A method of decoding and outputting video information suitable for three-dimensional [3D] display, the video information comprising encoded main video information suitable for displaying on a 2D display and encoded additional video information for enabling three-dimensional [3D] display, the method comprising: receiving or generating three-dimensional [3D] overlay information to be overlayed over the video information; buffering a first part of the overlay information to be overlayed over the main video information in a first buffer; buffering a second part of overlay information to be overlayed over the additional video information in a second buffer; decoding the main video information and the additional video information and generating as a series of time interleaved video frames, each outputted video frame being either main video frame or additional video frame; determining a type of an video frame to be outputted being either main video frame or additional video frame; overlaying either first or second part of the overlay information on an video frame to be outputted in agreement with the determined type of frame outputting the video frames and the overlayed information.
FIG. 3
METHODOLOGY AND DEVICE FOR OVERLAYING 3D GRAPHICS OVER 3D VIDEO

FIELD OF THE INVENTION

[0001] The invention relates to a method of decoding and outputting video information suitable for three-dimensional (3D) display, the video information comprising encoded main video information suitable for displaying on a 2D display and encoded additional video information for enabling three-dimensional (3D) display, 3D overlay information being overlayed onto the video information.

[0002] The invention further relates to a device for decoding and outputting video information suitable for three-dimensional (3D) display, the video information comprising encoded main video information suitable for displaying on a 2D display and encoded additional video information for enabling three-dimensional (3D) display, the device adapted to overlay 3D overlay information onto the video information.

[0003] The invention relates to the field playback of 3D video information and 3D overlay information by a playback device, information to be displayed onto a 3D enabled display.

BACKGROUND OF THE INVENTION

[0004] Devices for rendering video data are well known, for example video players like DVD players, BD players or set top boxes for rendering digital video signals. The rendering device is commonly used as a source device to be coupled to a display device like a TV set. Image data is transferred from the source device via a suitable interface like HDMI.

[0005] With respect to the coded video information stream, for example this may under the format known as stereoscopic, where left and right (L+R) images are encoded. Alternatively, coded video information stream may comprise a 2D picture and an additional picture (L+D), a so-called depth map, as described in Oliver Sheer: “3D Video Communication”, Wiley, 2005, pages 29-34. The depth map conveys information about the depth of objects in the 2D image. The grey scale values in the depth map indicate the depth of the associated pixel in the 2D image. A stereo display can calculate the additional view required for stereo by using the depth value from the depth map and by calculating the required pixel transformation. The 2D video-depth map may be extended by adding occlusion and transparency information (DOT).

[0006] Currently in 3D systems, a known solution for the output video data to be transferred via the HDMI interface to the 3D display is time interleaving, wherein frames corresponding to Left or 2D information are interleaved with Right or DOT frames.

[0007] It is known that, for 2D video systems, application formats like for distribution of video content and playback device support overlay or real time generated graphics on top of the video. Overlay graphics being for example for example internally generated by the player device for on screen display (OSD) menus, or received, such as subtitles or other graphics.

[0008] However extending the known overlay models to 3D systems creates the problem that the performance requirements of drawing routines for the real-time generated overlay graphics are increased.

SUMMARY OF THE INVENTION

[0009] It is an object of the invention to provide a method for decoding and outputting video information and overlay information which is suitable for 3D systems

[0010] For this purpose, according to a first aspect of the invention, the method further comprises receiving or generating three-dimensional (3D) overlay information to be overlayed over the video information; buffering a first part of the overlay information to be overlayed over the main video information in a first buffer; buffering a second part of overlay information to be overlayed over the additional video information in a second buffer; decoding the main video information and the additional video information and generating as a series of time interleaved video frames, each outputted video frame being either main video frame or additional video frame; determining a type of an video frame to be outputted being either a main video frame or an additional video frame; overlaying either first or second part of the overlay information on an video frame to be outputted in agreement with the determined type of frame-outputting the video frames and the overlayed information.

