensures continuous relative movement between the retina and a fixed scene. Without this tremor, retinal adaptation would cause the perceived image to fade out.

[0209] Although peripheral attention usually leads to foveal attention via fixation, the brain is also capable of attending to a peripheral point of interest without fixating on it.

[0210] Light emitted by a point source creates a series of spherical wavefronts centered on the point source. When the wavefronts impinge on the human eye, the human visual system is able to change the shape of the crystalline lens to bring the wavefronts to a point of focus on the retina. This is referred to as accommodation. The curvature of each wavefront as it impinges on the eye is the inverse of the

distance from the point source to the eye. The smaller the distance, the greater the wavefront curvature, and the greater the accommodation required. The greater the distance, the flatter the wavefronts, and the smaller the accommodation required.

[0211] In order to fixate on a point source, the human visual system rotates each eye so that the point source is aligned with the visual axis of each eye. This is referred to as vergence. Vergence in turn helps control the accommodation response, and a mismatch between vergence and accommodation cues can therefore cause eye strain.

[0212] The state of accommodation and vergence of the eyes in turn provides the visual system with a cue to the distance from the eyes to the point source, i.e. with a sense of depth.

[0213] The disparity between the relative positions of multiple point sources in the two eyes' fields of view provides the visual system with a cue to their relative depth. This disparity is referred to as binocular parallax. The visual system's process of fusing the inputs from the two eyes and thereby perceiving depth is referred to as stereopsis. Stereopsis in turn helps achieve vergence and accommodation.

[0214] Binocular parallax and motion parallax, i.e. parallax induced by relative motion, are the two most powerful depth cues used by the human visual system. Note that parallax may also lead to an occlusion disparity.

[0215] The visual system's ability to locate a point source in space is therefore determined by the center and radius of curvature of the wavefronts emitted by the point source as they impinge on the eyes. Furthermore, the discussion of point sources applies equally to extended objects in general, by considering the surface of each extended object as consisting of an infinite number of point sources. In practice, due to the finite resolving power of the visual system, a finite number of point sources is suffice to model an extended object.

[0216] Persistence of vision describes the inability of the human visual system, and the retina in particular, to detect changes in intensity occurring above a certain critical frequency. This critical fusion frequency (CFF) is between 50 and 60 Hz, and is somewhat dependent on contrast and luminance conditions. It provides the basis for the human visual system's flicker-free perception of projected film and video.

Three-Dimensional Displays

[0217] If one imagines a spherical camera capable of capturing three-dimensional images of its surrounding space, and a corresponding spherical display capable of displaying them, then a defining characteristic of the display is that it becomes invisible when placed in the same location as the camera, no matter how it is viewed. The display emits the same light as would have been emitted by the space it occupies had it not been present. More conventionally, one can imagine a camera surface capable of recording all light penetrating it from one side, and a corresponding display surface capable of emitting corresponding light. This is illustrated in **FIG. 13**, where the camera **308** is shown capturing a subset of rays **310** emitted by a pair of point sources **312**. **FIG. 14** shows the display **314** is shown emitting corresponding rays **316**. In reality, a larger number

of rays are captured and displayed than shown in **FIG. 14**, so a viewer will perceive the point sources **312** as being correctly located at fixed points in three-dimensional space, independently of viewing position.

[0218] The capture and manipulation of true three-dimensional image data has been the subject of much research in recent years, mainly for the purpose of constructing novel views. The images captured by an infinite collection of infinitely small spherical cameras define the so-called plenoptic function [42], while the light penetrating an arbitrary surface in three dimensions defines a so-called light field [36,30]. Both functions, although theoretically continuous, are typically discretized for practical manipulation, and are resampled to construct novel views. Although the discussion so far has posited a 3D camera, the camera can be virtual and a light field can be generated from a virtual 3D model.

[0219] A light field has the advantage that it captures both position and occlusion parallax. It has the disadvantage that it is data-intensive compared with a traditional 2D image. Conceptually, compared with a view-dependent 2D image, a discretized view-independent light field is defined by an array of 2D images, each image corresponding to a pixel in the view-dependent image. Although a light field can be used to generate a 2D image for a novel view, it is expensive to directly display a 2D light field. Because of this, 3D light field displays such as the lenslet display described in [35] only support relatively low spatial resolution. Furthermore, although the light field samples can be seen as samples of a suitably low-pass filtered set of wavefronts, the discrete light field display does not reconstruct the continuous wavefronts which the samples represent, relying instead on approximate integration by the human visual system.

[0220] Synthetic holographic displays have similar resolution problems [52].

[0221] **FIG. 15** shows a simple wavefront display **322** of a virtual point source of light **318**. In contrast to a discrete light field display, a wavefront display emits a set of continuous spherical wavefronts **324**. The centre of curvature of each wavefront in the set is the virtual point source of light **318**. If the virtual point **318** was an actual point, it would be emitting spherical wavefronts **320**. The wavefronts **324** emitted from the display **322** are equivalent to the virtual wavefronts **320** had they passed through the display **322**.

[0222] The advantage of the wavefront display **322** is that the description of the input 3D image is much smaller than the description of the corresponding light field, since it consists of a 2D image augmented with depth information. The disadvantage of this representation is that it fails to represent occlusion parallax. However, in applications where occlusion parallax is not important, the wavefront display has clear advantages.

[0223] A volumetric display acts as a simple wavefront display [24], but has the disadvantage that the volume of the display must encompass the volume of the virtual object being displayed.

[0224] A virtual retinal display [27], as discussed in the next section, can act as a simple wavefront display when augmented with a wavefront modulator [43]. Unlike a volumetric display, it can simulate arbitrary depth. It can be further augmented with a spatial light modulator [32] to support occlusions.

[0225] Many simpler display technologies have been developed which provide some of the cues used by the human visual system to perceive depth. These display technologies are predominantly stereoscopic, i.e. they present a different view to each eye and rely on binocular disparity to stimulate depth perception. In a stereoscopic head-mounted display, left and right views are presented directly to each eye. Left and right views may also be spectrally multiplexed on a conventional display and viewed through glasses with a different filter for each eye, or time-multiplexed on a conventional display and viewed through glasses which shutter each eye in alternating fashion. Polarization is also commonly used for view separation. In an autostereoscopic display, so called because it allows stereoscopic viewing without encumbering the viewer with headgear or eyewear, strips of the left and right view images are typically interleaved and displayed together. When viewed through a parallax barrier or a lenticular array, the left eye sees only the strips comprising the left image, and the right eye sees only the strips comprising the right image. These displays often only provide horizontal parallax, only support limited variation in the position and orientation of the viewer, and only provide two viewing zones, i.e. one for each eye. As discussed above, arrays of lenslets can be used to directly display light fields and thus provide omnidirectional parallax [35], dynamic parallax barrier methods can be used to support wider movement of a single tracked viewer [50], and multi-projector lenticular displays can be used to provide a larger number of viewing zones to multiple simultaneous viewers [40]. In a head-mounted display, motion parallax results from rendering views according to the tracked position and orientation of the viewer, whereas in a multiview autostereoscopic system, motion parallax is intrinsic although typically of lower quality.

The Netpage Head-Mounted Display

[0226] The Netpage HMD utilises a virtual retinal display ⁷ (VRD) for each eye. A VRD projects a beam of light directly onto the eye, and scans the beam rapidly across the eye in a two-dimensional raster pattern. It modulates the intensity of the beam during the scan, based on a source video signal, to produce a spatially-varying image. The combination of human persistence of vision and a sufficiently fast and bright scan creates the perception of an object in the user's field of view.

⁷Also referred to as a Retinal Scanning Display (RSD).

[0227] The VRD utilises independent red, green and blue beams to create a colour display. The tri-stimulus nature of the human visual system allows a red-green-blue display system to stimulate the perception of most perceptible colours. Although a colour display capability is preferred, a monochromatic display capability also has utility.

[0228] Rendering the image presented to each eye differently according to eye separation and virtual object depth creates the perception of depth via stereopsis. Adjusting the projection angle into each eye to allow correct vergence further enhances depth perception, as does adjusting the divergence of each beam to allow correct accommodation. Apart from reinforcing depth perception, consistent depth cues maximise viewer comfort.

[0229] Key to the operation of the Netpage HMD is the registration of the image projected by the VRD with the

surface of the Netpage onto which the image is being virtually projected. By operating as a limited wavefront display, a VRD allows this registration to be achieved without requiring registration between the eye and the VRD. In this regard it differs from screen-based HMDs, which require careful calibration or monitoring of eye position relative to the HMD to achieve and maintain registration. Thus the view-independent nature of a wavefront display is exploited to avoid registration between the eye and the HMD, rather than its more conventional purpose of avoiding a HMD altogether in the context of an autostereoscopic display. As an alternative to exploiting a VRD for this purpose, a view-independent light field display can also be used, using a much faster laser scan.

[0230] A VRD provides only a limited wavefront display capability because of practical limits on the size of its exit pupil. Ideally its exit pupil is large enough to cover the eye's maximum entrance pupil, at any allowed position relative to the display. The position of the eye's pupil relative to the display can vary due to eye movements, variations in the placement of the HMD, and variations in individual human anatomy. In practice it is advantageous to track the approximate gaze direction of the eye relative to the display, so that limited system resources can be dedicated to generating display output where it will be seen and/or at an appropriate resolution.

