



(51) International Patent Classification:

H04N 5/783 (2006.01) *H04N* 21/422 (2011.01)
H04N 21/4728 (2011.01) *H04N* 21/4223 (2011.01)
H04N 21/4402 (2011.01) *H04N* 21/414 (2011.01)

(21) International Application Number:

PCT/GB2017/051180

(22) International Filing Date:

27 April 2017 (27.04.2017)

(25) Filing Language:

English

(26) Publication Language:

English

(30) Priority Data:

1611911.7 08 July 2016 (08.07.2016) GB

(71) Applicant: **SONY INTERACTIVE ENTERTAINMENT INC.** [JP/JP]; 1-7-1 Konan, Minato-ku, Tokyo 108-0075 (JP).

(71) Applicant (for AL only): **SONY INTERACTIVE ENTERTAINMENT EUROPE LIMITED** [GB/GB]; 10 Great Marlborough Street, London W1F 7LP (GB).

(72) Inventor: **CAMP, Philip Ronald**; c/o Sony Interactive Entertainment Europe Limited, 10 Great Marlborough Street, London W1F 7LP (GB).

(74) Agent: **EALEY, Douglas**; D Young & Co LLP, 120 Holborn, London EC1N 2DY (GB).

(81) Designated States (unless otherwise indicated, for every kind of national protection available): AE, AG, AL, AM, AO, AT, AU, AZ, BA, BB, BG, BH, BN, BR, BW, BY, BZ, CA, CH, CL, CN, CO, CR, CU, CZ, DE, DJ, DK, DM, DO, DZ, EC, EE, EG, ES, FI, GB, GD, GE, GH, GM, GT, HN, HR, HU, ID, IL, IN, IR, IS, JP, KE, KG, KH, KN, KP, KR, KW, KZ, LA, LC, LK, LR, LS, LU, LY, MA, MD, ME, MG, MK, MN, MW, MX, MY, MZ, NA, NG, NI, NO, NZ, OM, PA, PE, PG, PH, PL, PT, QA, RO, RS, RU, RW, SA, SC, SD, SE, SG, SK, SL, SM, ST, SV, SY, TH, TJ, TM, TN, TR, TT, TZ, UA, UG, US, UZ, VC, VN, ZA, ZM, ZW.

(84) Designated States (unless otherwise indicated, for every kind of regional protection available): ARIPO (BW, GH, GM, KE, LR, LS, MW, MZ, NA, RW, SD, SL, ST, SZ, TZ, UG, ZM, ZW), Eurasian (AM, AZ, BY, KG, KZ, RU, TJ, TM), European (AL, AT, BE, BG, CH, CY, CZ, DE, DK, EE, ES, FI, FR, GB, GR, HR, HU, IE, IS, IT, LT, LU, LV, MC, MK, MT, NL, NO, PL, PT, RO, RS, SE, SI, SK, SM, TR), OAPI (BF, BJ, CF, CG, CI, CM, GA, GN, GQ, GW, KM, ML, MR, NE, SN, TD, TG).

Published:

— with international search report (Art. 21(3))

(54) Title: AUGMENTED REALITY SYSTEM AND METHOD

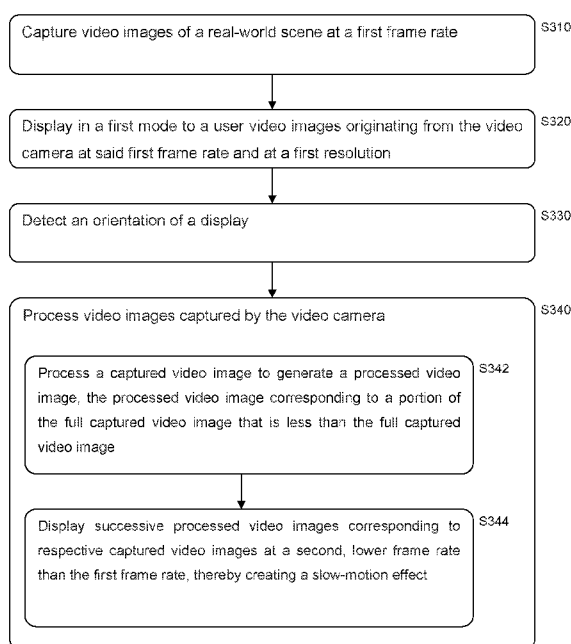


Figure 3

(57) Abstract: An augmented reality system comprises a video input adapted to receive video images of a real-world scene captured by a video camera at a first frame rate; a display operable in a first mode to display to a user video images originating from the video camera at said first frame rate and at a first resolution; a sensor operable to detect an orientation of the display; and an image processor operable to process video images captured by the video camera; and wherein the augmented reality system is operable in a second mode, in which the image processor is operable to process a captured video image to generate a processed video image, the processed video image corresponding to a portion of the full captured video image that is less than the full captured video image; the image processor is operable to present to the display successive processed video images corresponding to respective captured video images at a second, lower frame rate than the first frame rate, thereby creating a slow-motion effect; and wherein the position of the portion of the captured video image being processed to generate the current processed video image for display is responsive to the detected orientation of the display.

AUGMENTED REALITY SYSTEM AND METHOD

The present invention relates to an augmented reality system and method.

Current augmented reality (AR) systems comprise a means for a user to observe the real world, coupled with a means to augment that view with computer graphics.

- 5 Typical devices comprise transparent glasses onto which graphics can be projected, creating a semi-transparent augmentation overlay, or similar glasses with light emitting elements embedded therein, such as a sparse array of organic light emitting diodes (OLEDs), whose comparative brightness when active tend to obscure the scene behind them and so present an augmentation that appears more solid.
- 10 Another device that can provide augmented reality is a suitably adapted virtual reality (VR) device. Virtual reality devices typically only display computer graphics and hence to not comprise transparent gasses or a window onto the real world; however, if also equipped with a forward facing video camera, the video image captured by this camera can be displayed to provide a similar AR experience to the other devices. In this case, the video image can be
- 15 modified directly to include an augmentation layer of computer graphics of any effective degree of transparency.

These devices have the potential to bring information and entertainment into the user's everyday experience of the real world. However, there are still limits to how this experience can be augmented by such devices.