[0011] For this purpose, according to a second aspect of the invention, the device described in the opening paragraph comprises means for receiving three-dimensional (3D) overlay information to be overlayed over the video information or generation means for generating three-dimensional (3D) overlay information to be overlayed over the video information a decoder for decoding the main video information and the additional video information, the decoder further adapted to generating as a series of time interleaved video frames, each outputted video frame being either main video frame or additional video frame; generation means for generating three-dimensional (3D) overlay information to be overlayed over the video information; a graphics processing unit comprising a first buffer for buffering a first part of the overlay information to be overlayed over the main video information and a second buffer for buffering a second part of overlay information to be overlayed over the additional video information; a controller for determining a type of an video frame to be outputted being either a main video frame or an additional video frame; a mixer for overlapping either first or second part of the overlay information on an video frame to be outputted in agreement with the determined type of frame; output means for outputting the video frames and the overlayed information.

[0012] The invention is also based on the following recognition. 3D Overlay graphics can no longer simply be composited with the 3D video output in systems outputting frames corresponding to Left or 2D information interleaved with Right or DOT frames, since the 3D video output switches between the two different video streams every frame. As an example, at time T the video output could contain the 2D frame, and at time T+1 the video output contains accompanying depth information for the frame at time T. The graphics that need to be composited with the video at time T (the 2D graphics) greatly differ from the graphics that need to be composited with the video at time T+1 (the depth graphics or the R graphics). The graphics unit present in 2D video player devices is not fast enough to frame accurately update its graphics plane with these different graphics every frame. The solution according to the invention is to implement two buffers in the graphics unit. Each buffer is assigned to one of the output video streams. For example, for 2D+depth drawing, one buffer could be assigned for graphics overlay over the 2D frame and one buffer could be assigned for the graphics overlay over the depth frame. For 4K, similarly, one buffer
could be used for graphics overlay over the L frame, and one buffer could be assigned for overlay over the R frame. The advantage of this solution is that the slow graphics are decoupled from the frame accurate overlay engine, so that the processing requirements are significantly reduced.

[0013] Advantageously, the graphics control unit further comprises a controller is adapted to copy parts of a first overlay frame in the first buffer or parts of a second overlay frame in the second buffer at frame frequency for generating an overlay frame. When the player device handles 2D+DOL depth streams, this enables fast generation of occasion data, by copying the relevant areas from the buffered frames.

BRIEF DESCRIPTION OF THE DRAWINGS

[0014] These and other aspects of the invention will be apparent from and elucidated further with reference to the embodiments described by way of example in the following description and with reference to the accompanying drawings, in which

[0015] FIG. 1 shows schematically a system for receiving and displaying 3D video information in parts of which the invention may be practiced

[0016] FIG. 2 shows schematically a graphics processing unit of a known 2D video player.

[0017] FIG. 3 shows schematically the composition of video planes in known Blu-Ray (BD) systems

[0018] FIG. 4 illustrates schematically a graphics processing unit according to the invention

[0019] In the Figures, elements which correspond to elements already described have the same reference numerals.

DETAILED DESCRIPTION OF EMBODIMENTS

[0020] A system 1 for playback of 3D video information wherein the invention may be practiced is shown in FIG. 1. The system comprises a player device 10 and a display device 11 communicating via an interface 12. The player device 10 comprises a front end unit 12 responsible for receiving and pre-processing the coded video information stream to be displayed, and a processing unit for decoding, processing and generation a video stream to be supplied to the output 14. The display device comprises a rendering unit for rendering 3D views from the received.

[0021] With respect to the coded video information stream, for example this may under the format known as stereoscopic, where left and right (L+R) images are encoded. Alternatively, coded video information stream may comprise a 2D picture and an additional picture (L+D), a so-called depth map, as described in Oliver Sheer: "3D Video Communication", Wiley, 2005, pages 20-24. The depth map conveys information about the depth of objects in the 2D image. The grey scale values in the depth map indicate the depth of the associated pixel in the 2D image. A stereo display can calculate the additional view required for stereo by using the depth value from the depth map and by calculating the required pixel transformation. The 2D video-depth map may be extended by adding occlusion and transparency information (DOT). In a preferred embodiment, a flexible data format comprising stereo information and depth map, adding occlusion and transparency, as described in EP 08305420.5 (Attorney docket PH1010082), to be included herein by reference, is used.