[0231] Tracking the pupil also allows the system to determine an approximate point of fixation, which it can use to identify a document of interest. In a Netpage context, projecting virtual imagery onto the surface region to which the user is directing foveal attention is most important. It is less critical to project imagery into the periphery of the user's field of view. Gaze tracking can also be used to navigate a virtual cursor, or to indicate an object to be selected or otherwise activated, such as a hyperlink.

[0232] In a Netpage context, the surface onto which the virtual imagery is being projected can generally be assumed to be planar, and for most applications the projected virtual object can similarly be assumed to be planar. This simplifies the wavefront display requirements of the Netpage HMD. In particular, the wavefront curvature is not required to vary abruptly within a scanline. Alternatively, if the curvature modulation mechanism is slow, then the wavefront curvature can be fixed for an entire frame, e.g. based on the average depth of the virtual object. If the wavefront curvature cannot be varied automatically at all, then the system may still provide the user with a manual adjustment mechanism for setting the curvature, e.g. based on the user's normal viewing distance. Alternatively, the wavefront curvature may be fixed by the system based on a standard viewing distance, e.g. 50 cm, to maximise viewer comfort. FIG. 16 shows a block diagram of a VRD suitable for use in the Netpage HMD, similar in structure to VRDs described in [27, 28, 37 and 38].

[0233] The VRD as a whole scans a light beam across the eye 326 in a two-dimensional raster pattern. The eye 326 focuses the beam 390 onto the retina to produce a spot which traces out the raster pattern over time. At any given time, the intensity of the beam and hence the spot represents the value of a single colour pixel in a two-dimensional input image. Human persistence of vision fuses the moving spot into the perception of a two-dimensional image. The required pixel

rate of the VRD is the product of the image resolution and the frame rate. The frame rate in turn is at least as high as the critical fusion frequency, and ideally higher (e.g. 100 Hz or more). By way of example, a frame rate of 100 Hz and a spatial resolution 2000 pixels by 2000 pixels gives a pixel rate of 400 MHz and a line rate of 200 kHz.

[0234] A video generator 328 accepts a stream of image data 330 and generates the requisite data and control signals 332 for displaying the image data 330.

[0235] Light beam generators 334 generate red, green and blue beams 336, 338 and 340 respectively. Each beam generator 334 has a matching intensity modulator 342, for modulating the intensity of each beam according to the corresponding component of the pixel colour 344 supplied by the video generator 328.

[0236] The beam generator 334 may be a gas or solid-state laser, a light-emitting diode (LED), or a super-luminescent LED. The intensity modulator 342 may be intrinsic to the beam generator or may be a separate device. For example, a gas laser may rely on a downstream acousto-optic modulator (AOM) for intensity modulation, while a solid-state laser or LED may intrinsically allow intensity modulation via its drive current.

[0237] Although FIG. 16 shows multiple beam generators 334 and colour intensity modulators 342, a single monochrome beam generator may be utilised if color projection is not required.

[0238] Furthermore, multiple beam generators and intensity modulators may be utilised in parallel to achieve a desired pixel rate. In general, any component of the VRD whose fundamental operating rate limits the achievable pixel rate may be replicated, and the replicated components operated in parallel, to achieve a desired pixel rate.

[0239] A beam combiner 346 combines the intensity modulated colored beams 348, 350 and 352 into a single beam 354 multiple colored beams into a single beam suitable for scanning. The beam combiner may utilise multiple beam splitters.

[0240] A wavefront modulator 356 accepts the collimated input beam 354 and modulates its wavefront to induce a curvature which is the inverse of the pixel depth signal 358 supplied by the video generator 328. The pixel depth 358 is clipped at a reasonable depth, beyond which the wavefront modulator 356 passes a collimated beam. The wavefront modulator 356 may be a deformable membrane mirror (DMM) [43, 51], a liquid-crystal phase corrector [47], a variable focus liquid lens or mirror operating on an electrowetting principle [16, 25], or any other suitable controllable wavefront modulator. Depending on the time constant of the modulator 356, it may be utilised to effect pixel-wise, line-wise or frame-wise wavefront modulation, corresponding to pixel-wise, line-wise or frame-wise constant depth. Furthermore, as mentioned earlier, multiple wavefront modulators may be utilised in parallel to achieve higher-rate wavefront modulation. If the operation of the wavefront modulator is wavelength-dependent, then multiple wavefront modulators may be employed beam-wise before the beams are combined. Even if the wavefront modulator is incapable of random pixel-wise modulation, it may still be capable of ramped modulation corresponding to the linear change of depth within a single scanline of the projection of a planar object.

[0241] FIG. 17a shows a simplified schematic of a DMM 360 used as a wavefront modulator (see FIG. 16). When the DMM 360 is flat, i.e. with no applied voltage (shown on the left), it reflects a collimated beam 362. This corresponds to infinite pixel depth. FIG. 17b shows the DMM 360 deformed with an applied voltage. The deformed DMM now reflects a converging beam 364 which becomes a diverging beam 368 beyond the focal point 366. This corresponds to a particular finite pixel depth.

[0242] FIG. 18a shows a simplified schematic of a variable focus liquid lens 370 used as a wavefront modulator (and as part of the beam expander). The lens is at rest with no applied voltage and produces a converging beam 364 which is collimated by the second lens 372. FIG. 18b shows the lens 370 deformed by an applied voltage so that it produces a more converging beam 364 which is only partially collimated by the second lens 372 to still produce a diverging beam 368. A similar configuration can be used with a variable focus liquid mirror instead of a liquid lens.

[0243] Referring again to FIG. 16, a horizontal scanner 374 scans the beam in a horizontal direction, while a subsequent vertical scanner 376 scans the beam in a vertical direction. Together they steer the beam in a two-dimensional raster pattern. The horizontal scanner 374 operates at the pixel rate of the VRD, while the vertical scanner operates at the line rate. To prevent possible beating between the frame rate and the frequency of microsaccades, which are of the same order, it is useful for the pixel-rate scan to occur horizontally with respect to the eye, since many detail-oriented microsaccades, such as occur during reading, are horizontal.

[0244] The horizontal scanner may utilise a resonant scanning mirror, as described in [37]. Alternatively, it may utilise an acousto-optic deflector, as described in [27,28], or any other suitable pixel-rate scanner, replicated as necessary to achieve the desired pixel rate.

[0245] Although FIG. 16 shows distinct horizontal and vertical scanners, the two scanners may be combined in a single device such as a biaxial MEMS scanner, as described in [37].

[0246] Similarly, FIG. 16 shows the video generator 328 producing video timing signals 378 and 380, it may be convenient to derive video timing from the operation of the horizontal scanner 374 if it utilises a resonant design, since a resonant scanner's frequency is determined mechanically. Furthermore, since a resonant scanner generates a sinusoidal scan velocity, it is crucial to vary pixel durations accordingly to ensure that their spatial extent is constant [54].

[0247] An optional eye tracker 382 determines the approximate gaze direction 384 of the eye 326. It may image the eye to detect the position of the pupil as well as the position of the corneal reflection of an infrared lightsource, to determine the approximate gaze direction. Typical corneal reflection eye tracking systems are described in [20,34].

[0248] Eye tracking in general is discussed in [23].

[0249] Multiple off-axis light sources may be positioned within the HMD, as prefigured in [14]. These can be lit in succession, so that each successive image of the eye contains the reflection of a single light source. The reflection data resulting from multiple successive images can then be

combined to determine gaze direction 384, either analytically or using least squares adjustment, without requiring prior calibration of eye position with respect to the HMD. An image of the infrared corneal reflection of a Netpage coded surface in the user's field of view may also serve as the basis for un-calibrated detection of gaze direction.

[0250] If the gaze direction 384 of both eyes is tracked, then the resultant two fixation points can be averaged to determine the likely true fixation point.

[0251] The tracked gaze direction 384 may be low-pass filtered to suppress fine saccades and microsaccades.

[0252] An optional beam offsetter 386 acts on the gaze direction 384 provided by the eye tracker 382 to align the beam with the pupil of the eye 326. The gaze direction 384 is simultaneously used by a high-level image generator to generate virtual imagery offset correspondingly.

[0253] Projection optics 388 finally project the beam 390 onto the eye 326, magnifying the scan angle to provide the required field of view angle. The projection optics include a visor-shaped optical combiner which simultaneously reflects the generated imagery onto the eye while passing light from the environment. The VRD thereby acts as a see-through display. The visor is ideally curved, so that it magnifies the projected imagery to fill the field of view.

[0254] The HMD as a whole, discussed below, ensures that the projected imagery is registered with a physical Netpage coded surface in the user's field of view. The optical transmission of the combiner may be fixed, or it may be variable in response to active control or ambient light levels. For example, it may incorporate a liquid-crystal layer switchable between transmissive and opaque states, either under user or software control. Alternatively or additionally, it may incorporate a photochromic material whose opacity is a function of ambient light levels.

