



(51) International Patent Classification:

H04N 7/26 (2006.01) *H04N 7/46* (2006.01)
H04N 7/32 (2006.01)

(21) International Application Number:

PCT/GB2013/052390

(22) International Filing Date:

12 September 2013 (12.09.2013)

(25) Filing Language:

English

(26) Publication Language:

English

(30) Priority Data:

1216263.2 12 September 2012 (12.09.2012) GB

(71) Applicant: **THE UNIVERSITY OF WARWICK**
[GB/GB]; Warwick Ventures, University House, Coventry,
West Midlands CV4 8UW (GB).(72) Inventors: **DEBATTISTA, Kurt**; 28 Clarence Mansions,
5 Clarence Terrace, Leamington Spa, Warwickshire CV32
5LD (GB). **SELMANOVIC, Elmedin**; Grbavicka 2A, 71
000 Sarajevo (BA). **CHALMERS, Alan**; 78 Woodcote
Avenue, Kenilworth, Warwickshire CV8 1BE (GB).
BASHFORD-ROGERS, Thomas; 41B New High Street,
Headington, Oxford, Oxfordshire OX3 7A (GB).(74) Agent: **DEHNS**; St Bride's House, 10 Salisbury Square,
London, Greater London EC4Y 8JD (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, DK, DM, DO, DZ, EC, EE, EG, ES, FI, GB, GD, GE, GH, GM, GT, HN, HR, HU, ID, IL, IN, IS, JP, KE, KG, KN, KP, KR, KZ, LA, LC, LK, LR, LS, LT, 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, 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: MULTI-VIEW HIGH DYNAMIC RANGE IMAGING

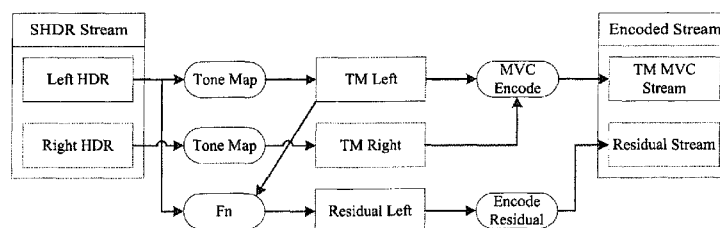


Figure 9

(57) **Abstract:** An encoding method for first and second relatively high dynamic range original images, representing different views of the same scene. The method comprises tone mapping the first relatively high dynamic range original image to create a first relatively low dynamic range image; creating difference data which represents the difference between the first relatively low dynamic range image and the first relatively high dynamic range image; tone mapping the second relatively high dynamic range original image to create a second relatively low dynamic range image; using the first and second relatively low dynamic range images to create encoded image data representing the first and second relatively low dynamic range images; and creating a package comprising the encoded image data and the difference data. In some embodiments the method is carried out on a series of first original images constituting a first relatively high dynamic range original video stream, and a series of second original images constituting a second relatively high dynamic range original video stream.

Multi-View High Dynamic Range Imaging

This invention relates to multi-view high dynamic range imaging, and more particularly to methods of producing compressed multiview, for example stereoscopic, images, either still or moving. Preferred implementations of the invention are concerned with compression techniques which are backwards compatible with established methods such as JPEG or MPEG and/or similar established methods.

Stereoscopic imagery and high dynamic range (HDR) imagery are relatively new imaging methods that both provide advantages over traditional monocular low dynamic range (LDR) imagery. Possibly due to their novelty and perhaps due to technical complications, these methods have not been combined together as yet; however, there is no reason to envisage that they cannot be complementary.

Stereoscopic high dynamic range (SHDR) imagery, as this combination may be termed, has the potential of providing richer fidelity when compared to the traditional LDR methods by combining the advantages of HDR and stereoscopic methods. Many challenges exist for this coupling to prove fruitful and the present invention concerns compression of SHDR images and video for storage or transmission. In the preferred embodiment there are provided compression methods that are not only efficient but also provide backward compatibility with the traditional JPEG LDR and JPEG HDR (as in WARD G.: JPEG-HDR: A backwards-compatible, high dynamic range extension to JPEG. *ACM SIGGRAPH 2006 Courses* (2006), 8. 1, 2) to improve the possibilities of early adoption.

Stereoscopic (frequently called 3D) imaging is a technology that captures and displays images representing the two human eyes. The advantages of such a method rely on enabling or improving the illusion of depth. It can also improve task performance and provide a strong cue for distance judgements. Stereoscopic imagery recently became more widely available as it has now become popular in the consumer market. Consumer products that both capture and display stereoscopic images and video can be purchased at reasonable price and are slowly becoming standard. The entertainment industry is also adopting these techniques with more movies being released in 3D format and video games

- 2 -

allowing 3D visualisation. Even some of the latest smart- phones implement this technology.

5 HDR imagery is a novel technology that has seen several advances over the past few years. HDR enables the capture, storage and display of real world luminance, thus producing images that are more representative of the real world. While a traditional LDR image is limited to around eight exposure stops, i.e. eight fold doubling of the amount of light, HDR does not have such limitations and is able to handle any luminance values that the human eye can see and more. HDR capture
10 of static images has now become widespread, and is available on standard smartphones. HDR video is less common at present although some cameras exist.