[0022] With respect to the display device 11, this can be either a display device that makes use of controllable glasses to control the images displayed to the left and right eye respectively, or, in a preferred embodiment, the so-called autostereoscopic displays are used. A number of auto-stereoscopic devices that are able to switch between 2D and 3D displays are known, one of them being described in U.S. Pat. No. 6,069,650. The display device comprises an LCD display comprising actively switchable Liquid Crystal lenticular lens. In auto-stereoscopic displays, the processing side a rendering unit 16 converts the decoded video information received via the interface 12 from the player device 10 to multiple views and maps these onto the sub-pixels of the display panel 17. It is duly noted that the rendering unit 16 may reside either inside the player device 10, in such case the multiple views being sent via the interface.

[0023] With respect to the player device 10, this may be adapted to read the video stream from an optical disc, another storage media such as flash, or receive the video information via wired or wireless network, such as an internet connection. A known example of a Blu-Ray™ player is the PlayStation™ 3, as sold by Sony Corporation.


[0025] In the following, when referring to the BD application format, we refer specifically to the application formats as disclosed in the US application No. 2006-011011 (Attorney docket N021359) and in white paper “Blu-ray Disc Format 2.B Audio Visual Application Format Specifications for BD-ROM, March 2005” as published by the Blu-ray Disc Association.

[0026] It is known that BD systems also provide a fully programmable application environment with network connectivity thereby enabling the Content Provider to create interactive content. This platform is based on the Java™ platform and is known as “BD-J”. BD-J defines a subset of the Digital Video Broadcasting (DVB)-Multimedia Home Platform (MHP) Specification 1.0, publicly available as ETSI TS 101 812.

[0027] FIG. 2 illustrates a graphics processing unit (part of the processing unit 13) of a known 2D video player, namely a Blu-Ray player. The graphics processing unit is equipped with two read buffers (1304 and 1305), two preloading buffers (1302 and 1303) and two switches (1306 and 1307). The second read buffer (1305) enables the supply of an Out-of-Mux audio stream to the decoder even while the main MPEG stream is being decoded. The preloading buffers cache Text subtitles, Interactive Graphics and sounds effects (which are presented at button selection or activation). The preloading buffer 1303 stores data before movie playback begins and supplies data for presentation even while the main MPEG stream is being decoded.

[0028] This switch 1301 between the data input and buffers selects the appropriate buffer to receive packet data from any one of read buffers or preloading buffers. Before starting the main movie presentation, effect sounds data (if it exists), text subtitle data (if it exists) and Interactive Graphics (if preloaded Interactive Graphics exist) are preloaded and sent to each buffer respectively through the switch. The main MPEG stream is sent to the primary read buffer (1304) and the Out-of-Mux stream is sent to the secondary read buffer (1305) by the switch 1301.
FIG. 3 shows schematically the composition of video planes in known Blu-Ray (BD) systems.

As shown, two independent full graphics planes (32, 33) for graphics which are composited on the video plane (31) are present. One graphics plane (32) is assigned for substituting applications (Presentation Graphics or Text Subtitles) and the other plane is assigned to interactive applications (33) (HDMV or BD-J mode interactivity graphics).

Returning to FIG. 3, the main video plane (1310) and the presentation (1309) and graphics plane (1308) are supplied by the corresponding decoders, and the three planes are overlaid by an overlay plane 1311 and outputted. FIG. 4 illustrates schematically a graphics processing unit (13) according to the invention. This specific example constitutes an improvement of the known graphics processing unit in BD systems, but the concept described herein is directly applicable to all graphics processing units in video players, as the decoder models for various types of video players are similar.