- 20 The present invention seeks to address or mitigate this problem.

In a first aspect of the present invention, an augmented reality system is provided in accordance with claim 1.

In another aspect, a method of video processing for an augmented reality system is provided in accordance with claim 3.

- 25 Further respective aspects and features of the invention are defined in the appended claims.

Embodiments of the present invention will now be described by way of example with reference to the accompanying drawings, in which:

Figure 1 is a schematic diagram of an entertainment system in accordance with an embodiment of the present invention.

Figure 2 is a schematic diagram of a head mounted display in accordance with an embodiment of the present invention.

- 5 Figure 3 is a flow chart of a method of video processing in accordance with an embodiment of the present invention.

An augmented reality system and method are disclosed. In the following description, a number of specific details are presented in order to provide a thorough understanding of the embodiments of the present invention. It will be apparent, however, to a person skilled in the art that these specific details need not be employed to practice the present invention. Conversely, specific details known to the person skilled in the art are omitted for the purposes of clarity where appropriate.

An example arrangement of an entertainment system comprising a videogame console and a head mounted display unit (HMD) is illustrated in Figure 1.

- 15 Figure 1 schematically illustrates the overall system architecture of a Sony® PlayStation 4® entertainment device. A system unit 10 is provided, with various peripheral devices connectable to the system unit.

The system unit 10 comprises an accelerated processing unit (APU) 20 being a single chip that in turn comprises a central processing unit (CPU) 20A and a graphics processing unit (GPU) 20B. The APU 20 has access to a random access memory (RAM) unit 22. The APU 20 communicates with a bus 40, optionally via an I/O bridge 24, which may be a discreet component or part of the APU 20. Connected to the bus 40 are data storage components such as a hard disk drive 37, and a Blu-ray ® drive 36 operable to access data on compatible optical discs 36A. Additionally the RAM unit 22 may communicate with the bus 40. Optionally also connected to the bus 40 is an auxiliary processor 38. The auxiliary processor 38 may be provided to run or support the operating system.

The system unit 10 communicates with peripheral devices as appropriate via an audio/visual input port 31, an Ethernet ® port 32, a Bluetooth ® wireless link 33, a Wi-Fi ® wireless link 34, or one or more universal serial bus (USB) ports 35. Audio and video may be output via an AV output 39, such as an HDMI port.

The peripheral devices may include a monoscopic or stereoscopic video camera 41 such as the PlayStation Eye ®; wand-style videogame controllers 42 such as the PlayStation Move ® and conventional handheld videogame controllers 43 such as the DualShock 4 ®; portable entertainment devices 44 such as the PlayStation Portable ® and PlayStation Vita ®; a
5 keyboard 45 and/or a mouse 46; a media controller 47, for example in the form of a remote control; and a headset 48. Other peripheral devices may similarly be considered such as a printer, or a 3D printer (not shown).

The GPU 20B, optionally in conjunction with the CPU 20A, generates video images and audio for output via the AV output 39. Optionally the audio may be generated in conjunction
10 with or instead by an audio processor (not shown). The video and optionally the audio may be presented to a television 51. Where supported by the television, the video may be stereoscopic. The audio may be presented to a home cinema system 52 in one of a number of formats such as stereo, 5.1 surround sound or 7.1 surround sound. Video and audio may likewise be presented to a head mounted display unit 53 worn by a user 60.

15 In operation, the entertainment device defaults to an operating system such as a variant of FreeBSD 9.0. The operating system may run on the CPU 20A, the auxiliary processor 38, or a mixture of the two. The operating system provides the user with a graphical user interface such as the PlayStation Dynamic Menu. The menu allows the user to access operating system features and to select games and optionally other content.

20 Referring now to Figure 2, the user 60 is wearing an exemplary HMD 53 on their head. In this example, the HMD comprises a frame 240 formed of a rear strap and a top strap, and a display portion 250 that comprises a separate respective display for each of the user's eyes.

In this example, the HMD of Figure 2 completely (or at least substantially completely) obscures the user's view of the surrounding environment. All that the user can see is the pair
25 of images displayed within the HMD. The HMD has associated headphone audio transducers or earpieces 260 which fit into the user's left and right ears 270. The earpieces 260 replay an audio signal provided from an external source, which may be the same as the video signal source which provides the video signal for display to the user's eyes.

The combination of the fact that the user can see only what is displayed by the HMD and,
30 subject to the limitations of the noise blocking or active cancellation properties of the

earpieces and associated electronics, can hear only what is provided via the earpieces, mean that this HMD may be considered as a so-called “full immersion” HMD.

As noted above, in embodiments of the present invention, however, in a first mode the user is provided with a real-time view of the real, ‘outside’ world. This can be provided by projecting a view of the outside (captured using a camera, for example a camera mounted on the HMD) via the HMD’s displays, and by providing a microphone to generate an input sound signal (for transmission to the earpieces) dependent upon the ambient sound.

A front-facing camera 222 may capture images to the front of the HMD, in use. A Bluetooth® antenna 224 may provide communication facilities or may simply be arranged as a directional antenna to allow a detection of the direction of a nearby Bluetooth transmitter.

In operation, a video signal is provided for display by the HMD. This could be provided by an external video signal source 280 such as the PlayStation 4 or a data processing apparatus (such as a personal computer), in which case the signals could be transmitted to the HMD by a wired or a wireless connection 282 or 39. For example, video signals from the front-facing camera may be sent to the PlayStation 4 for processing, and then are relayed back to the HMD for display. Examples of suitable wireless connections include Bluetooth® connection 33. Audio signals for the earpieces 260 can be carried by the same connection. Similarly, any control signals passed from the HMD to the video (audio) signal source, such as position or orientation signals obtained from one or more accelerometers or gyroscopes in the HMD (not shown) may be carried by the same connection. Furthermore, a power supply 283 (including one or more batteries and/or being connectable to a mains power outlet) may be linked by a cable 284 to the HMD. Note that the power supply 283 and the video signal source 280 may be separate units or may be embodied as the same physical unit.