There have been a number of ways of storing HDR content, particularly for still images and some of these are discussed later. Display of HDR images has been
15 mostly limited to tone mapping methods that compress the luminance for traditional displays. Whilst there are relatively few HDR displays at present, their numbers will increase. While certain benefits of HDR are straightforward such as the ability to capture and visualise content that would be otherwise over or under exposed, HDR can also benefit stereoscopic content in one of its stronger areas, depth perception,
20 as it can provide further depth information via contrast. SHDR content thus has the potential of achieving the best of both worlds, resulting in images with higher fidelity than current imaging methods. One drawback of SHDR may be the size of images. SHDR images can be quite large if left uncompressed. A raw SHDR image at an HD resolution of 1,920×1,080 can be just short of 48 MB as it stores 96-bits per
25 pixel. With the emergence of higher resolutions, such as super HD's 7,680×4,320 resolution, this would balloon to 759 MB.

Embodiments of the present invention provide the ability to efficiently store SHDR images. A number of novel methods are disclosed which make it possible to store
30 SHDR images more efficiently, in some cases with sizes which are only marginally larger than traditional LDR images or videos. Some embodiments provide the ability of opening any SHDR image in a traditional LDR viewer and in a traditional HDR viewer, and some embodiments are backwards compatible with LDR stereo viewers.

35

- 3 -

Viewed from one aspect the invention provides a method of compressing first and second relatively high dynamic range original images, representing different views of the same scene, comprising the steps of tone mapping the first relatively high dynamic range original image to create a first relatively low dynamic range image; creating difference data which represents the difference between the first relatively low dynamic range image and the first relatively high dynamic range image; tone mapping the second relatively high dynamic range original image to create a second relatively low dynamic range image; using the first and second relatively low dynamic range images to create encoded image data representing the first and second relatively low dynamic range images; and creating a package comprising the encoded image data and the difference data.

The first and second images may represent stereoscopic views, i.e. left and right images of the same scene.

Preferably the method is carried out on a series of first original images constituting a first relatively high dynamic range original video stream, and a series of second original images constituting a second relatively high dynamic range original video stream, the first and second video streams representing different views of the same scene.

In such embodiments there will be provided a method of compressing first and second relatively high dynamic range original video streams, representing different views of the same scene, comprising the steps of tone mapping the first relatively high dynamic range original video stream to create a first relatively low dynamic range video stream; creating difference data which represents the difference between the first relatively low dynamic range video stream and the first relatively high dynamic range video stream; tone mapping the second relatively high dynamic range original video stream to create a second relatively low dynamic range video stream; using the first and second relatively low dynamic range video streams to create encoded video data representing the first and second relatively low dynamic range video streams; and creating a package comprising the encoded video data and the difference data.

Preferably, the first and second relatively low dynamic range video streams are

- 4 -

used to create a tone mapped multiview video encoded stream; and the package comprises the multiview video encoded stream and the difference data.

The multiview video encoded stream may be created by using the first and second relatively low dynamic range video streams together to create a single encoded
5 video stream using for example MVC (Multiview Video Coding) based on the H.264/MPEG-4 AVC, a block-oriented motion-compensation-based codec standard.

Viewed from another aspect the invention provides a decoding method for producing first and second original high dynamic range images each of which
10 represents a different view of the same scene, from a package comprising encoded image data representing first and second relatively low dynamic range tone mapped images and difference data which represents the difference between the first relatively low dynamic range image and the first original relatively high dynamic range image, comprising the steps of decoding the encoded video data to generate
15 first relatively low dynamic range image data representing a first view of the scene, and second relatively low dynamic range image data representing a second view of the scene; using the difference data and the first relatively low dynamic range image data to generate the first high dynamic range image data representing the first view of the scene; using the second relatively low dynamic image data and the
20 difference data to create second difference data; and using the second relatively low dynamic range image data and the second difference data to generate the second high dynamic range image representing a second view of the scene.

The first and second images may represent stereoscopic views, i.e. left and right
25 images of the same scene.

Preferably the method is carried out on a package comprising encoded video stream data representing first and second relatively low dynamic range tone mapped video streams and difference data which represents the difference
30 between the first relatively low dynamic range video stream and a first original relatively high dynamic range video stream, to produce a series of first original images constituting the first relatively high dynamic range original video stream, and a series of second original images constituting a second relatively high dynamic range original video stream, the first and second video streams representing
35 different views of the same scene.

- 5 -

In such a method there will be provided a method of producing first and second original high dynamic range video streams each of which represents a different view of the same scene, from a package comprising a tone mapped encoded stream of relatively low dynamic range video stream data and difference data which represents the difference between one of the multiview streams and the corresponding original relatively high dynamic range video stream data, comprising the steps of decoding the multiview video encoded stream to generate first relatively low dynamic range video stream data representing the first view of the scene, and second relatively low dynamic range video stream data representing the second view of the scene; using the difference data and the first relatively low dynamic range video stream data to generate the first high dynamic range video stream representing the first view of the scene; using the second relatively low dynamic range video stream data and the difference data to create second difference data; and using the second relatively low dynamic range video stream data and the second difference data to generate the second high dynamic range video stream representing a second view of the scene.

The multiview video encoded stream to be decoded may have been created by using the first and second relatively low dynamic range video streams together to create a single encoded video stream using for example MVC (Multiview Video Coding) based on the H.264/MPEG-4 AVC.

Various methods disclosed herein share the same concept of exploiting the natural spatial coherence that exists between the left and right views, so most methods either store both views or as much of one view as is possible and some form of extra information to construct the other view.