[0255] The HMD correctly renders occlusions as part of any displayed virtual imagery, according to the user's current viewpoint relative to a tagged surface. It does not, however, intrinsically support occlusion parallax according to the position of the user's eye relative to the HMD unless it uses eye tracking for this purpose. In the absence of eye tracking, the HMD renders each VRD view according to a nominal eye position. If the actual eye position deviates from the assumed eye position, then the wavefront display nature of the VRD prevents misregistration between the real world and the virtual imagery, but in the presence of occlusions due to real or virtual objects, it may lead to object overlap or holes.

[0256] Referring to FIG. 19, the VRD can be further augmented with a spatial light (amplitude) modulator (SLM) such as a digital micromirror device (DMD) [32, 48] to support occlusion parallax. The SLM 392 is introduced immediately after the wavefront modulator 356 and before the raster scanner 374, 376. Alternatively, the SLM 392 is introduced immediately before the wavefront modulator (but after its beam expander). The video generator 328 provides the SLM 392 with an occlusion map 394 associated with the current pixel. The SLM passes non-occluded parts of the wavefront but blocks occluded parts. The amplitude-modulation capability of the SLM may be multi-level, and each map entry in the occlusion map may be correspondingly multi-level. However, in the limit case the SLM is a

binary device, i.e. either passing light or blocking light, and the occlusion map is similarly binary.

[0257] To prevent holes appearing when a nominally invisible part of the virtual scene becomes visible due to eye movement, the HMD can make multiple passes to display multiple depth planes in the virtual scene. The HMD can either render and display each depth plane in its entirety, or can render and display only enough of each depth plane to support the maximum eye movement possible.

[0258] FIG. 20 shows the wavefront display of FIG. 14 augmented with support for displaying an occlusion 396.

[0259] FIG. 21 shows the DMM 360 of FIGS. 17a and 17b augmented with a DMD SLM 392 to produce a VRD with occlusion support. The "shadow" 398 of the virtual occlusion is a gap formed in the cross-section of the beam reflected by the DMD 360 by the SLM 392.

[0260] Per-pixel occlusion maps are easily calculated during rendering of a virtual model. They may also be derived directly from a depth image. Where the occluding object is an object in the real world, such as the user's hand (as discussed further below), it may be represented as an opaque black virtual object during rendering.

[0261] Table 5 gives examples of the viewing angle associated with common media at various viewing distances. In the table, specified values are shown shaded, while derived values are shown un-shaded. For print media, various common viewing distances are specified and corresponding viewing angles are derived. Required VRD image sizes are then derived based representing a maximum feature frequency of 30 cycles per degree. For display media, various common viewing angles are specified and corresponding viewing angles (and maximum feature frequencies) are derived. For both media types the corresponding surface resolution is also shown.

[0262] Based on their native resolution and human visual acuity, display media such as HDTV video monitors are suited to a viewing angle of between 30 and 40 degrees. This is consistent with viewing recommendations for such display media. Based on their native size and human accommodation limits, print media such as US Letter pages are also suited to a viewing angle of 30 to 40 degrees.

[0263] A VRD image size of around 2000 pixels by 2000 pixels is therefore adequate for virtualising these media. Significantly less is required if knowledge of gaze direction is used to project non-foveated parts of the image at lower resolution.

TABLE 5

Viewing parameters for different media					
format	viewing distance (cm)	viewing angle (deg)	max. freq. (cyc/deg)	VRD size (pixels)	pixels per inch
US Letter page (portrait, 8.5" wide)	20	57	30	3420	402
	30	40		2400	282
	40	30		1800	212
US Letter page (landscape, 11" wide)	50	24		1440	169
	20	70		4200	382
	30	50		3000	273
	40	39		2340	213
	50	31		1860	169

TABLE 5-continued

Viewing parameters for different media					
format	viewing distance (cm)	viewing angle (deg)	max. freq. (cyc/deg)	VRD size (pixels)	pixels per inch
cinema screen (Panavision 2.35:1)	2.5 ^b	50	30	3000	1277 ⁹
	3.2 ^a	40 ^c		2400	1021 ^b
	4.4 ^a	30 ^d		1800	766 ^b
32" diag. video monitor (16:9 HDTV, 1920 wide)	76	50	19	1920	69
	97	40 ¹⁰	24		
21" diag. computer monitor (4:3 XVGA, 1600 wide)	132	30 ¹¹	32		
	46	50	16	1600	95
	59	40 ^c	20		
	80	30 ^d	27		

⁸In units of screen height

⁹Per unit of screen height

¹⁰THX recommends 36 degrees in back row of theatre

¹¹SMPTE EG-18-1994 recommends 30 degrees viewing angle

[0264] FIG. 22 shows a block diagram of a Netpage HMD 300 incorporating dual VRDs 304 and 306 for binocular stereoscopic display as shown in FIG. 14. Dual earphones 800 and 802 provide stereophonic sound. Although dual VRDs are preferred, a single VRD providing a monoscopic display capability also has utility (see FIG. 13). Similarly, a single earphone also has utility.

[0265] Although VRDs or similar display devices are preferred for incorporation in the Netpage HMD because they allow the incorporation of wavefront curvature modulation, more conventional display devices such as liquid crystal displays may also be utilised, but with the added complexity of requiring more careful head and eye position calibration or tracking. Conventional LCD-based HMDs are described in detail in [45].

[0266] To maximise the operating range of the VRDs with respect to eye movement, and to maximise user comfort, the optical axes of the VRDs can be approximately aligned with the resting positions of the two eyes by adjusting the lateral separation of the VRDs and adjusting the tilt of the visor. This can be achieved as part of a fitting process and/or performed manually by the user at any time. Note again that the wavefront display capability of the VRDs means that these adjustments are not required to achieve registration of virtual imagery with the physical world.

[0267] A Netpage sensor 804 acquires images 806 of a Netpage coded surface in the user's field of view. It may have a fixed viewing direction and a relatively narrow field of view (of the order of the minimum field of view required to acquire and decode a tag); a variable viewing direction and a relatively narrow field of view; or a fixed viewing direction and a relatively wide field of view (of the order of the VRD viewing angle or even greater). In the first case, the user is constrained to interacting with a Netpage coded surface in the fixed and narrow field of view of the sensor, requiring the head to be turned to face the Netpage of interest. In the second case, the gaze-tracked fixation point can be used to steer the image sensor's field of view, for example via a tip-tilt mirror, allowing the user to interact with a Netpage by fixating on it. In the third case, the gaze-tracked fixation point can be used to select a sub-region

of the sensor's field of view, again allowing the user to interact with a Netpage by fixating on it. In the second and third cases, and as described earlier, the user's effective viewing angle is widened by using the tracked gaze direction to offset the beam.

[0268] A controlling HMD processor **808** accepts image data **330** from the Netpage sensor **804**. The processor locates and decodes the tags in the image data to generate a continuous stream of identification, position and orientation information for the Netpage being imaged. A suitable Netpage image sensor with an on-board image processor, and the corresponding image processing algorithm, tag decoding algorithm and pose (position and orientation) estimation algorithm, are described in [9,59]. In the HMD **300**, the image sensor resolution is higher than described in [9] to support a greater range of tag pattern scales. The sensor utilises a small aperture to ensure good depth of field, and an objective lens system for focusing, approximately as described in [4].

[0269] The Netpage sensor **804** incorporates a longpass or bandpass infrared filter matched to the absorption peak of the infrared ink used to encode the HMD-oriented Netpage tag pattern. It also includes a source of infrared illumination matched to the ink. Alternatively it relies on the infrared component of ambient illumination to adequately illuminate the tag pattern for imaging purposes. In addition, large and/or distant SVDs (such as cinema screens, billboards, and even video monitors) are usefully self-illuminating, either via front or back illumination, to avoid reliance on HMD illumination.

[0270] Alternatively or additionally to determining the actual viewing distance of the tagged surface by analysing the scale and perspective distortion of the tagged pattern images **806**, the Netpage sensor **804** may include an optical range finder. Time-of-flight measurement of an encoded optical pulse train is a well-established technique for optical range finding, and a suitable system is described in [17].

[0271] The depth determined via the optical range finder can be used by the HMD to estimate the expected scale of the imaged tag pattern, thus making tag image processing more efficient, and it can be used to fix the z depth parameter during pose estimation, making the pose estimation process more efficient and/or accurate. It can also be used to adjust the focus of Netpage sensor's optics, to provide greater effective depth of field, and can be used to change the zoom of the Netpage sensor's optics, to allow a smaller image sensor to be utilised across a range of viewing distances, and to reduce the image processing burden.

[0272] Zoom and/or focus control may be effected by moving a lens element, as well as by modulating the curvature of a deformable membrane mirror [43,51], a liquid-crystal phase corrector [47], or other suitable device. Zoom may also be effected digitally, e.g. simply to reduce the image processing burden.

[0273] Range-finding, whether based on pose estimation or time-of-flight measurement, can be performed at multiple locations on a surface to provide an estimate of surface curvature. The available range data can be interpolated to provide range data across the entire surface, and the virtual imagery can be projected onto the resultant curved surface. The geometry of a tagged curved surface may also be known a priori, allowing proper projection without additional range-finding.