If one or more cables are used, the physical position at which the cable 282 and/or 284 enters or joins the HMD is not particularly important from a technical point of view. Aesthetically, and to avoid the cable(s) brushing the user’s face in operation, it would normally be the case that the cable(s) would enter or join the HMD at the side or back of the HMD (relative to the orientation of the user’s head when worn in normal operation). Accordingly, the position of the cables 282, 284 relative to the HMD in Figure 2 should be treated merely as a schematic representation.

Accordingly, the arrangement of Figure 2 provides an example of a head-mounted display unit (HMD) comprising a frame to be mounted onto an observer's head, the frame defining one or two eye display positions which, in use, are positioned in front of a respective eye of the observer and a display element mounted with respect to each of the eye display positions, the display element providing a virtual image of a video display of a video signal from a video signal source to that eye of the observer.

Figure 2 shows just one example of an HMD. Other formats are possible: for example an HMD could use a frame more similar to that associated with conventional eyeglasses, namely a substantially horizontal leg extending back from the display portion to the top rear of the user's ear, possibly curling down behind the ear.

The above system provides an illustrative example of an augmented reality system.

More generally, in an embodiment of the present invention, the augmented reality system comprises a video input adapted to receive video images of a real-world scene captured by a video camera at a first frame rate; a display operable to display to a user video images originating from the video camera at a first resolution; a sensor operable to detect an orientation of the display; and an image processor operable to process video images captured by the video camera.

These components may be integral to or inserted/attached to a head mounted unit (HMD) worn by a user, or (for example in the case of the image processor) may be remote to the head mounted unit but in communication with it. It will be appreciated that the video camera may be a stereoscopic video camera providing left-eye and right-eye video images; consequently subsequent references to a video camera herein may be assumed to encompass such a stereoscopic video camera unless otherwise stated, and similarly subsequent references to an image or video image can be assumed to encompass stereoscopic images unless otherwise stated.

Typically the video camera is integral to the head-mounted unit so that the captured video images track the head movements of the user. In this case, the video input is internal to the system and refers to the link between the camera and the processor. Alternatively however, the video camera may be attachable to the head mounted unit to achieve a similar effect but may optionally be reusable for other purposes. Hence for example the camera may be part of a mobile phone that can be inserted into the head mounted unit, or may be a webcam or

sports video camera designed to be mountable on headgear. In this case the video input will typically be a wired or wireless interface with the camera.

The display may comprise two separate display units for the left and right eyes, or may comprise a single display unit logically split into left and right eye display regions. Again the display may be integral to the head mounted unit, or may be attachable (for example in the form of a display on a mobile phone that is inserted into the head mounted unit).

The sensor may comprise one or more from a list consisting of an accelerometer, gyroscope or optical tracking means, selected to allow detection of the orientation of the head mounted unit (and by extensions, the display within the head mounted unit) as the user moves their head around. Optionally lateral movement horizontally, vertically and forwards/backwards may also be detected in addition to rotation. The sensor may be integral to the head mounted unit, or may be part of a detachable device such as a mobile phone inserted into the head mounted unit.

The image processor may comprise one or more central processing units or one or more graphics processing units, or any combination of these either as separate components or a single component (for example, the APU 20 of the PlayStation 4). At least one processor of the image processor may be integral to the head mounted unit, or may be attachable, for example in the form of one or more processors of a mobile phone attached to the head mounted display; similarly at least one processor of the image processor may be remote from the head mounted display, for example in a videogame console operating in conjunction with the head mounted display. Hence the image processor may comprise one or more processors located as applicable integrally or detachably within the head mounted unit, or remotely in an associated computer.

In an embodiment of the present invention, the augmented reality system is operable in two modes. In the first mode, the video camera is operable to capture video images of a real-world scene at a first frame rate (for example 60fps, but equally it may be 50, 30, 25 or any other suitable rate as the first frame rate), and the display is operable to display to a user video images originating from the video camera at a first resolution, and typically at a frame rate corresponding to the first frame rate (for example if the first frame rate is 60fps, then showing each captured image at 60fps, or showing every other captured image at 30 fps, which is the effective capture rate for alternate frames). As such, the user is presented with a real-time view from the video camera feed. The view may correspond to the whole of the

captured video image, or may correspond to only a portion of the captured video image, for example where there is a difference in aspect ratio between the resolutions of the camera and the display. Optionally, the presented video images are augmented with computer graphics to generate an augmented image.

- 5 However, the augmented reality system is also operable in a second mode. In this second mode, the captured video is sent to a buffer, and the buffered video images originating from the video camera are displayed to the user at a second, different, slower rate than the capture rate.

This creates a slow-motion effect for the real-world scene perceived by the user, optionally
10 also augmented with computer graphics as in the first mode. For example, if the first frame rate was 60fps, then the second frame rate for showing each captured image may be 30fps, representing a 50% slowdown. As noted above, this could be achieved by changing the frame rate for each frame from 60fps to 30fps, or by switching from showing alternate frames captured at 60fps at 30fps to showing each frame at 30fps (which may reduce processing and
15 display overheads during the majority of the time when the device is in the first mode).

In the event that the desired slow-down would result in an effective second frame rate that is below an acceptable threshold (for example, a 75% slowdown resulting in an unacceptable 15fps output), then optionally the image processor is operable to perform inter-image interpolation to generate one or more processed video images for presentation in between
20 processed video images corresponding to respective captured video images. In this way, the rate of images directly derived from the capture video can still be reduced by 75% whilst increasing the rate of images presented to the user to an acceptable rate. For example, for a capture rate of 60fps, and a slowdown of 75% to a slow motion frame rate of 15fps, if one image was interpolated between each original image, then an acceptably smooth presentation
25 frame rate of 30fps can be achieved whilst still maintaining the 75% slow down.

Alternatively or in addition, the video camera may be operable to capture video images at a third, higher frame rate than the first frame rate. For example, increasing the capture frame rate to 120fps would enable a 75% slow down to 30 fps without interpolation.

Hence more generally, the image processor is operable to present to the display successive
30 processed video images corresponding to respective captured video images at a second, lower frame rate than the first frame rate, thereby creating a slow-motion effect.

As was noted previously, a characteristic feature of virtual reality and augmented reality is the ability to look around a scene by moving your head to change the orientation of the head mounted unit. In the first mode of operation, this is straightforward to achieve because the camera is generating a live feed from which images can be obtained for presentation to the user in real time.