When stored as a raw data, considering by way of example only the situation when using RGB colour space, HDR images are generally considered to be composed of three floating point values, one for each of the red, green and blue channels, for a total of 96-bits per pixel (bpp); this results in 24 MB per frame at HD resolution. Similarly large image sizes will result if using other colour spaces such as HSV or any form of spectral data. Due to such prohibitively large sizes, a number of

- 6 -

methods for storing HDR images exist. WARD G.: Real pixels. Graphics Gems II (1991), 80–83. 2 introduced the RGBE format for use with the Radiance software. This format stores RGB values as an 8-bit mantissa per pixel and furthermore stores an 8-bit exponent for a total of 32-bpp. The LogLuv format also proposed by
5 Ward stores images by storing colour and luminance separately. LogLuv supports two formats, a 24-bpp format and a 32-bpp format. The 24-bpp stores 10 bits for the luminance channel and 14 bits for the colour channels, but only supports just short of 5 orders of luminance magnitude. The 32-bpp format can achieve 38 orders of magnitude with 16 bits for chroma and 16 for luminance. The OpenEXR format is a
10 common HDR format used by the entertainment industry, and stores each of the channels as 16-bit half precision floating point values for a total of 48-bpp. It is frequently further compressed using lossless methods.

The above methods are considered storage methods for representing HDR.
15 Another set of HDR imaging method formats are more akin to traditional compression. In particular, there are methods which are compatible with JPEG and JPEG2000. Ward and Simmons (WARD G., SIMMONS M.: Subband encoding of high dynamic range imagery. In Proceedings of the 1st Symposium on Applied perception in graphics and visualization - APGV '04 (New York, New York, USA, Aug. 2004), ACM Press, p. 83. 2; and WARD G.: JPEG-HDR: A backwards-compatible, high dynamic range extension to JPEG. ACM SIGGRAPH 2006
20 Courses (2006), 8. 1, 2) presented JPEG-HDR, an HDR compression method that is backward compatible with the JPEG format. This method stores a tone mapped version of the HDR image which is encoded using JPEG and the extra information used to reconstruct the HDR image, termed the ratio image, is stored in a sub-band
25 of the JPEG format. The ratio image can be downsampled significantly as the human visual system is not very sensitive to low frequency changes in luminance. When opened by a traditional JPEG viewer the sub-band is ignored and the tone mapped image can be seen. The decoding process recovers the missing
30 information from the tone mapped image from the sub-band image (which is upsampled back to the original size) to recover the HDR image. Xu et al. presented HDR-JPEG2000 (XU R., PATTANAIK S., HUGHES C.: High-dynamic-range still-image encoding in jpeg 2000. Computer Graphics and Applications, IEEE 25, 6 (2005), 57–64. 2). HDR-JPEG2000 transforms the HDR data into the 16-bit
35 unsigned shorts supported natively by JPEG2000. Okuda and Adami (OKUDA M.,

ADAMI N.: Two-layer coding algorithm for high dynamic range images based on luminance compensation. Journal of Visual Communication and Image Representation 18, 5 (Oct. 2007), 377–386. 2) presented a method similar to JPEG-HDR which includes wavelet compression for the residuals and minimisation for the tone mapping parameters.

Stereoscopic images, in their most basic form are stored as two separate images. Perhaps, the most popular stereo image format is the JPEG stereo (JPS) which supports a number of formats, storing images side by side, potentially at half the resolution or interleaved. A good overview of different techniques for stereoscopic content, focusing on stereoscopic video is provided by Smolic et al. (SMOLIC A., MUELLER K., STEFANOSKI N.: Coding Algorithms for 3DTV - A Survey. Circuits and Systems for Video Technology, IEEE Transactions on 17, 11 (2007), 1606–1621. 2).

In embodiments of the present invention, as with the Ward and Simmons approach the JFIF (JPEG File Interchange Format) wrapper is used as a format for storing all the extra data that the SHDR methods produce. However, other formats that allow multiple streams of data or metadata to be stored can also equally be used. The main entry is used to store a tone mapped version of one of the stereo pairs (and for one embodiment a side by side version). This format provides storage channels for metadata. This is used to store the additional information such as ratio images, disparity maps and motion compensation information, depending on the technique proposed in order to restore the full SHDR content, as shall be discussed below.

While, the number (16) of metadata channels and size (64 KB) are limited, they are sufficient for the proposed methods. Furthermore, if required, this limitation can be overcome by using more storage channels which have the same identifier. Some embodiments aim to achieve a balance between image size, quality and backward compatibility. For videos there could be stored separate streams of data within a single media container, such as AVI.

In embodiments of the invention discussed below, JPEG and MPEG are the coding methods used but the invention is applicable to other codecs. Where there are references to JPEG or MPEG below they extend to other still image or video codecs.

A first technique aims to store images together before or after the compression process. Even though the preferred embodiment involves stereo, multi-view is achieved simply by increasing the number of images stored. While in the
5 description of an embodiment of this technique images are stored side-by-side, the invention includes storing images in any predetermined spatiotemporal pattern which is used for reconstruction. For example images/frames can be stored next to each other horizontally (as in the SBS embodiment disclosed below), vertically (over-under), one after the other (time sequence) or interleaved in blocks of pixels
10 (which could be of different sizes and even generated dynamically).

A first embodiment of this technique (a "Side-by-Side" or SBS method) aims to preserve quality of the original image and minimise data loss at the expense of larger sizes compared to the other methods. It is also a good foundation for further
15 examination as it provides an initial reference point in terms of quality which other approaches should try to attain. Both of the images in this case are coded using JPEG-HDR so that the quality of the resorted image is kept high and further file size reduction is not considered.