[0274] Rather than utilising a two-dimensional image sensor, the Netpage sensor **804** may instead utilise a scanning laser, as described in [5]. Since the image produced by the scanning laser is not distorted by perspective, pose estimation cannot be used to yield the z depth of the tagged surface. Optical (or other) range finding is therefore crucial in this case. Pose estimation may still be performed to determine three-dimensional orientation and two-dimensional position. The optical range finder may be integrated with the laser scanner, utilising the same laser source and photodetector, and operating in multiplexed fashion with respect to scanning.

[0275] The frame rate of the Netpage sensor **804** is matched to the frame rate of the image generator **328** (e.g. at least 50 Hz, but ideally 100 Hz or more), so that the displayed image is always synchronised with the position and orientation of the tagged surface. Decoding of the page identifier embedded in the surface coding can occur at a lower rate, since it changes much less often than position. Decoding of the page identifier can be triggered when a tag pattern is re-acquired, and when the decoded position changes significantly. Alternatively, if the least significant bits of the page identifier are encoded in the same codewords which encode position, then full page identifier decoding can be triggered by a change in the least significant page identifier bits.

[0276] The imaging axis of the Netpage sensor emerges from the HMD **300** between and slightly above the eyes, and is roughly normal to the face. Alternatively, the Netpage sensor **804** is arranged to image the back of the visor, so that its imaging axis roughly coincides with one eye's resting optical axis.

[0277] Although the HMD **300** incorporates a single Netpage sensor **804**, it may alternatively incorporate dual Netpage sensors and be configured to perform pose estimation across both image sensor's acquired images. It may also incorporate multiple tag sensors to allow tag acquisition across a wider field of view.

[0278] Various scenarios for connecting the HMD **300** to a Netpage server **812** are illustrated in FIG. 23, FIG. 24 and FIG. 25.

[0279] A radio transceiver **810** (see FIG. 22) provides a communications interface to a server such as a video server or a Netpage server **812**. The architecture of the overall Netpage system with which the Netpage HMD **300** communicates is described in [1, 3].

[0280] The radio interface **810** may utilise any of a number of protocols and standards, including personal-area and local-area standards such as Bluetooth, IEEE 802.11, 802.15, and so on; and wide-area mobile standards such as GSM, TDMA, CDMA, GPRS, etc. It may also utilise different standards for outgoing and incoming communication, for example utilising a broadcast standard for incoming data, such as a satellite, terrestrial analogue or terrestrial digital standard.

[0281] The HMD **300** may effect communication with a server **812** in a multi-hop fashion, for example using a personal-area or local-area connection to communicate with a relay device **816** which in turn communicates with a server via communications network **814** for a longer-range connection. It may also utilise multiple layers of protocols, for

example communicating with the server via TCP/IP overlaid on a point-to-point Bluetooth connection to a relay as well as on the broader Internet.

[0282] Alternatively or additionally, the HMD may utilise a wired connection to a relay or server, utilising one or more of a serial, parallel, USB, Ethernet, Firewire, analog video, and digital video standard.

[0283] The relay device 816 may, for example, be a mobile phone, personal digital assistant or a personal computer. The HMD may itself act as a relay for other Netpage devices, such as a Netpage pen [4], or vice versa.

[0284] In the Netpage architecture, the identifier of a Netpage is used to identify a corresponding server which is able to provide information about the page and handle interactions with the page. When the HMD first encounters a new page identifier, it looks up a corresponding server, for example via the DNS. Having identified a server, it retrieves static and/or dynamic data associated with the page from the server. Having retrieved the page data, an image generator 328 renders the page data stereoscopically for the two eyes according to the position and orientation of the Netpage with respect to the HMD, and optionally according to the gaze directions of the eyes. The generated stereo images include per-pixel depth information which is used by the VRDs 304 and 306 to modulate wavefront curvature (see FIG. 22).

[0285] Static page data may include static images, text, line art and the like. Dynamic page data may include video 822, audio 824, and the like.

[0286] A sound generator 820 renders the corresponding audio, if any, optionally spatialised according to the relative positions of the HMD and the coded surface, and/or the virtual position(s) of the sound source(s) relative to the coded surface. Suitable audio spatialisation techniques are described in [41].

[0287] The HMD may download dynamic data such as video and audio into a local memory or disk device, or it may obtain such data in streaming fashion from the server, with some degree of local buffering to decouple the local playback rate from any variations in streaming rate due to network behaviour.

[0288] Whether the image data is static or dynamic, the image generator 328 constantly re-renders the page data to take into account the current position and orientation of the Netpage with respect to the HMD 300 (and optionally according to gaze direction).

[0289] The frame rate of the image generator 328 and the VRDs 304, 306 is at least the critical fusion frequency and is ideally faster. The frame rate of the image generator and the VRDs may be different from the frame rate of a video stream being displayed by the HMD 808. Ideally the image generator utilises motion estimation to generate intermediate frames not explicitly present in the video stream. Applicable techniques are described in [21, 39]. If the video stream utilises a motion-based encoding scheme such as an MPEG variant, then the HMD uses the motion information inherent in the encoding to generate intermediate frames.

[0290] As an alternative to the image generator in the HMD performing full page image rendering, the server may perform page image rendering and transmit a corresponding video sequence to the HMD. Because of the latency between

pose estimation, image rendering and subsequent display in this scenario, it is advantageous to still transform the resultant video stream according to pose in the HMD at the display frame rate.

[0291] More generally, whether image generation occurs on the server or in the HMD, a dedicated image warper 826 can be utilised to perspective-project the video stream according to the current pose, and to generate image data at a rate and at a resolution appropriate to the display, independent of the rate and resolution of the image data generated by the image generator 328. This is illustrated in FIG. 26.

[0292] Multi-pass perspective projection techniques are described in [58]. Single-pass techniques and systems are described in [31, 2]. General techniques based on three-dimensional texture mapping are described in [13]. Transforming an input image to produce a perspective-projected output image involves low-pass filtering and sampling the input image according to the projection of each output pixel into the space of the input image, i.e. computing the weighted sum of input pixels which contribute to each output pixel. In most hardware implementations, such as described in [22], this is efficiently achieved by trilinearly interpolating an image pyramid which represents the input image at multiple resolutions. The image pyramid is often represented by a mipmap structure [57], which contains all power-of-two image resolutions. A mipmap only directly supports isotropic low-pass filtering, which leads to a compromise between aliasing and blurring in areas where the projection is anisotropic. However, anisotropic filtering is commonly implemented using mipmap interpolation by computing the weighted sum of several mipmap samples.

[0293] In general, image generation for or in the HMD can make effective use of multi-resolution image formats such as the wavelet-based JPEG2000 image format, as well as mixed-resolution formats such as Mixed Raster Content (MRC), which treats line art and text differently to contone image data, and which is also incorporated in JPEG2000.

[0294] If there is noticeable latency between initial acquisition of a surface by the HMD, and subsequent display of virtual imagery associated with that surface, then the HMD can signal acquisition of the surface to the user to provide immediate feedback. For example, the HMD can highlight or outline the surface. This also serves to distinguish Netpage tagged surfaces from un-tagged surfaces in the user's field of view. The tags themselves can contain an indication of the extent of the surface, to allow the HMD to highlight or outline the surface without interaction with a server. Alternatively, the HMD can retrieve and display extent information from the server in parallel with retrieving full imagery.

[0295] The HMD may be split into a head-mounted unit and a control unit (not shown) which may, for example, be worn on a belt or other harness. If the beam generators are compact, then the head-mounted unit may house the entire VRDs 304 and 306. Alternatively, the control unit may house the beam generators and modulators, and the combined beams may be transmitted to the head-mounted unit via optic fibers.

[0296] As described earlier, the user may utilise gaze to move a cursor within the field of view and/or to virtually

“select” an object. For example, the object may represent a virtual control button or a hyperlink. The HMD can incorporate an activation button, or “clicker” **828**, as shown in **FIG. 27**, to allow the user to activate the currently selected object. The clicker **828** can consist of a simple switch, and may be mounted in any of a number of convenient locations. For example, it may be incorporated in a belt-mounted control unit, or it may be mounted on the index finger for activation by the thumb. Multiple activation buttons can also be provided, analogously to the multiple buttons on a computer mouse.

[0297] Gaze-directed cursor movement can be particularly effective because the precision of the movement of the cursor relative to a surface can be increased by simply bringing the surface closer to the eye.

[0298] In the absence of precise gaze tracking, the user may move their head to move a cursor and/or select an object, based simply on the optical axis of the HMD itself

[0299] The HMD can also provide cursor navigation buttons **830** and/or a joystick **832** to allow the user to move a cursor without utilising gaze. In this case the cursor is ideally tied to the currently active tagged surface, so that the cursor appears attached to the surface when relative movement between the HMD and the surface occurs. The cursor can be programmed to move at a surface-dependent rate or a view-dependent rate or a compromise between the two, to give the user maximum control of the cursor.