However, in slow motion, the live feed is buffered and presented to the user at a reduced rate; this means that movements of the user's head will not be reflected in the next image from the camera that is being supplied to the user, since that image was obtained earlier than the head movement, and supplied to the buffer. Hence in principle it is not possible to provide the ability to look around in the second mode of operation because the slowed down video can no longer track the user's movements in real time.

However, in an embodiment of the present invention, in the second mode the image processor is operable to process a captured video image so that it corresponds to a portion of the full captured video image that is less than the full captured video image.

Hence for example the image processor may only present a 75% x 75% central portion of the image, compared to either the original image or the portion of the image presented in the first mode if this is already cropped for reasons of aspect ratio or the like.

Consequently, the position of the portion of the captured video image being processed to generate the current processed video image for display is responsive to the detected orientation of the display.

In other words, when the user subsequently moves their head, the selected portion pans within the uncropped/original/larger image in a direction corresponding with the direction of head movement, to give the illusion of being able to look around the view.

It will be appreciated that this illusion will break once the user's head movements correspond to a position where the selected portion would extend beyond the limits of the original captured image, but within these limits, the illusion of time-dilated AR can be maintained.

It will be appreciated however that the smaller the size of the selected portion relative to the original captured image, the greater the relative amount of head movement that can be accommodated before the illusion is broken. However, there is also a corresponding trade-off

between image portion size and the effective resolution of the presented image if this portion is blown-up to occupy the full display for the user.

Accordingly, in an embodiment of the present invention, in the second mode the image processor performs intra-image interpolation to re-scale a portion of a captured video image to the first resolution to generate a processed video image.

It will similarly be appreciated that a sudden switch from the first mode, with a first field of view and a real time presentation, to a second mode with a significantly smaller field of view and a slow motion presentation could be disorientating for a user.

Accordingly, in an embodiment of the present invention the image processor is operable to present a transition phase in which the portion of the full captured video image changes over successive presented images from a first size portion to a second, smaller size portion. Hence for example it may transition from 100% to 75%, or some other size (66%, 50%, 33% or any other suitable size). Typically this transition may take between 1/10 and 1 second, but may be faster or slower.

In the above example, it was assumed that the selected portion was initially central and/or that the portion at the end of a transition phase was similarly central. However, in an embodiment of the present invention, the position of the portion (or the portion at the end of a transition phase) may be selected responsive to a predetermined feature of either the captured video image being processed, or a the location of a computer generated augmentation superposed over the presented processed image.

Examples of such features in a capture image or image sequence include a face (or a specific recognised face), a fiduciary marker that may be a target for an augmentation, an object being tracked by an AR application, an element of the image that appears to be moving, or moving in a different direction to the overall image motion, or moving at a speed within the image above a predetermined threshold, or an element of the image having a brightness or colour that is either above or (separately) below a threshold, or a threshold difference for a mean of the image. Examples of such features in a computer generated augmentation include the appearance of a new virtual object (for example a ball, or an explosion / gun shot hit or the like); or a change of state for a virtual object (for example the ball crossing a goal line, or a virtual character pulling a taunting expression).

It will be appreciated that not every instance of one of these features may trigger a transition between first and second mode.

Alternatively to the above, the user may be presented with only a portion of the original captured image in both first and second modes, so that the switch between modes is not accompanied by a change in the field of view. As a non-limiting example, the images presented to the user in both modes may comprise the central 75% x 75% portion of the captured image, before any panning of that portion in the second mode.

In other words, in this embodiment then by default the camera has a wider field of view than the images presented to the user in either mode.

10 This may be the case for example, if the video images are received from a remote camera that is used by plural AR users, for example as spectators at a sports match; this enables each user to look around independently of the others, since no actual motion of the camera is required. Optionally the camera may have a much higher resolution than the augmented reality system display; for example, the camera may be a so-called 4K UHD camera, whilst the display is a 15 1080 or 720 HD display. This enables presentation of good quality images whilst still permitting the user to look around, at least to some extent. In a variant of this embodiment, the orientation of the user's display is transmitted to a broadcast server, so that only the relevant portion of the 4K captured image is transmitted to the respective user, thereby saving bandwidth. Optionally a small buffer region of additional image may be sent so that the 20 selected portion can be slightly adjusted into such a region to account for any incremental motion of the user / display during the frame transmission time.

Given such an arrangement, the transition from first to second mode may be triggered remotely by the broadcaster, and the portion may also be selected by the broadcaster, for example to let the user see if a ball was out in a tennis match, or if a foul was committed in a 25 game of football. It will be appreciated therefore that such a system could be used to present highlights to all viewers, and also by individual viewers to focus on features of personal interest during any suitably adapted broadcast, whether in sports, news, the arts or any appropriate field.

Separately to the above, the rate of slow-down may also transition from the first frame rate to 30 the second frame rate over a finite period, which would typically correspond to that of the portion transition.

In any of the above cases, the transition between fields of view or effective rates of time may also be accompanied by a characteristic sound effect that the user can learn to associate with the transition from first to second modes, such as a dropping audio tone evocative of a recording slowing down.

- 5 Alternatively or in addition, the augmented reality system may comprise a microphone (for example integral to the head mounted unit, or attachable thereto, for example as part of a mobile phone that can be attached to the head mounted unit, or remote, for example associated with a remote camera in a broadcast scenario).

10 Then, in the first mode, external sounds are captured by the microphone and presented to the user in real time, but in the second mode, the sounds are processed to slow them down by an extent corresponding to the change in rate between the first video capture rate and the second video capture rate.

This can provide an additional level of realism to the user's experience. Optionally stereo balance or a directional mix of the recorded sounds can be adjusted responsive to the user's
15 head motions (and by extension, the orientation of the display).

The above apparatus and techniques enable an augmented reality system to provide a slow-motion or time dilation view of the external, real world, in which the user is still able to look around such that their viewpoint changes in real time, at least to some extent, within this slow-motion representation.

- 20 However, by decoupling the user from a real-world video feed (particularly in the case of a feed from a camera associated with their own head mounted unit, there is a risk that the user will no longer know what is currently happening in the outside world.