20 This method starts by appending the right HDR image of a stereo pair to the left one. The result is a single side-by-side HDR image. This image is then compressed using JPEG-HDR. The resulting size is almost equivalent to compressing each image separately, and the fact that there is a large amount of correspondence between the stereo pair is not taken advantage of. Alternatively, each image could
25 be compressed first using JPEG-HDR, and then the results would get formatted as explained below.

There are two ways of formatting the output depending on the viewer the image is aimed for. Data from the second image can be put into JPEG sub-bands making it
30 available for traditional monocular image viewers which would show only one of the image pair. Alternatively, the tone-mapped images can be left side-by-side and saved as a stereo JPEG (JPS) which can be then viewable in LDR stereo viewers. Leaving all or some (more than two) images like this in the multi-view case would not be compatible with LDR stereo viewers. This variant can also be opened in
35 traditional viewers but all the views would be displayed. While such behaviour

- 9 -

provides at least some insight into the content of the file, it may not be desirable for the user.

5 SBSV (side by side video) SHDR MPEG stores double the number of streams, the one half containing streams of tone mapped images and the other streams containing ratio information. Encoding of the tone mapped streams will be conducted using MPEG. For the ratio streams, MPEG or other lossy or lossless methods can be used. The ratio streams can also be decreased in size before being compressed using filtering methods such as bilateral filters. This method is
10 backward compatible with both stereo MPEG and HDR MPEG and consequently with traditional MPEG also.

Another technique disclosed is a "Half Side-by-Side" (HSBS) Method. In this the pair of images is put side by side; the horizontal resolution of each is halved such
15 that both can fit in the space of a single image. All the alternative arrangements of the images discussed for the SBS method apply here. In this technique, the coding process involves resizing the images of the HDR pair as described and proceeds in the same manner as SBS. When the image is being decompressed images are resized back up again. Compared to the SBS method the image size is roughly
20 halved but so is its resolution. This method is backward compatible with LDR stereo viewers. Images could be resized using any scaling factor, and not only 0.5.

In a technique of HSBSV (Half Side by Side Video) there is stored one stream composed of side by side tone mapped images of the video stored at half vertical or
25 horizontal resolution each and a second data stream with side by side ratio information also at the same half resolution. It is also possible to store interlaced images in a side by side method by storing each line from each image alternately, or any of the image arrangements discussed earlier. The tone mapped image is encoded using MPEG and similar methods to what are described above are used
30 for the ratio stream. HSBSV is backward compatible with stereo MPEG. It can also be viewed on traditional LDR displays but the images will be displayed side by side.

In an alternative method there is exploited the correlation between the left and right views and the correspondence between most of the pixels. The image that
35 represents these correspondences is a disparity map. The precise maps can be

obtained by using specialist equipment while taking stereoscopic photographs or they can be provided by the rendering software in case of computer generated (CG) images. However, for the majority of stereo images, disparity data is not available and needs to be calculated using the stereo pair. This problem has been
5 researched by previous publications. Scharstein et al. (SCHARSTEIN D., SZELISKI R., ZABIH R.: A taxonomy and evaluation of dense two-frame stereo correspondence algorithms. Proceedings IEEE Workshop on Stereo and Multi-Baseline Vision, 1 (2001), 131–140. 3) provide a taxonomy of stereo correspondence algorithms. They also created a test bed which evaluates the
10 performance of more than one hundred of these algorithms.

Any disparity map can be used with this method (calculated or captured); however the smooth low frequency ones are preferred because of better compression rates. In some embodiments the disparity maps are obtained employing techniques
15 suggested by Mei et al. (MEI X., CUI C., SUN X., ZHOU M., WANG Q., WANG H.: On Building an Accurate Stereo Matching System on Graphics Hardware. In Computer Vision Workshops (ICCV Workshops) (2011), pp. 467—474. 3) which is named AD-census (ADC). This method utilises the Graphics Processing Unit (GPU) during calculations leading to fast performance, and also the technique produces a
20 low number of errors.

An "Image Plus Disparity Video" (IPDV) method stores sources of data (in the preferred embodiment, at least three streams, although all the data can be stored in one or more streams arranged in any predefined manner as explained for the case
25 of the SBS method) containing the tone mapped sequence representing one of the eye sequences, the ratio sequence of the same eye and the disparity. The disparity can be computed using any pixel or stereo correspondence methods or calculated as a depth map from hardware or computed using software. The tone mapped and ratio sequences are stored as above and the disparity method can be compressed
30 using lossy or lossless compression, depending on the quality required.

Another method disclosed is an "Image Plus Disparity with Corrections" (IPDC) method. During testing of the IPD method it has been observed that some occluded regions did not restore well during image warping stage. Here edges were
35 misplaced and a number of cases had major offset issues. The cause of such

problems was that disparity maps were smoothed and background pixels that the restored image required, were occluded by the foreground objects in the available image. This resulted, for example, in those foreground objects being warped to wrong positions and some of the objects being perceived at incorrect depths.

5

An alternative to the method used for IPD is to use a disparity map that maps only the closeness of the values, such as the sum of absolute differences (SAD) method (CYGANNEK B., SIEBERT J. P.: An Introduction to 3D Computer Vision Techniques and Algorithms. John Wiley & Sons, Ltd, Chichester, UK, Jan. 2009. 4). Such
10 disparity maps do without the smoothness condition that maps such as AD-census (ADC) use and are therefore higher frequency. The high frequency means that these disparity maps would not compress well if used by themselves. A method disclosed herein avoids the problems of both these methods by combining them and using SAD or other similar matching methods in regions with large differences
15 only.