[0300] The HMD can also incorporate a brain-wave monitor **834** to allow the user to move the cursor, select an object and/or activate the object by thought alone [60].

[0301] The HMD can provide a number of dedicated control buttons **836**, e.g. for changing the cursor mode (e.g. between gaze-directed, manually controlled, or none), as well as for other control functions.

[0302] It is sometimes useful to dissociate a SVD from the physical surface to which it is attached. The HMD can therefore provide a control button **836** which allows the user to “lift” an SVD from a surface and place it at a fixed location and in a fixed orientation relative to the HMD field of view. The user may also be able to move the lifted SVD, zoom in and zoom out etc., using virtual or dedicated control buttons. The user may also benefit from zooming the SVD in situ, i.e. without lifting it, for example to improve readability without reducing the viewing distance.

[0303] Referring back to **FIG. 22**, the HMD can include a microphone **838** for capturing ambient audio or voice input **840** from the user, and a still or video camera for capturing still or moving images **844** of the user’s field of view. All captured audio, image and video input can be buffered indefinitely by the HMD as well as streamed to a Netpage or other server **812** (**FIGS. 23, 24** and **25**) for permanent storage. Audio and video recording can also operate continuously with a fixed-size circular buffer, allowing the user to always replay recent events without having to explicitly record them.

[0304] The still or video camera **842** can be in line with the HMD’s viewing optics, allowing the user to capture essentially what they see. The camera can also be stereoscopic. In a simpler configuration, a single camera is mounted centrally and has an imaging axis parallel to the viewing axes.

In a more sophisticated configuration, using appropriate beam-steering optics coupled with the gaze tracking mechanism, the camera can follow the user’s gaze. The camera ideally provides automatic focus, but provides the user with zoom control. Multiple cameras pointing in different directions can also be deployed to provide panoramic or rear-facing capture. Direct imaging of the cornea can also capture a wide-angle view of the world from the user’s point of view [49].

[0305] If the camera is placed in line with the viewing optics, then the corresponding beam combiner can be an LCD shutter, which can be closed during exposure to allow the optical path to be dedicated to the camera during exposure. If the camera is a video camera, then display and capture can be suitably multiplexed, although with a concomitant loss of ambient light unless the exposure time is short.

[0306] If the HMD incorporates a video camera, then the Netpage sensor can be configured to use it. If the HMD incorporates a corneal imaging video camera, then it can be utilized by the gaze-tracking system as well as the Netpage sensor.

[0307] Audio and video control buttons, for settings as well as for recording and playback, can be provided by the HMD virtually or physically.

[0308] Binocular disparity between the images captured by a stereo camera can be used by the HMD to detect foreground objects, such as the user’s hand or coffee cup, occluding the Netpage surface of interest. It can use this to suppress rendering and/or projection of the SVD where it is occluded. The HMD can also detect occlusions by analysing the entire visible tagging of the Netpage surface of interest.

[0309] An icon representing a captured image or video clip can be projected by the HMD into the user’s field of view, and the user can select and operate on it via its icon. For example, the user can “paste” it onto a tagged physical surface, such as a page in a Netpage notebook. The image or clip then becomes permanently associated with that location on the surface, as recorded by the Netpage server, and is always shown at that location when viewed by an authorized user through the HMD. Arbitrary virtual objects, such as electronic documents, programs, etc., can be attached to a Netpage surface in a similar way.

[0310] The source of an image or video clip can also be a separate camera device associated with the user, rather than a camera integrated with the HMD.

[0311] The HMD’s microphone **838** and earphones **800, 802** allow it to conveniently support telephony functions, whether over a local connection such as Bluetooth or IEEE 802.11, or via a longer-range connection such as GSM or CDMA. Voice may be carried via dedicated voice channels, and/or over IP (VoIP). Telephony control functions, such as dialling, answer and hangup, may be provided by the HMD via virtual or physical buttons, may be provided by a separate physical device associated with the HMD or more loosely with the user, or may be provided by a virtual interface tied to a physical surface [7].

[0312] The HMD’s earphones allow it to support music playback, as described in [8]. Audio can be copied or

streamed from a server, or played back directly from a storage device in the HMD itself

[0313] The HMD ideally incorporates a unique identifier which is registered to a specific user. This controls what the wearer of the HMD is authorized to see.

[0314] The HMD can incorporate a biometric sensor, as shown in FIG. 28, to allow the system to verify the identity of the wearer. For example, the biometric sensor may be a fingerprint sensor 846 incorporated in a belt-mounted control unit, or it may be a iris scanner 848 incorporated in either or both the displays 304, 306 (see FIG. 22), possibly integrated with the gaze tracker 382 (see FIG. 16).

[0315] The HMD can include optics to correct for deficiencies in a user's vision, such as myopia, hyperopia, astigmatism, and presbyopia, as well as non-conventional refractive errors such as aberrations, irregular astigmatism, and ocular layer irregularities. The HMD can incorporate fixed prescription optics, e.g. integrated into the beam-combining visor, or adaptive optics to measure and correct deficiencies on a continuous basis [18,56].

[0316] The HMD can incorporate an accelerometer so that the acceleration vector due to gravity can be detected. This can be used to project a three-dimensional image properly if desired. For example, during remote conferencing it may be desirable to always render talking heads the right way up, independently of the orientation of the surfaces to which they are attached. As a side-effect, such projections will lean if centripetal acceleration is detected, such as when turning a corner in a car.

[0317] The HMD incorporates a battery, recharged by removal and insertion into a battery charger, or by direct connection between the charger and the HMD. The HMD may also conveniently derive recharging power on a continuous basis from an item of clothing which incorporates a flexible solar cell [53]. The item may also be in the shape of a cap or hat worn on the head, and the HMD may be integrated with the cap or hat.

Surface Coding

[0318] The scale of the HMD-oriented Netpage tag pattern disposed on a particular medium is matched to the minimum viewing distance expected for that medium. The tag pattern is designed to allow the Netpage sensor in the HMD to acquire and decode an entire tag at the minimum supported viewing distance. The pixel resolution of the Netpage image sensor then determines the maximum supported viewing distance for that medium. The greater the supported maximum viewing distance, the smaller the tag pattern projected on the image sensor, and the greater the image sensor resolution required to guarantee adequate sampling of the tag pattern. Surface tilt also increases the feature frequency of the imaged tag pattern, so the maximum supported surface tilt must also be accommodated in the selected image sensor resolution.

[0319] The basis for a suitable Netpage tag pattern is described in [6]. The hexagonal tag pattern described in the reference requires a sampling field of view with a diameter of 36 features. This requires an image sensor with a resolution of at least 72×72 pixels, assuming minimal two-times sampling. By way of example, assuming arbitrarily that the Netpage sensor in the HMD has an angular field of view of

10 degrees, and assuming the minimum supported viewing distance for a hand-held printed page is 30 cm, an appropriate HMD-oriented Netpage tag pattern has a scale of about 1.5 mm per feature (i.e. $30 \text{ cm} \times \tan(5)/(36/2)$). Further assuming the maximum supported viewing distance is 120 cm (i.e. 4×30 cm), the required image sensor resolution is 288×288 pixels (i.e. 4×72). Greater image sensor resolution allows for a greater range of viewing distances. By comparison, assuming the minimum supported viewing distance for a large-screen "HDTV" Netpage is 2 m, an appropriate HMD-oriented Netpage tag pattern has a scale of about 1 cm per feature (i.e. $2 \text{ m} \times \tan(5)/(36/2)$), and the same image sensor supports a maximum viewing distance of 8 m (i.e. $4 \times 2 \text{ m}$). By way of further comparison, assuming the minimum supported viewing distance for a billboard Netpage mounted on the side of a building is 30m, an appropriate HMD-oriented Netpage tag pattern has a scale of about 15 cm per feature (i.e. $30 \text{ m} \times \tan(5)/(36/2)$), and the same image sensor supports a maximum viewing distance of 120m (i.e. $4 \times 30 \text{ m}$).

[0320] Although it is useful for particular media types to utilise a consistent tag pattern scale, it is also possible for individual users to select a tag pattern scale suited to their particular viewing preferences. This is particularly convenient when the Netpages in question are printed on demand.

[0321] It is useful to encode the scale of a tag pattern in the data encoded in the pattern, so that a decoding device such as the Netpage HMD can determine the scale and hence the absolute viewing distance without reference to associated information. However, if it is not convenient to encode a scale factor in the tag data, then the scale factor can be recorded by the corresponding Netpage server, either per page instance or per page type. The HMD then obtains the scale factor from the server once it has identified the page. In general, the server records the scale factor as well as an affine transform which relates the coordinate system of the tag pattern to the coordinate system of the physical page.

[0322] As described earlier, if a Netpage surface also supports pen interaction, then it may be coded with two sets of tags utilising different infrared inks, one set of tags printed at a pen-oriented scale, and the other set of tags printed at a HMD-oriented scale, as discussed above. Alternatively the surface may be coded with multi-resolution tags which can be imaged and decoded at multiple scales. In another option, the HMD tag sensor is capable of acquiring and decoding pen-scale tags, then a single set of tags is sufficient. A laser scanning Netpage sensor is capable of acquiring pen-scale tags at normal viewing distances such as 30 cm to 120 cm.