Accordingly, in an embodiment of the present invention, the second mode can be terminated in the event of one or more circumstances, including but not limited to:

- 25 detection of a predetermined event within a currently captured video image not yet presented for display, such as detection of movement within the video towards the user (for example an object appearing to get larger over successive frames);

detection of a nearby object in the environment of the user, for example by use of proximity detectors on the head mounted display, or analysis of an image from a camera monitoring the
30 user;

detection of a sound in the environment of the user that exceeds a threshold volume, and/or has a predetermined property, such as a harmonic/formant structure indicative of a person's voice; and

a termination signal from a broadcaster that had initiated the second mode.

- 5 In addition to these externally driven means of terminating the second mode, in embodiments of the invention the user themselves may terminate the second mode, for example by pressing a predetermined button on a controller, or performing a predetermined motion or sequence of motions, such as shaking their head as if to say 'no'.

Similarly, the system may have one or more termination criteria. One such criteria may be
10 that the buffer used to hold captured video frames for processing has reached capacity. Another may be that the system has been in the second mode for a predetermined period of time. It will be appreciated that the predetermined period of time may be shorter than the time taken to reach buffer capacity.

If the purpose of the time limit is to preserve the illusion of being able to look around within
15 the time dilated view, then the predetermined period may also vary depending on the comparative size of the portion of the image used in the second mode, since the user is more likely to move outside the available area of the original image in a short time if the portion of the image is comparatively large. Hence the predetermined period of time for a 75%x75% portion may be shorter than for a 40%x40% portion.

20 It will be appreciated that the user will not necessarily be able to hear or respond to current events during use of the head mounted device in the second mode. Accordingly, the head mounted display may comprise an externally detectable indicator that is operable to be activated for the duration of the second mode. This tells anyone intending to interact with the user that the user is in the second mode and may not be responsive to normal audio or visible
25 cues.

Referring now to Figure 3, a method of video processing for an augmented reality system according to an embodiment of the present invention comprises:

In a first step s310, capturing video images of a real-world scene at a first frame rate;

In a second step s320, displaying in a first mode to a user video images originating from the
30 video camera at said first frame rate and at a first resolution;

In a third step 330, detecting an orientation of a display; and

In a fourth step 340, processing video images captured by the video camera,

and wherein in a second mode, the processing step comprises the substeps of:

in a first sub-step 342, processing a captured video image to generate a processed video

5 image, the processed video image corresponding to a portion of the full captured video image that is less than the full captured video image; and

in a second sub-step 344, displaying successive processed video images corresponding to respective captured video images at a second, lower frame rate than the first frame rate, thereby creating a slow-motion effect; and

10 in the first sub-step, presenting a transition phase in which the position of the portion selected from the full captured video corresponds to one or more selected from the list consisting of a predetermined feature of the captured video image being processed, and the location of a computer generated augmentation superposed over the presented processed image, the position of the portion of the captured video image then being processed to generate the
15 current processed video image for display responsive to the detected orientation of the display.

It will be apparent to a person skilled in the art that variations in the above method corresponding to operation of the various embodiments of the apparatus as described and claimed herein are considered within the scope of the present invention, including but not
20 limited to:

- in the second mode, performing intra-image interpolation to re-scale a portion of a captured video image to the first resolution to generate a processed video image;
- in the second mode, capturing video images at a third, higher frame rate than the first frame rate;
- 25 - in the second mode, performing inter-image interpolation to generate one or more processed video images for presentation in between processed video images corresponding to respective captured video images;
- presenting a transition phase in which the portion of the full captured video image changes over successive presented images from a first size portion to a second, smaller size portion;
30
- ending the second mode of operation in the event of one or more selected from the list consisting of: a predetermined time elapsing; detection of a predetermined motion or sequence of motions; activation of a predetermined control by a user; detection of a

predetermined event within a currently captured video image not yet presented for display; detection of a nearby object in the environment; and a memory reaching a predetermined occupancy;

- activating an externally detectable indicator for the duration of the second mode;
- 5 - generating a sound effect specifically associated with a transition from the first mode to the second mode;
- in the first mode, capturing external sounds with a the microphone and presenting them to the user; and in the second mode, processing the sounds to slow them down by an extent corresponding to the change in rate between the first video capture rate
- 10 and the second video capture rate; and
- in the first mode, processing a captured video image to generate a processed video image, the processed video image corresponding to a portion of the full captured video image that is less than the full captured video image.

It will be appreciated that the above methods may be carried out on conventional hardware

15 suitably adapted as applicable by software instruction or by the inclusion or substitution of dedicated hardware.

Thus the required adaptation to existing parts of a conventional equivalent device may be implemented in the form of a computer program product comprising processor implementable instructions stored on a non-transitory machine-readable medium such as a

20 floppy disk, optical disk, hard disk, PROM, RAM, flash memory or any combination of these or other storage media, or realised in hardware as an ASIC (application specific integrated circuit) or an FPGA (field programmable gate array) or other configurable circuit suitable to use in adapting the conventional equivalent device. Separately, such a computer program may be transmitted via data signals on a network such as an Ethernet, a wireless network, the

25 Internet, or any combination of these or other networks.

CLAIMS

1. An augmented reality system, comprising
a video input adapted to receive video images of a real-world scene captured by a
5 video camera at a first frame rate;
a display operable in a first mode to display to a user video images originating from
the video camera at a first resolution;
a sensor operable to detect an orientation of the display; and
an image processor operable to process video images captured by the video camera;
10 and wherein
the augmented reality system is operable in a second mode, in which
the image processor is operable to process a captured video image to generate a
processed video image, the processed video image corresponding to a portion of the full
captured video image that is less than the full captured video image;
15 the image processor is operable to present to the display successive processed video
images corresponding to respective captured video images at a second, lower frame rate than
the first frame rate, thereby creating a slow-motion effect; and wherein
the image processor is operable to present a transition phase in which the position of
the portion selected from the full captured video corresponds to one or more selected from
20 the list consisting of:
 - i. a predetermined feature of the captured video image being processed; and
 - ii. the location of a computer generated augmentation superposed over the
presented processed image,
the positioning of the portion of the captured video image then being processed by the
25 image processor to generate the current processed video image for display responsive to the
detected orientation of the display.
2. The augmented reality system of claim 1, in which in the second mode, the image
processor performs intra-image interpolation to re-scale a portion of a captured video image
to the first resolution to generate a processed video image.
- 30 3. The augmented reality system of claim 1 or claim 2, in which in the second mode, the
video camera is operable to capture video images at a third, higher frame rate than the first
frame rate.