For clarity, the overlaying of one graphics plane over the main video plane will be discussed, but the concept is directly applicable to overlaying more than one graphics plane.

For 3D video, extra information is needed besides the 2D video that is stored and send to the display in normal Blu-ray movies. For stereoscopic 3D, it is necessary to send both the left view and the right view to the stereoscopic display. The display then uses a certain technique to make sure only the left eye of the viewer sees the left picture and only the right eye sees the right picture. Common techniques to achieve this are shutter glasses or polarized glasses.

Autostereoscopic displays requires a different interface format: the 2D-depth video format. Besides the 2D video, an additional video stream is used to send depth information. The display combines the video stream in the rendering stage and calculates the resulting 3D picture.

For both 3D techniques it is necessary to send the 2 video streams to the display in a certain interface format, which depends on the display type. A possible interface format is sending the frames from both videos into the display. This means that at time T a frame from the first video stream (left or 2D) is send, and at time T+1 a frame from the second video stream (right or depth) is send.

Application formats like Blu-ray format as mentioned above, support overlay graphics on top of the video. Overlay graphics are for example used to display subtitles of create a selection menu. Blu-ray overlay graphics are read from disc (presentation graphics and interactive graphics) or generated in real time (BD-J graphics, OSD displays and text based subtitles).

Outputting the video in a time-sequential interface format greatly effects the performance requirements of drawing routines for the real-time generated overlay graphics, in particular that of BD-J graphics. This is because the graphics plane can no longer simply be composited with the video output, since the video output switches between the two different video streams each frame. As an example, at time T the video plane could contain the 2D view, and at time T+1 the video plane contains accompanying depth information for the frame at time T. The BD-J graphics that need to be composited with the video at time T (the 2D graphics) greatly differ from the BD-J graphics that need to be composited with the video at time T+1 (the depth graphics).

A graphics processing unit, in particular the BD-J drawing is not fast enough to frame accurately update its graphics plane with these different graphics every frame. The solution according to the invention is to implement two buffers in the graphics unit. Each buffer is assigned to one of the output video streams. For example, for 2D+depth drawing, one buffer could be assigned for graphics overlay over the 2D frame and one buffer could be assigned for the graphics overlay over the depth frame. For L+R, similarly, one buffer could be used for graphics overlay over the L frame, and one buffer could be assigned for overlay over the R frame. The advantage of this solution is that the slow graphics are decoupled from the frame accurate overlaying engine, so that the processing requirements are significantly reduces.

In FIG. 4, a Java application 41 running on a Java Virtual machine generating overlay information and sending it to the graphics processing unit (API). It is noted that the source of the overlay information is not important, such overlay information for a graphics plane could be other graphics from disc or OSD (On Screen display) information. The graphics processing unit comprises two buffers 42 and 43. Each buffer communicate with a controller (45), the controller preferably comprising a frame accurate area copier. Timing information is sent from the drawing application (41) and from the video decoder (47) to the to the graphics processing unit. Based on the received timing information, the frame accurate area copier then can composite the correct buffer onto the graphics output plane, according to what video frame is currently being decoded onto the video output plane (this is known by the Time info from video source). By doing this, the frame accurate area copier ensures that the mixer composites the correct BD-J graphics over the video frame that is currently outputted (for 2D+depth this means that the 2D graphics buffer is copied onto the graphics plane when a 2D video frame is decoded, and the depth DOT graphics buffer is copied onto the graphics plane when a depth frame is decoded). For L+R graphics, this ensure that L real time graphics is overlaid over the L frame and the R real time graphics is overlaid over the R frame.

It is to be noted that the invention may be implemented in hardware and/or software, using programmable components. A method for implementing the invention has the processing steps corresponding to the rendering system elucidated with reference to FIG. 1. Although the invention has been mainly explained by embodiments using optical record carriers or the internet, the invention is also suitable for any image processing environment, like authoring software or broadcasting equipment. Further applications include a 3D personal computer [PC] user interface or 3D media center PC, a 3D mobile player and a 3D mobile phone.