[0323] Since the virtual imagery displayed by the HMD is effectively added to the user's view of the real world, the physical Netpage surface region onto which the imagery is virtually projected is ideally printed black. It is impractical to selectively change the opacity of the HMD visor, since the beam associated with a single pixel may cover the entire exit pupil of the VRD, depending on its depth.

[0324] Tags are ideally disposed on a surface invisibly, e.g. by being printed using an infrared ink. However, visible tags may be utilised where invisibility is impractical. Although printing is an effective mechanism for disposing tags on a surface, tags may also be manufactured on or into a surface, such as via embossing. Although inkjet printing is

an effective printing mechanism, other printing mechanisms may also be usefully employed, such as laser printing, dye sublimation, thermal transfer, lithography, offset, gravure, etc.

[0325] Neither pen-oriented nor HMD-oriented Netpage tags are limited in their application to surfaces traditionally associated with publications, displays and computer interfaces. For example, tags can also be applied to skin in the form of temporary or permanent tattoos; they can be printed on or woven into textiles and fabric; and in general they can be applied to any physical surface where they have utility. HMD-oriented tags, because of their intrinsically larger scale, are more easily applied to a wide range of surfaces than pen-oriented tags.

Applications

[0326] FIG. 29 shows a mockup of a printed page 850 containing a typical arrangement of text 858, graphics and images 842. The page 850 also includes two invisible tag patterns 854 and 856. One tag pattern 854 is scaled for close-range imaging by a Netpage stylus or pen or other device typically in contact with or in close proximity to the page 850. The other tag pattern 856 is scaled for longer-range imaging by a Netpage HMD. Either tag pattern may be optional on any given page.

[0327] FIG. 30 shows the page 850 of FIG. 29 augmented with a virtual embedded video clip 860 when viewed through the Netpage HMD, i.e. the video clip 860 is a dedicated situated virtual display (SVD) on the page. The video clip appears with playback controls 862. A playback control buttons can be activated using a Netpage stylus or pen 8 (see FIG. 31). Alternatively a control button can be selected and activated via the HMD's clicker as described earlier. The control buttons 862 can also be printed on the page 850. Alternatively still, a generic Netpage remote control may be utilised in conjunction with the Netpage HMD. The remote control may provide generic media playback control buttons, such as play, pause, stop, rewind, skip forwards, skip backwards, volume control, etc. The Netpage system can interpret playback control commands received from a Netpage remote control associated with a user as pertaining to the user's currently selected media object (e.g. video clip 860).

[0328] The video clip 860 is just one example of the use of an SVD to augment a document. In general, an arbitrary interactive application with a graphical user interface can make use of an SVD in the same manner.

[0329] FIG. 31 shows a four-function calculator application 864 embedded in a page 850, with the page augmented with a virtual display 866 for the calculator. The input buttons 868 for the calculator are printed on the page, but could also be displayed virtually.

[0330] FIG. 32 shows a page 850 augmented with a display 870 for confidential information only intended for the user.

[0331] As described earlier, apart from registration of the HMD as belonging to the user, the HMD may verify user identify via a biometric measurement. Alternatively, the user may be required to provide a password before the HMD will display restricted information.

[0332] FIG. 33 shows the page 850 of FIG. 29 augmented with virtual digital ink 9 drawn using a non-marking Netpage stylus or pen 8. Virtual digital ink has the advantage that it can be virtually styled, e.g. with stroke width, colour, texture, opacity, calligraphic nib orientation, or artistic style such as airbrush, charcoal, pencil, pen, etc. It also has the advantage that it is only seen by authorized users via their HMDs (or via Netpage browsers).

[0333] If all "pen" input is virtual, then multiple physical instances of the same logical Netpage page instance can be printed and used as a basis for remote collaboration or conferencing. Any digital ink 9 drawn virtually by one authorized user instantaneously appears "on" the other instances of the page 850 when viewed by other authorized users.

[0334] Even on different logical instances of a page a subregion can be mapped to a shared "whiteboard" for remote collaboration and conferencing purposes.

[0335] Physical and virtual digital ink can also co-exist on the same physical page.

[0336] Whether Netpage pen input actually marks the page or is only displayed virtually, and whether pen input is created relative to page content printed physically or displayed virtually, the pen input is captured by the Netpage system as digital ink and is interpreted in the context of the corresponding page description. This can include interpreting it as an annotation, as streaming input to an application, as form input to an application (e.g. handwriting, a drawing, a signature, or a checkmark), or as control input to an application (e.g. a form submission, a hyperlink activation, or a button press) [3].

[0337] FIG. 34 shows another version of the page 850 of FIG. 29, where even the static page content 858 and 852 is virtual and is only seen via the Netpage HMD (or the Netpage browser). In this case the entire page can be thought of as a dedicated SVD for the static and dynamic content of the page. Only the tag pattern(s) 854, 856 exist on the physical page, and the virtual content is associated with the page, possibly by "printing" onto the page by passing it through a virtual "printer" device. The virtual Netpage printer simply determines the page ID of each page which passes through it and associates it with the next document page. The association between page ID and page content is still recorded by the Netpage server in the usual way.

[0338] Physical pages can be manufactured from durable plastic and can be tagged during manufacture rather than being tagged on demand. They can be re-used repeatedly. New content can be "printed" onto a page by passing it through a virtual Netpage printer. Content can be wiped from a page by passing it through a virtual Netpage shredder. Content can also be erased using various forms of Netpage erasers. For example, a Netpage stylus or pen operating in one eraser mode may only be capable of erasing digital ink, while operating in another eraser mode may also be capable of erasing page content.

[0339] Fully virtualising page content has the added advantage that pages can be viewed and read in ambient darkness.

[0340] Although not shown in the figures, regions which are augmented with virtual content (such as video clips and

the like) are ideally printed in black. Since the output of the Netpage HMD is added to the page, it is ideally added to black to create color and white. It cannot be used to subtract color from white to create black. In regions where black is impractical, such as when annotating physical page content with virtual digital ink, the brightness of the HMD output is sufficiently high to be clearly visible even with a white page in the background.

[0341] If plastic blanks are used and all page content is virtual, then the blanks are also ideally black, and matte to prevent specular reflection of ambient light.

[0342] FIG. 35 shows a mobile phone device 872 incorporating an SVD. Like the document page discussed above, the display surface 874 includes a tag pattern scaled for longer-range imaging by a Netpage HMD 856. It also optionally includes a tag pattern 854 scaled for close-range imaging by a Netpage stylus or pen 8, for “touch-screen” operation.

[0343] The extent of the SVD 876 need not be constrained by the physical size of the device to which it is “attached”. As shown in FIG. 36, the display 876 can protrude laterally beyond the bounds of the device 872.

[0344] The SVD 876 can also be used to virtualise the input functions on the device 872, such as the keypad in this case, as shown in FIG. 37.

[0345] Generally also, the SVD 876 can overlay the conventional display 874 of the device 872, such as an LCD or OLED. The user may then choose to use the built-in display 874 or the SVD 876 according to circumstance.

[0346] Although the examples show a mobile phone device 872, the same approach applies to any portable device incorporating a display and/or a control interface, including a personal digital assistant (PDA), an music player, A/V remote control, calculator, still or video camera, and so on.

[0347] Since, as discussed earlier, the physical surface 874 of an SVD 876 is ideally matte black, it provides an ideal place to incorporate a solar cell into the device 872 for generating power from ambient light.

[0348] FIG. 38 shows an SVD 876 used as a cinema screen 878. Note that the scale of the HMD-oriented tag pattern 856 is much larger than in the cases described above, because on the much larger average viewing distance.

[0349] The movie is virtually projected from a video source 880, either via direct streaming from a video transmitter 882 to the Netpage HMDs of the members of the audience 884, or via a Netpage server 812 and an arbitrary communications network 814.

[0350] Individual delivery of content to each audience member during an otherwise “shared” viewing experience has the advantage that it can allow individual customisation. For example, specific edits can be delivered according to age, culture or other preference; each individual can specify language, subtitle display, audio settings such as volume, picture settings such as brightness, contrast, color and format; and each individual may be provided with personal playback controls such as pause, rewind/replay, skip etc.

[0351] In a public performance scenario, a Netpage-encoded printed ticket can act as a token which gives a HMD

access to the movie. The ticket can be presented in the field of view of the tag sensor in the HMD, and the HMD can present the scanned ticket information to the projection system to gain access.

[0352] FIG. 39 shows an SVD used as a video monitor 886, e.g. to display pre-recorded or live video from any number of sources including a television (TV) receiver 888, video cassette recorder (VCR) 890, digital versatile disc (DVD) player 892, personal video recorder (PVR) 894, cable video receiver/decoder 896, satellite video receiver/decoder 898, Internet/Web interface 900, or personal computer 902. Again note that the scale of the HMD-oriented tag pattern 856 is larger than in the page and personal device cases described above, but smaller than in the cinema case.