4. The augmented reality system of any one of the preceding claims, in which in the second mode, the image processor is operable to perform inter-image interpolation to generate one or more processed video images for presentation in between processed video images corresponding to respective captured video images.
5. The augmented reality system of any preceding claim, in which the image processor presents a transition phase in which the portion of the full captured video image changes over successive presented images from a first size portion to a second, smaller size portion.
6. The augmented reality system of any preceding claim, in which the image processor is operable to end the second mode of operation in the event of one or more selected from the list consisting of:
- i. a predetermined time elapsing;
 - ii. detection of a predetermined motion or sequence of motions;
 - iii. activation of a predetermined control by a user;
 - iv. detection of a predetermined event within a currently captured video image not yet presented for display;
 - v. detection of a nearby object in the environment; and
 - vi. a memory reaching a predetermined occupancy.
7. The augmented reality system of any preceding claim, comprising:
an externally detectable indicator that is operable to be activated for the duration of the second mode.
8. The augmented reality system of any preceding claim, comprising
an audio output; and in which
the augmented reality system is operable to generate a sound effect specifically associated with a transition from the first mode to the second mode.
9. The augmented reality system of any preceding claim, comprising
a microphone; and
an audio output; and in which
in the first mode, external sounds are captured by the microphone and presented to the user;

and in the second mode, the sounds are processed to slow them down by an extent corresponding to the change in rate between the first video capture rate and the second video capture rate.

10. The augmented reality system of any proceeding claim, in which the image processor
5 is operable in the first mode to process a captured video image to generate a processed video image, the processed video image corresponding to a portion of the full captured video image that is less than the full captured video image.

11. A method of video processing for an augmented reality system, comprising the steps of:

10 receiving captured video images of a real-world scene at a first frame rate;
displaying in a first mode to a user video images originating from the video camera at said first frame rate and at a first resolution;

detecting an orientation of a display; and

processing video images captured by the video camera;

15 and wherein

in a second mode,

the processing step comprises processing a captured video image to generate a processed video image, the processed video image corresponding to a portion of the full captured video image that is less than the full captured video image;

20 the processing step comprises displaying successive processed video images corresponding to respective captured video images at a second, lower frame rate than the first frame rate, thereby creating a slow-motion effect; and wherein

the processing step comprises presenting a transition phase in which the position of the portion selected from the full captured video corresponds to one or more selected from
25 the list consisting of

i. a predetermined feature of the captured video image being processed; and

ii. the location of a computer generated augmentation superposed over the presented processed image, and

the position of the portion of the captured video image then being processed to
30 generate the current processed video image for display responsive to the detected orientation of the display.

12. The method of video processing of claim 11, in which in the second mode, the video camera is operable to capture video images at a third, higher frame rate than the first frame rate.

13. The method of video processing of any one of claims 11 or 12, in which in the second
5 mode, the image processor is operable to perform inter-image interpolation to generate one or more processed video images for presentation in between processed video images corresponding to respective captured video images.

14. A computer readable medium having computer executable instructions adapted to cause a computer system to perform the method of any one of claims 11 to 13.