It is noted, that in this document the word 'comprising' does not exclude the presence of other elements or steps than those listed and the word 'a' or 'an' preceding an element does not exclude the presence of a plurality of such elements, that any reference signs do not limit the scope of the claims, that the invention may be implemented by means of both hardware and software, and that several 'means' or 'units' may be represented by the same item of hardware or software, and a processor may fulfill the function of one or more units, possibly in cooperation with hardware elements. Further, the invention is not limited to the embodiments, and lies in each and every novel feature or combination of features described above.

A method of decoding and outputting video information suitable for three-dimensional [3D] display, the video information comprising encoded main video information suitable for displaying on a 2D display and encoded additional video information for enabling three-dimensional [3D] display,
the method comprising:
receiving or generating three-dimensional [3D] overlay information to be overlaid over the video information;
buffering a first part of the overlay information to be overlaid over the main video information in a first buffer;
buffering a second part of overlay information to be overlaid over the additional video information in a second buffer;
decoding the main video information and the additional video information and generating as a series of time interleaved video frames, each outputted video frame being either main video frame or additional video frame;
determining a type of an video frame to be outputted being either a main video frame or an additional video frame;
overlaying either first or second part of the overlay information on an video frame to be outputted in agreement with the determined type of frame;
outputting the video frames and the overlaid information.

2. A method according to claim 1 wherein the main video information is a left video frame and the additional video information is a right video frame.

3. A method according to claim 2 wherein the overlay information is real time graphics.

4. A method according to claim 3 wherein the real time graphics is generated by a Java application running on a Java Virtual machine.

5. A method according to claim 3 wherein timing information is used to controlling the overlaying either first or second part of the overlay information on an video frame to be outputted in agreement with the determined type of frame.

6. A method according to claim 1 wherein the additional video information comprised depth information with respect to the video information.

7. A method according to claim 1 wherein the additional video information further comprises depth and occlusion information.

8. A device for decoding and outputting video information suitable for three-dimensional [3D] display, the video information comprising encoded main video information suitable for displaying on a 2D display and encoded additional video information for enabling three-dimensional [3D] display,

- the device comprising:
  input means for receiving three-dimensional [3D] overlay information to be overlaid over the video information or generation means for generating three-di-
mensional [3D] overlay information to be overlaid over the video information;
  a decoder for decoding the main video information and the additional video information, the decoder further adapted to generating as a series of time interleaved video frames, each outputted video frame being either main video frame or additional video frame;
  means for receiving or generating three-dimensional [3D] overlay information to be overlaid over the video information;
  a graphics processing unit comprising a first buffer for buffering a first part of the overlay information to be overlaid over the main video information and a second buffer for buffering a second part of overlay information to be overlaid over the additional video information;
  the graphics processing unit further comprising a controller for determining a type of an video frame to be outputted being either a main video frame or an additional video frame;
  a mixer for overlaying either first or second part of the overlay information on an video frame to be outputted in agreement with the determined type of frame;
  output means for outputting the video frames and the overlaid information.

9. A device according to claim 8 wherein the main video information is a left video frame and the additional video information is a right video frame.

10. A device according to claim 9 wherein the overlay information is real time graphics.

11. A device according to claim 10 wherein the real time graphics is generated by a Java application running on a Java Virtual machine.

12. A device according to claim 11 wherein timing information is used to controlling the overlaying either first or second part of the overlay information on an video frame to be outputted in agreement with the determined type of frame.

13. A device according to claim 8 wherein the additional video information comprised depth information with respect to the video information.

14. A device according to claim 9 wherein the additional video information further comprises depth and occlusion information.

15. A device according to claim 8 wherein the controller is adapted to copy parts of a first overlay frame in the first buffer or parts of a second overlay frame in the second buffer at frame frequency for generating an overlay frame.