A further method disclosed is based on the observation that views differ only by the camera position, which is not dissimilar from the temporal motion between subsequent frames in videos, and may thus be termed a "Motion Compensation"
20 (MC) method. Known video coders are used to compress the SHDR image. Two views of a stereo pair (or multiple views in case of multiview) are treated as consecutive frames in the video which is then processed. Any video coder or camera motion compensation can be used here. In some embodiments H.264 codec is used.

25

"Motion Compensation Video" (MCV) uses MPEG (or similar codecs) not only for compressing over consecutive frames but also for adjacent views, similar to MC for still images. Using this method tone mapped streams are stored, one for each camera position; the first is compressed traditionally but the remainder use motion
30 compensation information from the first also. Further streams are stored with the ratio sequence information, possibly using a similar method to the first. This method is backward compatible with HDR MPEG and traditional MPEG and also with stereo MPEG (in the case of two streams). It is possible to compress the first eye information by using the information in both sequences, such as is used by Multi
35 View Codec in H.264, in which case the backward compatibility with HDR MPEG

and traditional MPEG is compromised but it remains compatible with versions of stereo MPEG.

5 In embodiments of the invention, it is also possible to store residuals (computed via subtraction or division from the original streams) as a separate stream to improve the quality of the image. Any tone mapping and inverse tone mapping methods can be applied to these methods and would not affect the quality. In alternative embodiments of the invention there is stored one tone mapped stream for one eye, the other eye, and only one ratio stream. The other is reconstructed automatically
10 as part of the decoding process.

In some embodiments of the invention the methods can be expanded for multiview HDR video by storing more than a single stream per view or using disparity similar to IPDV.

15 In embodiments of the invention, there is provided Stereoscopic High Dynamic Range (SHDR) Imagery which is a novel technique that combines high dynamic range imaging and stereoscopy. Stereoscopic imaging captures two images representing the views of both eyes and allows for better depth perception. High
20 dynamic range (HDR) imaging is an emerging technology which allows the capture, storage and display of real world lighting as opposed to traditional imagery - which only captures a restricted range of light due to limitation in hardware capture and displays. HDR provides better contrast and more natural looking scenes. One of the main challenges that needs to be overcome for SHDR to be successful is an
25 efficient storage format that compresses the very large sizes obtained by SHDR if left uncompressed; stereoscopic imaging requires the storage of two images and uncompressed HDR requires the storage of a floating point value per colour channel per pixel. In this specification there are presented a number of SHDR compression methods that are backward compatible with traditional JPEG, stereo
30 JPEG and JPEG-HDR, traditional MPEG, stereo MPEG and MPEG-HDR. The methods of some embodiments can encode SHDR content to little more than that of a traditional LDR image and the backward compatibility property encourages early adopters to adopt the format since their content will still be viewable by any of the legacy viewers. The methods can also be used with multi-view images and video.

Some embodiments of the invention will now be described in more detail and with reference to the accompanying drawings, in which:

5 Figure 1 illustrates a process for encoding a stereoscopic High Dynamic Range image using an Image Plus Disparity (IPD) technique;

Figure 2 illustrates a process for decoding the image using the Image Plus Disparity (IPD) technique;

10 Figure 3 illustrates part of a process for encoding a stereoscopic High Dynamic Range image using an Image Plus Disparity with Corrections (IPDC) technique;

Figure 4 illustrates a process for encoding a stereoscopic High Dynamic Range image using a Motion Compensation technique;

15 Figure 5 illustrates a process for decoding a stereoscopic High Dynamic Range image using the using the Motion Compensation technique;

20 Figure 6 illustrates a process for encoding a stereoscopic High Dynamic Range video stream using an Image Plus Disparity (IPD) technique;

Figure 7 illustrates part of a process for encoding a stereoscopic High Dynamic Range video stream using an Image Plus Disparity with Corrections (IPDC) technique.

25 Figure 8 illustrates a process for decoding a stereoscopic High Dynamic Range video stream using an Image Plus Disparity (IPD) technique (or an Image Plus Disparity with Corrections (IPDC) technique):

30 Figure 9 illustrates a process for encoding a stereoscopic High Dynamic Range video stream using a Multiview Video Coding technique;

Figure 10 illustrates a process for decoding a stereoscopic High Dynamic Range video stream using the Multiview Video Coding technique;

35

- 14 -

Figure 11 illustrates a process for encoding a stereoscopic High Dynamic Range video stream using a two stream technique;

5 Figure 12 illustrates a process for decoding a stereoscopic High Dynamic Range video stream using the two stream technique;

Figure 13 illustrates a process for encoding a stereoscopic High Dynamic Range video stream using a Side by Side technique;

10 Figure 14 illustrates a process for decoding a stereoscopic High Dynamic Range video stream using the Side by Side technique;

Figure 15 illustrates a process for encoding a stereoscopic High Dynamic Range video stream using a Half Side by Side technique; and

15

Figure 16 illustrates a process for decoding a stereoscopic High Dynamic Range video stream using the Half Side by Side technique.