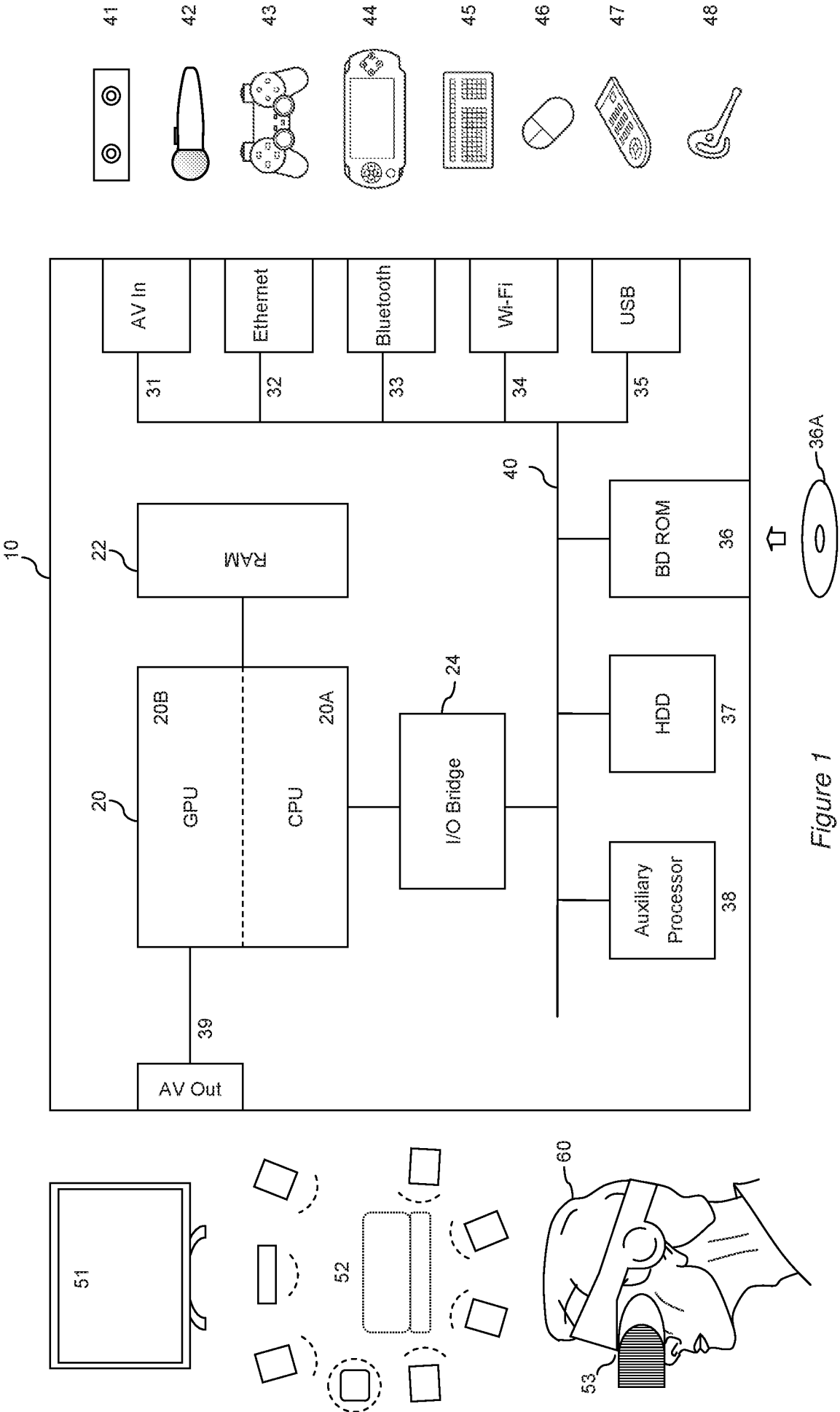
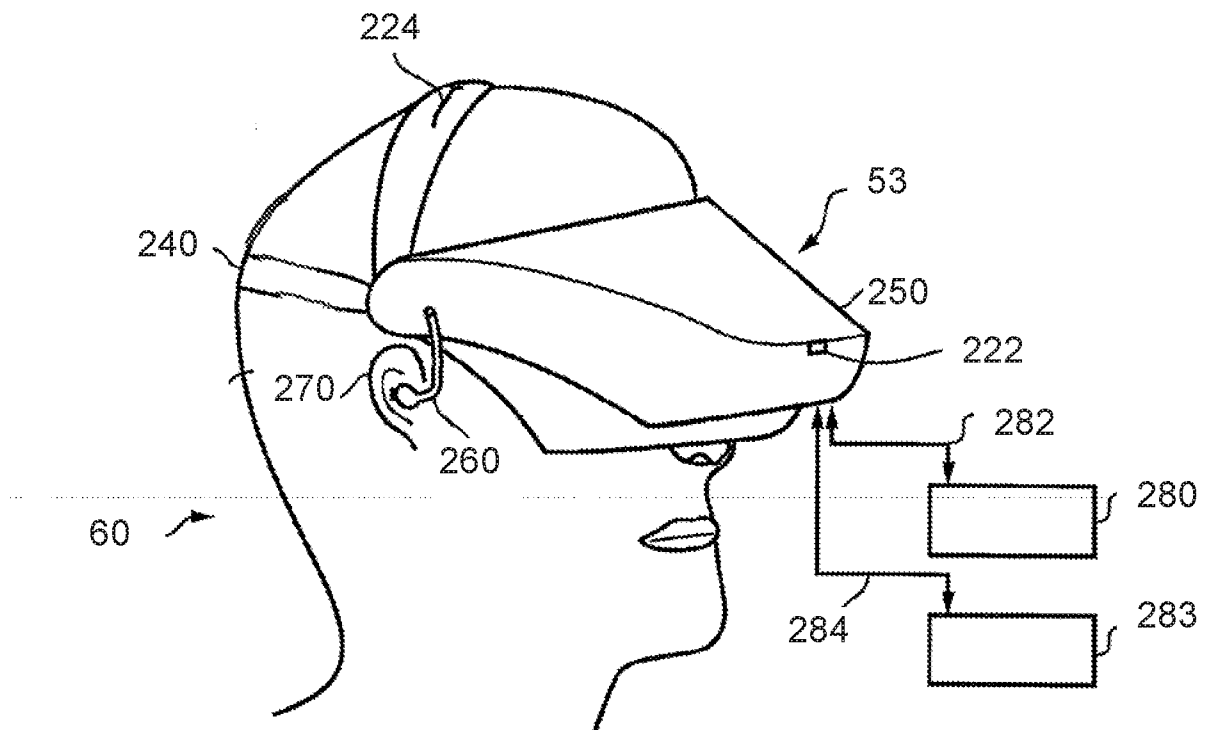
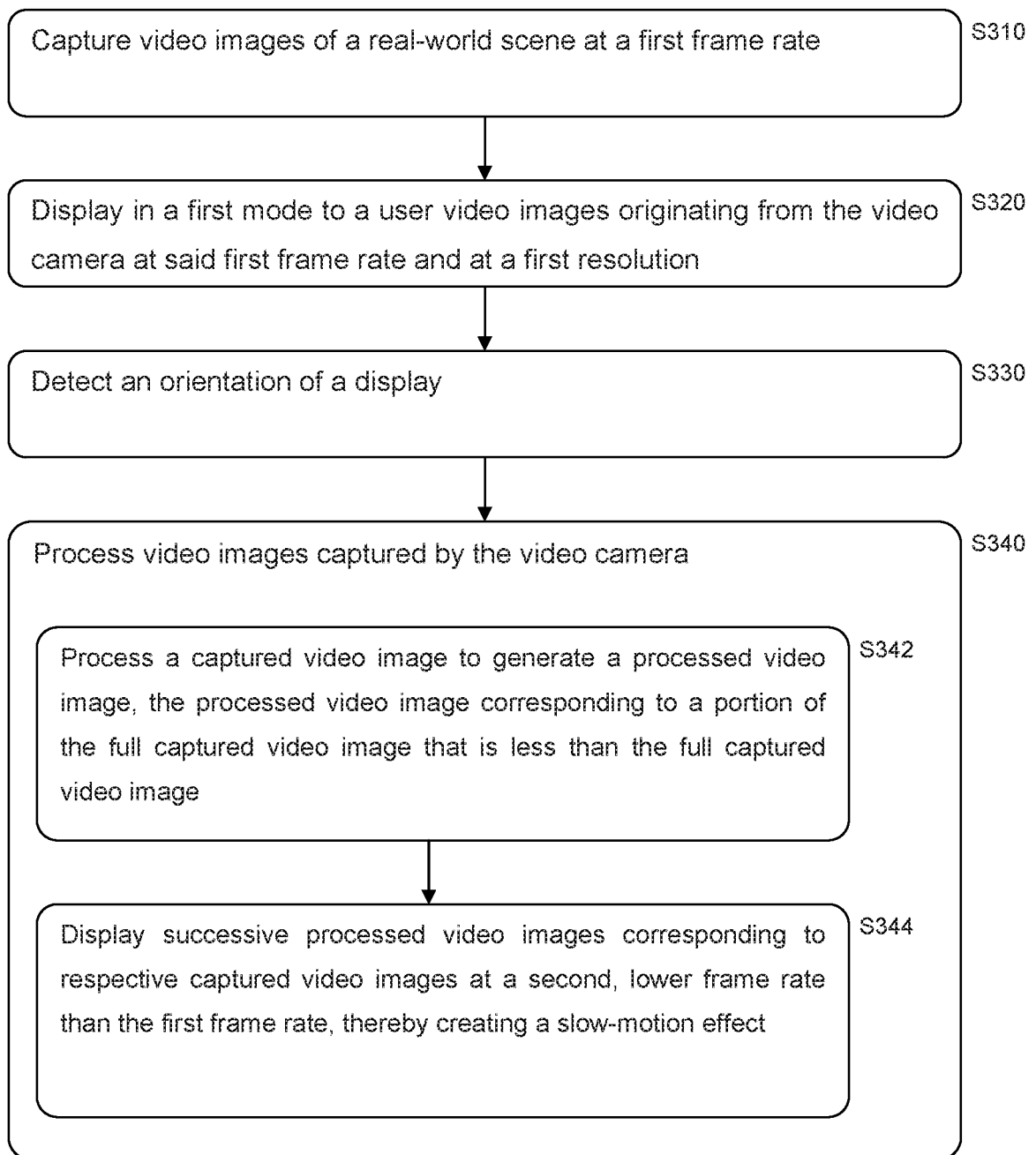