[0353] The video switch 906 directs the video signal from one of the video sources (888-902), to the Netpage HMDs 300 of one or more users. The video is delivered via direct streaming from a video transmitter 882 or a Netpage server 812 and an arbitrary communications network 814.

[0354] As in the case of cinema described above, video delivered via an SVD has the advantage can be individually customised.

[0355] FIG. 40 shows an SVD used as a computer monitor 914. The monitor surface includes a tag pattern scaled for imaging by a Netpage HMD 856. It also optionally includes a tag pattern scaled for close-range imaging 854 by a Netpage stylus or pen 8, for “touch-screen” operation. Video output from the personal computer 902 or workstation is delivered either via direct streaming from a video transmitter 882 to the Netpage HMDs 300 of one or more users, or via a Netpage server 812 and an arbitrary communications network 814.

[0356] Another input device 908 is also optionally provided, tagged with a stylus-oriented tag pattern 854. The input device can be used to provide a tablet and/or a virtualised keyboard 910, as well as other functions. Input from the stylus or pen 8 is transmitted to a Netpage server 912 in the usual way, for interpretation and possible forwarding. Although shown separately, the Netpage server 812 may be executing on the personal computer 902.

[0357] Multiple monitors 908 may be used in combination, in various configurations.

[0358] Advertising in public spaces, if virtually displayed, can be targeted according to the demographic of each individual viewer. People may be rewarded for opting in and providing a demographic profile. Virtually displayed advertising can be more finely segmented, both time-wise, according to how much an advertiser is willing to pay, and according to demographic. Targeting can also occur according to time-of-day, day-of-week, season, weather, external event etc.

[0359] If the advertising appears in (or is attached to) a movable object such as a magazine, newspaper, train, bus or taxi poster, or product packaging, then the advertising content can also be targeted according the instantaneous location of the viewer, as indicated by a location device associated with the user, such as a GPS receiver.

[0360] If the HMD incorporates gaze tracking, then gaze direction information can be used to provide statistical information to advertisers on which elements of their adver-

tising is catching the gaze of viewers, i.e. to support so-called “copy testing”. More directly, gaze direction can be used to animate an advertising element when the user’s gaze strikes it.

[0361] The Netpage HMD can be used to search a physical space, such as a cluttered desktop, for a particular document. The user first identifies the desired document to the Netpage system, perhaps by browsing a virtual filing cabinet containing all of the user’s documents. The HMD is then primed to highlight the document if it is detected in the user’s field of view. The Netpage system informs the HMD of the relation between the tags of the desired document and the physical extent of the document, so that the HMD can highlight the outline of the document when detected.

[0362] The user’s virtual filing cabinet can be extended to contain, either actually or by reference, every document or page the user has ever seen, as detected by the Netpage HMD. More specifically, in conjunction with gaze tracking, the system can mark the regions the user has actually looked at. Furthermore, by detecting the distinctive saccades associated with reading, the system can mark, with reasonable certainty, text passages actually read by the user. This can subsequently be used to narrow the context of a content search.

[0363] One of the advantages of the Netpage HMD is that it allows the user to consume and interact with information privately, even when in a public place. However, because each pixel is projected in succession, a snooper can build a simple detection device to collect each pixel in turn from any stray light emitted by the HMD, and re-synchronise it after the fact to regenerate a sequence of images. To combat this, the HMD can emit random stray light at the pixel rate, to swamp any meaningful stray light from the display itself.

[0364] A non-planar three-dimensional object, if unadorned but tagged on some or all of its faces, may act as a proxy for a corresponding adorned object. For example, a prototyping machine may be used to fabricate a scale model of a concept car. Disposing tags on the surface of the prototype then allows color, texture and fine geometric detail to be virtually projected onto the surface of the car when viewed through a Netpage HMD.

[0365] More simply, a pre-manufactured and pre-tagged shape such as a sphere, ellipsoid, cube or parallelepiped of a certain size can be used as a proxy for a more complicated shape. Virtual projection onto its surface can be used to imbue it with apparent geometry, as well as with color, texture and fine geometric detail.

References

[0366] The following references are incorporated herein by cross-reference.

[0367] Lapstun, P. and K. Silverbrook, “Method and System for Printing a Document”, U.S. Pat. No. 6,728,000, issued 27 Apr. 2004

[0368] [2] Silverbrook, K. and P. Lapstun, “Digital Image Warping System”, U.S. Pat. No. 6,636,216, issued 21 Oct. 2003

[0369] [3] see Appendix A

[0370] Silverbrook Research, “Sensing device for coded data”, U.S. Patent Application U.S. Ser. No. 10/815,636 (Docket Number HYJ001), filed 2 Apr. 2004, claiming priority from [9,11,12]

[0371] [5] Silverbrook Research, “Laser scanner device for printed product identification codes”, U.S. Patent Application U.S. Ser. No. 10/815,609 (Docket Number HYT001), filed 2 Apr. 2004, claiming priority from [11, 12]

[0372] [6] Silverbrook Research, “Rotationally symmetric tags”, U.S. Patent Application U.S. Ser. No. 10/309,358, filed 4 Dec. 2002

[0373] Silverbrook Research, “Method and system for telephone control”, U.S. Patent Application U.S. Ser. No. 09/721,895, filed 25 Nov. 2000

[0374] [8] Silverbrook Research, “Viewer with code sensor”, U.S. Patent Application U.S. Ser. No. 09/722,175, filed 25 Nov. 2000

[0375] [9] Silverbrook Research, “Image sensor with digital framestore”, U.S. Patent Application U.S. Ser. No. 10/778,056 (Docket Number NPS047), filed 17 Feb. 2004, claiming priority from [10]

[0376] [10] Silverbrook Research, “Methods, systems and apparatus”, Australian Provisional Patent Application 2003900746 (Docket Number NPS041), filed 17 Feb. 2003

[0377] [11] Silverbrook Research, “Methods and systems for object identification and interaction”, Australian Provisional Patent Application 2003901617 (Docket Number NIR002), filed 7 Apr. 2003

[0378] [12] Silverbrook Research, “Methods and systems for object identification and interaction”, Australian Provisional Patent Application 2003901795 (Docket Number NIR005), filed 15 Apr. 2003

[0379] [13] Akenine-Möller, T. and E. Haines, Real-Time Rendering, Second Edition, A K Peters 2002

[0380] [14] Amir, A., M. D. Flickner, D. B. Koons and C. H. Morimoto, “System and Method for Eye Gaze Tracking Using Corneal Image Mapping”, U.S. Pat. No. 6,659,611, issued 9 Dec. 2003

[0381] [15] Behringer, R., G. Klinker, and D. W. Mizell, eds., Augmented Reality: Placing Artificial Objects in Real Scenes: Proceedings of IWAR ’98, AK Peters 1999

[0382] [16] Berge, B., and J. Peseux, “Lens with variable focus”, U.S. Pat. No. 6,369,954, issued 9 Apr. 2002

[0383] [17] Bloebaum, F., “Method and Apparatus for Determining the Light Transit Time Over a Measurement Path Arranged Between a Measuring Apparatus and a Reflecting Object”, U.S. Pat. No. 5,805,468, issued 9 Sep. 1998

[0384] [18] Blum, R. D., D. P. Dustin, and D. Katzman, “Method for refracting and dispensing electro-active spectacles”, U.S. Pat. No. 6,733,130, issued 11 May 2004

[0385] [19] Cameron, C. D., D. A. Pain, M. Stanley, and C. W. Slinger, “Computational challenges of emerging novel true 3D holographic displays”, Critical Technologies for the Future of Computing, Proceedings of SPIE Vol. 4109, 2000, pp. 129-140

[0386] [20] Cleveland, D., J. H. Cleveland and P. L. Norloff, “Eye Tracking Method and Apparatus”, U.S. Pat. No. 5,231,674, issued 27 Jul. 1993