20 An encoding process using an IPD method is shown in Figure 1. It includes known HDR JPEG compression to generate a tone mapped image from one of a left and right image pair, in this case the left image, and a sub-band ratio image, which effectively will permit the original HDR left image to be retrieved from the tone mapped image if the equipment reading the coded image is compatible with HDR images. The other image of the pair, in this case the right image, is used together
25 with the first image to generate a disparity map. This is compressed by lossless LZW and Huffman Coding, or another suitable method. This disparity data will be used to construct the right image from the left image, in stereoscopic compatible equipment. The disparity data is then formatted with the tone mapped left image and the ratio image to create the final coded image. Both the ratio image and the
30 disparity data are in sub-bands.

Figure 2 shows how the image is decoded. In equipment which is compatible with stereoscopic HDR images, the ratio image is used together with the tone mapped image to create a restored HDR left image. The disparity data is decoded according
35 to LZW and Huffman decoding to create a disparity map, which is used to warp the

- 15 -

restored HDR left image data to create a restored HDR right image. This technique is therefore backwards compatible as either the ratio image data or the disparity data, or both can be used, or just the tone mapped image, depending on the capabilities of the display equipment.

5

The output file size is not fixed but it is a fraction of the original image and can easily fit in JPEG sub-bands.

For improved results this method can be modified by using an IPDC method, in which the disparity map is generated differently. This process is shown in Figure 3. A high frequency disparity map is obtained from the original HDR left and right images using the SAD method or any other similar method. A low frequency disparity map is also obtained from the original HDR left and right images using the ADC method or other correspondence methods. One of the pairs of images, in this case the right image, is warped using the low frequency disparity map to create a restored version of the other original image, in this case the left image. This is then used in conjunction with the original left image, for example by division, to create data representing the difference between the restored image and the original image, which may be termed a ratio image. Differences above an empirically obtained threshold are identified. Those pixels on the ADC disparity map are updated with ones from the high frequency map, which was obtained using SAD or any other similar method, to create a final disparity map. The rest of the coding process is identical to IPD one. Decoding is the same as in the IPD case.

Figure 4 illustrate an MC process for creating a coded image from an original pair of left and right stereoscopic HDR images. Encoding starts by compressing the left and right images separately using JPEG-HDR. This creates a pair of tone mapped images, each with a sub-band ratio image. The two tone-mapped images are merged to create a two frame video, which is processed by the video encoder. One is treated as the key frame, in this case the left image. The other image is used to create predicted frame data, in this case the right image. In the case of H.264 encoding the second frame corresponds to the p-frame data. The left image is left stored in the original JPEG-HDR format and is not encoded in any other way, as are the sub-band ratio images for both the left and right images. The extracted second frame data depends on the compression quality used but for moderate

compression values it is reasonably small and can fit in additional JPEG sub-bands together with the JPEG-HDR ratio images for both frames.

The decoding process is the inverse of the coding technique as shown in Figure 5.

5 A video file consisting of two frames is generated using the tone-mapped image of the left frame which is appended to the second frame data. This video is then decoded which provides the second tone mapped image for the reconstruction using JPEG-HDR. A standard viewer would only open the stored JPEG and a JPEG-HDR viewer will open only the HDR image of the stored view.

10

Figure 6 is similar to the IPD embodiment of Figure 1, but shows the encoding of left and right video streams. Tone mapping is used to create a tone mapped left HDR video data. The tone mapped stream is used together with the original left stream to create residual data representing the difference between the original and the tone mapped stream, and this is encoded. The right HDR video stream is used together with the original left HDR stream, to create disparity data, which is also encoded. The resultant encoded SHDR stream comprises the encoded tone mapped left HDR data, the encoded residual data and the encoded disparity data.

15

20 Figure 7 corresponds to the process of generating disparity data for an IPDC embodiment, described with reference to Figure 3, but for left and right HDR video streams, and thus shows in detail the steps marked by the dotted lines on Figure 6.

Figure 8 shows the decoding method for an IPD or IPDC encoded video stream.

25

The tone mapped stream is used together with decoded residual data to create decoded HDR left image data. Decoded disparity data is used together with the decoded left HDR data to create decoded right HDR data.

Figure 9 shows an encoding arrangement for a stereoscopic video stream with left and right HDR channels. The left and right HDR stream are both tone mapped and then are used together to create a single encoded video stream using for example MVC (Multiview Video Coding) based on the H.264/MPEG-4 AVC, a block-oriented motion-compensation-based codec standard. The tone mapped version of one stream, in this case the left, is compared with the original HDR stream, to create

30

residual data which then encoded so that the final encoded stream includes both the tone mapped video data and the residual data.

The decoding process is shown in Figure 10. The tone mapped stream is decoded to create a left and a right tone mapped stream. The left tone mapped stream is then used together with the residual data stream to create the left HDR stream. The right tone mapped stream is first used with the residual data, which come from the left stream, to create residual right data. The right residual data is used together with the right tone mapped stream to create the right HDR stream.

Figure 11 shows an encoding arrangement for a stereoscopic video stream with left and right HDR channels. The left and right HDR stream are both tone mapped and then are used together to create an encoded video stream with separate left and right tone mapped data. The tone mapped version of one stream, in this case the left, is compared with the original HDR stream, to create residual data which then encoded so that the final encoded stream includes both the tone mapped video data and the residual data.

The decoding process is shown in Figure 12. The data is decoded to create a left and a right tone mapped stream. The left tone mapped stream is then used together with the residual data stream to create the left HDR stream. The right tone mapped stream is first used with the residual data, which come from the left stream, to create residual right data. The right residual data is used together with the right tone mapped stream to create the right HDR stream.