Figure 1

*Figure 2*

*Figure 3*

INTERNATIONAL SEARCH REPORT

International application No
PCT/GB2017/051180

A. CLASSIFICATION OF SUBJECT MATTER

INV. H04N5/783 H04N21/4728 H04N21/4402 H04N21/422 H04N21/4223
H04N21/414

ADD.

According to International Patent Classification (IPC) or to both national classification and IPC

B. FIELDS SEARCHED

Minimum documentation searched (classification system followed by classification symbols)
H04N

Documentation searched other than minimum documentation to the extent that such documents are included in the fields searched

Electronic data base consulted during the international search (name of data base and, where practicable, search terms used)

EPO-Internal, WPI Data

C. DOCUMENTS CONSIDERED TO BE RELEVANT

Category*	Citation of document, with indication, where appropriate, of the relevant passages	Relevant to claim No.
Y	US 9 143 693 B1 (ZHOU CHANGYIN [US] ET AL) 22 September 2015 (2015-09-22) figures 1, 3, 4, 7 column 3, line 47 - column 5, line 21 column 6, line 15 - column 8, line 55 column 11, line 10 - line 24 column 12, line 10 - column 15, line 65 -----	1-14
Y	US 2011/043644 A1 (MUNGER REJEAN J Y B [CA] ET AL) 24 February 2011 (2011-02-24) figures 1-4 paragraph [0031] - paragraph [0055] -----	1,2,10, 11,14
Y	US 9 360 671 B1 (ZHOU CHANGYIN [US]) 7 June 2016 (2016-06-07) figures 1, 3, 7-9 column 5, line 51 - column 9, line 47 column 23, line 10 - column 28, line 39 ----- -/-	1-14



Further documents are listed in the continuation of Box C.



See patent family annex.

* Special categories of cited documents :

"A" document defining the general state of the art which is not considered to be of particular relevance

"E" earlier application or patent but published on or after the international filing date

"L" document which may throw doubts on priority claim(s) or which is cited to establish the publication date of another citation or other special reason (as specified)

"O" document referring to an oral disclosure, use, exhibition or other means

"P" document published prior to the international filing date but later than the priority date claimed

"T" later document published after the international filing date or priority date and not in conflict with the application but cited to understand the principle or theory underlying the invention

"X" document of particular relevance; the claimed invention cannot be considered novel or cannot be considered to involve an inventive step when the document is taken alone

"Y" document of particular relevance; the claimed invention cannot be considered to involve an inventive step when the document is combined with one or more other such documents, such combination being obvious to a person skilled in the art

"&" document member of the same patent family

Date of the actual completion of the international search

8 June 2017

Date of mailing of the international search report

16/06/2017

Name and mailing address of the ISA/

European Patent Office, P.B. 5818 Patentlaan 2
NL - 2280 HV Rijswijk
Tel. (+31-70) 340-2040,
Fax: (+31-70) 340-3016

Authorized officer

Döttling, Martin

INTERNATIONAL SEARCH REPORT

International application No

PCT/GB2017/051180

C(Continuation). DOCUMENTS CONSIDERED TO BE RELEVANT		
Category*	Citation of document, with indication, where appropriate, of the relevant passages	Relevant to claim No.
Y	US 9 197 864 B1 (STARNER THAD EUGENE [US] ET AL) 24 November 2015 (2015-11-24) figures 1-3, 6, 8, 9 column 3, line 25 - column 6, line 49 column 9, line 36 - column 13, line 50 -----	1-9, 11-14
Y	WO 00/72310 A1 (KONINKL PHILIPS ELECTRONICS NV [NL]) 30 November 2000 (2000-11-30) figure 2 page 1, line 18 - line 24 page 6, line 5 - line 29 -----	9

INTERNATIONAL SEARCH REPORT

Information on patent family members

International application No

PCT/GB2017/051180

Patent document cited in search report	Publication date	Patent family member(s)	Publication date
US 9143693	B1	22-09-2015	NONE
US 2011043644	A1	24-02-2011	CA 2781064 A1 26-05-2011 EP 2502410 A1 26-09-2012 US 2011043644 A1 24-02-2011 US 2016282624 A1 29-09-2016 WO 2011060525 A1 26-05-2011
US 9360671	B1	07-06-2016	NONE
US 9197864	B1	24-11-2015	US 9197864 B1 24-11-2015 US 9466112 B1 11-10-2016
WO 0072310	A1	30-11-2000	DE 60009827 D1 19-05-2004 DE 60009827 T2 17-03-2005 EP 1099216 A1 16-05-2001 JP 2003500703 A 07-01-2003 US 6944510 B1 13-09-2005 WO 0072310 A1 30-11-2000