- [0387] [21] Demos, G. E., "System and Method for Motion Compensation and Frame Rate Conversion", U.S. Pat. No. 6,442,203, issued 27 Aug. 2002
- [0388] [22] Dignam, D. L., "Circuit and method for tri-linear filtering using texels from only one level of detail", U.S. Pat. No. 6,452,603, issued 17 Sep. 2002
- [0389] [23] Duchowski, A. T., *Eye Tracking Methodology, Theory and Practice*, Springer-Verlag 2003
- [0390] [24] Favalora, G. E., J. Napoli, D. M. Hall, R. K. Dorval, M. G. Giovinco, M. J. Richmond, and W. S. Chun, "100 Million-voxel volumetric display", *Cockpit Displays IX: Displays for Defense Applications, Proceedings of SPIE Vol. 4712*, 2002, pp. 300-312
- [0391] [25] Feenstra, B. J., S. Kuiper, S. Stallinga, B. H. W. Hendriks, and R. M. Snoeren, "Variable focus lens", PCT Patent Application WO 03/069380, filed 24 Jan. 2003
- [0392] [26] Fulton, J. T., *Processes in Biological Vision*, <http://www.4colorvision.com>
- [0393] [27] Furness III, T. A., and J. S. Kollin, "Retinal Display Scanning of Image with Plurality of Image Sectors", U.S. Pat. No. 6,639,570, issued 28 Oct. 2003
- [0394] [28] Furness III, T. A., and J. S. Kollin, "Virtual Retinal Display", U.S. Pat. No. 5,467,104, issued 14 Nov. 1995
- [0395] [29] Gerhard, G. J., C. T. Tegreene, and B. Z. Eslam, "Scanned Display with Pinch, Timing, and Distortion Correction", 5 Aug. 1998
- [0396] [30] Gortler, S. J., R. Grzeszczuk, R. Szeliski, and M. F. Cohen, "The Lumigraph", *ACM Computer Graphics Proceedings, Annual Conference Series*, 1996, pp. 43-54
- [0397] [31] Heckbert, P. S., "Survey of Texture Mapping", *IEEE Computer Graphics & Applications* 6(11), pp. 56-67, November 1986
- [0398] [32] Hornbeck, L. J., "Active yoke hidden hinge digital micromirror device", U.S. Pat. No. 5,535,047, issued 9 Jul. 1996
- [0399] [33] Humphreys, G. W., and V. Bruce, *Visual Cognition*, Lawrence Erlbaum Associates, 1989, p. 15
- [0400] [34] Hutchinson, T. E., C. Lankford and P. Shannon, "Eye Gaze Direction Tracker", U.S. Pat. No. 6,152,563, issued 28 November 2000
- [0401] [35] Isaksen, A., L. McMillan, and S. J. Gortler, "Dynamically Reparameterized Light Fields", *ACM Computer Graphics Proceedings, Annual Conference Series*, 2000, pp. 297-306
- [0402] [36] Levoy, M. and P. Hanrahan, "Light Field Rendering", *ACM Computer Graphics Proceedings, Annual Conference Series*, 1996, pp. 31-42
- [0403] [37] Lewis, J. R., H. Urey and B. G. Murray, "Scanned Imaging Apparatus with Switched Feeds", U.S. Pat. No. 6,714,331, issued 30 Mar. 2004
- [0404] [38] Lewis, J. R., and N. Nestorovic, "Personal Display with Vision Tracking", U.S. Pat. No. 6,396,461, issued 28 May 2002
- [0405] [39] Maturi, G. V., V. Bhargava, S. L. Chen, and R.-Y. Wang, "Hybrid Hierarchical/Full-search MPEG Encoder Motion Estimation", U.S. Pat. No. 5,731,850, issued 24 Mar. 1998
- [0406] [40] Matusik, W., and H. Pfister, "3D TV: A Scalable System for Real-Time Acquisition, Transmission, and Autostereoscopic Display of Dynamic Scenes", *ACM Computer Graphics Proceedings, Annual Conference Series*, 2004
- [0407] [41] McGrath, D. S., "Methods and Apparatus for Processing Spatialised Audio", U.S. Pat. No. 6,021,206, issued 1 February 2000
- [0408] [42] McMillan, L. and G. Bishop, "Plenoptic Modeling: An Image-Based Rendering System", *ACM SIGGRAPH 95*, pp. 3946
- [0409] [43] McQuaide, S. C., E. J. Seibel, R. Burstein and T. A. Furness III, "50.4: Three-dimensional virtual retinal display system using a deformable membrane mirror", *SID 02 DIGEST*
- [0410] [44] Meisner, J., W. P. Donnelly, and R. Roosen, "Augmented Reality Technology", U.S. Pat. No. 6,625,299, issued 23 Sep. 2003
- [0411] [45] Melzer, J. E., and K. Moffitt, *Head Mounted Displays: Designing for the User*, McGraw-Hill 1997
- [0412] [46] Miller, G., "Volumetric Hyper-Reality, A Computer Graphics Holy Grail for the 21 st Century?", *Graphics Interface '95*, pp. 56-64
- [0413] [47] Naumov, A. F., and M. Yu. Loktev, "Liquid-crystal adaptive lenses with modal control", *OPTICS-LETTERS*, Vol. 23, No.13, Jul. 1, 1998, pp. 992-994
- [0414] [48] Nayar, S. K., V. Branzoi, and T. E. Boulton, "Programmable Imaging using a Digital Micromirror Array", *Proceedings of the IEEE Computer Society Conference on Computer Vision and Pattern Recognition*, July 2004, pp. 436-443
- [0415] [49] Nishino, K., and S. K. Nayar, "The World in an Eye", *Proceedings of IEEE Conference on Computer Vision and Pattern Recognition*, Washington D.C., June 2004
- [0416] [50] Perlin, K., S. Paxia, and J. S. Kollin, "An Autostereoscopic Display", *ACM Computer Graphics Proceedings, Annual Conference Series*, 2000, pp. 319-326
- [0417] [51] Silverman, N. L., B. T. Schowengerdt, J. P. Kelly, and E. J. Seibel, "58.5L: Late-News Paper: Engineering a Retinal Scanning Laser Display with Integrated Accommodative Depth Cues", *SID 03 DIGEST*, pp. 1538-1541
- [0418] [52] St.-Hilaire, P., M. Lucente, J. D. Sutter, R. Pappu, C. D. Sparrell, and S. A. Benton, "Scaling up the MIT holographic video system", *Fifth International Symposium on Display Holography, Proceedings of SPIE Vol. 2333*, 1992, pp. 374-380
- [0419] [53] Sverdrup, L. H. Jr., N. F. Dessel, and A. Pelkus, "Thin film flexible solar cell", U.S. Pat. No. 6,548,751, issued 15 Apr. 2003

[0420] [54] Urey, H., D. W. Wine, and T. D. Osborn, "Optical performance requirements for MEMS-scanner based microdisplays", Conference on MOEMS and Miniaturized Systems, SPIE Vol. 4178, pp. 176-185, Santa Clara, Calif. (2000)

[0421] [55] Urey, H., "Apparatus and Methods for Generating Multiple Exit-Pupil Images in an Expanded Exit Pupil", U.S. Patent Application 2003/0086173, published 8 May 2003

[0422] [56] Williams, D. R., and J. Liang, "Method and apparatus for improving vision and the resolution of retinal images", U.S. Pat. No. 5,949,521, issued 7 Sep. 1999

[0423] [57] Williams, L., "Pyramidal Parametrics", Computer Graphics (Proc. SIGGRAPH 1983) 17(3), July 1983, pp. 1-11

[0424] [58] Wolberg, G., *Digital Image Warping*, IEEE Computer Society Press, 1988

[0425] [59] Wolf, P. R., and B. A. Dewitt, *Elements of photogrammetry*, 3rd Edition, McGraw-Hill 2000

[0426] [60] Wolpaw, J. R., and D. J. McFarland, "Communication method and system using brain waves for multidimensional control", U.S. Pat. No. 5,638,826, issued 17 Jun. 1997

We claim:

1. An augmented reality device for inserting virtual imagery into a user's view of their physical environment, the device comprising:

a display device through which the user can view the physical environment;

an optical sensing device for sensing at least one surface in the physical environment; and,

a controller for projecting the virtual imagery via the display device; wherein during use,

the controller uses wave front modulation to match the curvature of the wave fronts of light reflected from the display device to the user's eyes with the curvature of the wave fronts of light that would be transmitted through the device display if the virtual imagery were situated at a predetermined position relative to the surface, such that the user sees the virtual imagery at the predetermined position regardless of changes in position of the user's eyes with respect to the see-through display.

2. An augmented reality device according to claim 1 wherein the display device has a see-through display for one of the user's eyes.

3. An augmented reality device according to claim 1 wherein the display device has two see-through displays, one for each of the user's eyes respectively.

4. An augmented reality device according to claim 1 wherein the surface has a pattern of coded data disposed on it, such that the controller uses information from the coded data to identify the virtual imagery to be displayed.

5. An augmented reality device according to claim 1 wherein the display device, the optical sensing device and the controller are adapted to be worn on the user's head.

6. An augmented reality device according to claim 1 wherein the optical sensing device is camera-based and during use, provides identity and position data related to the coded surface to the controller for determining the virtual imagery displayed.

7. An augmented reality device according to claim 1 wherein display device has a virtual retinal display (VRD) for each of the user's eyes, each of the VRD's scans at least one beam of light into a raster pattern and modulates the or each beam to produce spatial variations in the virtual imagery.

8. An augmented reality device according to claim 7 wherein the VRD scans red, green and blue beams of light to produce color pixels in the raster pattern.

9. An augmented reality device according to claim 8 wherein the VRDs present a slightly different image to each of the user's eyes, the slight differences being based on eye separation, and the distance to the predetermined position of the virtual imagery to create a perception of depth via stereopsis.

10. An augmented reality device according to claim 1 wherein the wavefront modulator uses a deformable membrane mirror, liquid crystal phase corrector, a variable focus liquid lens or a variable focus liquid mirror.

11. An augmented reality device according to claim 1 wherein the virtual imagery is a movie, a computer application interface, computer application output, hand drawn strokes, text, images or graphics.

12. An augmented reality device according to claim 1 wherein the display device has pupil trackers to detect an approximate point of fixation of the user's gaze such that a virtual cursor can be projected into the virtual imagery and navigated using gaze direction.

* * * * *