Figure 13 shows a Side by Side (SBS) encoding method for a stereoscopic video stream with left and right HDR channels. Each channel is tone mapped, and for each channel the tone mapped stream is compared with the original stream to create a residual stream. One tone mapped stream is appended to the other, and one residual stream is appended to the other, so that the resultant encoded stream contains a single tone mapped stream and a single residual stream.

The decoding method is shown in Figure 14. The tone mapped stream is separated into left and right tone mapped streams, and the residual stream is separated into left and right residual streams. Each tone mapped stream is used with its residual stream, to create decoded left and right HDR streams.

Figure 15 shows a Half Side by Side (HSBS) encoding method for a stereoscopic video stream with left and right HDR channels. The two channels are resized and one is appended to the other to create an HSBS HDR stream. This is then tone mapped to create a tone mapped HSBS stream, and a residual HSBS stream which are combined. Decoding is shown in Figure 16. The tone mapped HSBS stream and the residual HSBS stream are combined to create an HDR HSBS stream. The left and right images are separated and resized to create left and right HDR channels again.

A number of images were used to evaluate and test the methods described, for static images. All had been captured at an HD resolution of 1920×1080 . A variety of scenes was chosen which differed in dynamic range, depth, frequency, amount of noise and contrast. Some scenes were computer generated. When coding JPEG- HDR images it was desired to preserve the quality so it was set to 95. The disparity maps for IPD method were generated using default settings suggested. The SAD method used in IPDC had the window set to size 3.

A compatibility table of all methods is given below in Table 1. Please note that the SBS method is either compatible with a traditional LDR viewer or LDR stereo viewer depending on how the second frame data is formatted.

Table 1: Compatibility.

		SBS	HSBS	IPD	IPDC	MC
5	Mono LDR	✓*	✗	✓	✓	✓
	Stereo LDR	✓*	✓	✗	✗	✗
	Mono HDR	✓	✓	✓	✓	✓
10	Stereo HDR	✓	✓	✓	✓	✓

* depending on storage either Mono or Stereo LDR - not both

15 A small segment of a scene was studied to examine the differences between the methods in detail. All images were tone mapped so some of the artefacts observed might have been as a result of the tone mapping. As expected, the SBS method was the one which is the most similar to the original image as there were no prominent artefacts. A loss of resolution was visible in the HSBS image which is blurrier than the rest. Mistakes made during image warping due to the occlusion were present with the IPD method. IPDC managed to fix this problem to an extent but small mistakes were noticeable, such as the presence of grey values instead of original yellowish ones. Images produced by the MC method looked very similar to the original but on closer inspection some "blocky" artefacts due to MPEG compression were visible.

In order to evaluate the performance of the methods the following quantities were measured: file size of the compressed SHDR image (in kilobytes), normalised root-mean-square error (NRMSE) and peak signal-to-noise ratio (PSNR) between the decoded images and the original HDR images. NRMSE and PSNR were measured for left and right views separately and then averaged. File sizes are shown in Table 2. The last table column contains the sizes of a single LDR image of the same scene which is JPEG encoded using the same quality settings as the proposed methods. NRMSE, expressed as a percentage, is presented in Table 3 and PSNR is shown in Table 4. Average values for all of the scenes are shown in the last row. Table 5 provides a summary of the results. The compression ratio gives the average compression ratio compared to a raw HD stereo image. The "quality × compression" measure is a multiplication of the average image size for each method multiplied by the average NRMSE. This value is presented to give an idea

of what gives the best “bang for the buck” but should not be taken as a definitive measure as the different methods have distinct qualities. For "quality × compression" the smaller values are considered better.

5

Table 2: Sizes of compressed images (in KB) and single LDR image of the same scene shown for comparison.

10

15

20

Img	SBS	HSBS	IPD	IPDC	MC	LDR
1	1087	561	601	717	683	476
2	1088	558	613	728	671	482
3	868	459	488	616	544	370
4	897	493	494	589	557	384
5	891	506	474	514	549	383
6	1697	899	956	1392	1104	794
7	1021	512	572	704	616	449
8	784	417	436	458	484	328
9	1648	837	890	1119	977	769
10	1215	617	629	682	707	544
11	1087	549	563	577	639	487
12	436	267	245	263	285	168
13	1326	689	728	877	803	601
14	1004	530	549	589	617	439
15	888	455	549	656	543	362
16	1457	744	786	1024	913	685
17	1456	793	849	1132	883	676
18	954	491	529	603	590	412
19	1026	537	569	617	640	449
Avg	1096	574	606	729	674	487

Table 3: NRMSE measurements (%)

Img	SBS	HSBS	IPD	IPDC	MC
1	0.12	0.23	0.40	0.30	0.13
2	0.11	0.24	0.25	0.18	0.12
3	0.08	0.11	0.20	0.19	0.09
4	0.08	0.17	0.11	0.11	0.09
5	0.37	0.41	0.56	0.48	0.60
6	0.32	0.59	0.77	0.50	0.44
7	0.08	0.26	0.23	0.16	0.10
8	0.17	0.40	0.30	0.25	0.21
9	0.43	0.71	0.76	0.61	0.50
10	0.56	0.89	0.97	0.67	0.73
11	0.94	1.23	1.06	1.04	1.37
12	0.18	0.29	0.21	0.20	0.30
13	0.21	0.42	0.44	0.38	0.26
14	0.18	0.25	0.35	0.24	0.18
15	0.30	0.34	0.57	0.39	0.32
16	0.26	0.36	0.49	0.42	0.31
17	0.37	0.75	1.27	0.72	0.53
18	0.08	0.15	0.13	0.12	0.09
19	0.20	0.30	0.32	0.28	0.21
Average	0.27	0.43	0.49	0.38	0.35

Table 4: PSNR measurements

Img	SBS	HSBS	IPD	IPDC	MC
1	58	53	51	53	58
2	59	52	54	56	59
3	62	59	55	56	61
4	62	55	59	59	61
5	49	48	46	47	45
6	50	45	44	47	47
7	62	52	55	57	60
8	56	48	52	53	54
9	47	43	43	45	46
10	45	41	41	44	43
11	41	38	40	40	38
12	55	51	54	54	51
13	54	48	49	49	52
14	55	52	50	53	55
15	51	49	46	48	50
16	52	49	47	48	50
17	49	42	41	44	46
18	62	56	59	59	61
19	54	51	51	51	54
Average	54	49	49	51	52

Table 5: Results Summary. Compression Ratio (CR) and Quality $\times$ Compression ($Q \times C$). NRMSE is used for quality and for $Q \times C$ lower figures are better.

5

	SBS	HSBS	IPD	IPDC	MC
CR	44	85	80	67	72
$Q \times C$	291	245	300	278	234

On the whole it is the MC method that achieves the best overall "quality $\times$ compression"; it also has the added advantage that it is backward compatible with JPEG-HDR and JPEG. The HSBS achieves the second best overall score, this is primarily due to it having the highest compression ratio but this comes at the expense of quality, for which it is second from the bottom. It is backward compatible with LDR stereo JPEG but not fully compatible with JPEG-HDR and traditional JPEG; the image shown on a traditional JPEG viewer would be both images side by side. The SBS method is third overall and is backward compatible with all possible formats, however it is the largest in size, which may be to much of a prohibitive obstacle, potentially hampering the uptake of SHDR. The disparity methods IPDC and IPD are second from last and last respectively. However, the results depend on the quality of the disparity maps produced, and as these may become better, as this is an active research area the scope of such disparity based methods is likely to improve. In addition these methods are backward compatible with traditional JPEG and HDR JPEG and produce relatively small image sizes not much larger than JPEG-HDR images.

In this specification there is introduced SHDR as a combination of stereoscopic and high dynamic range imaging. There are proposed a number of SHDR compression methods based on the JPEG standard and the JPEG HDR compression method, and based on the MPEG standard. The MC, SBS, IPDC and IPD methods are all backward compatible with traditional JPEG and JPEG-HDR, and while the results demonstrate a better overall performance for MC, the disparity based methods may improve with time as disparity methods improve. The SBS and HSBS methods are backward compatible with stereo JPEG and could be used on occasions where this is sufficient.

The backward compatibility constraint that has been imposed may be removed, although this may be better to use once the technology is widely adopted and if great improvements are possible. The methods have been evaluated with NRMSE and PSNR, since no perceptual metric currently exists for evaluating SHDR.

5

It will be appreciated that the invention is not limited to JPEG and MPEG, and the encoded files could be made compatible or backwards compatible with equipment design to handle images or videos which have been encoded using other codecs.

10

In general, in some embodiments of the invention there is provided a method of encoding a plurality of relatively high dynamic range original images each of which represents a different view of the same scene, to provide relatively low dynamic range image data and difference data which represents the difference between the relatively low dynamic range image data and original relatively high dynamic range image data. In some embodiments of the invention there are two original images, representing a stereoscopic view. The images may be static images or frames of a video stream.

15

20

In general, in some embodiments of the invention there is provided a method of encoding a plurality of relatively high dynamic range original video streams each of which represents a different view of the same scene, to provide relatively low dynamic range video stream data and difference data which represents the difference between the relatively low dynamic range video stream data and original relatively high dynamic range video stream data. In some embodiments of the invention there are two original video streams, representing a stereoscopic view.

25

30

In general, in some embodiments of the invention there is provided a method of decoding a plurality of relatively high dynamic range original images each of which represents a different view of the same scene, from relatively low dynamic range image data and difference data which represents the difference between the relatively low dynamic range image data and original relatively high dynamic range image data. In some embodiments of the invention there are two original images, representing a stereoscopic view. The images may be static images or frames of a video stream.

35

- 24 -

In general, in some embodiments of the invention there is provided a method of a plurality of relatively high dynamic range video streams each of which represents a different view of the same scene, from relatively low dynamic range video stream data and difference data which represents the difference between the relatively low
5 dynamic range video stream data and original relatively high dynamic range video stream data. In some embodiments of the invention there are two original video streams, representing a stereoscopic view. The images may be static images or frames of a video stream.

10 Viewed from another aspect of the invention there is provided a method of compressing first and second relatively high dynamic range original images, representing different views of the same scene, comprising the steps of generating disparity data representing the disparity between the first relatively high dynamic range original image and the second relatively high dynamic range original image;
15 tone mapping the first relatively high dynamic range original image to create a relatively low dynamic range image; creating difference data which represents the difference between the relatively low dynamic range image and the first relatively high dynamic range image; and creating a package comprising the relatively low dynamic range image, the difference data, and the disparity data.

20 The disparity data may be compressed and included in the package in compressed form. The difference data may be compressed and included in the package in compressed form.

25 There may be additional relatively high dynamic range original images representing additional views of the scene. For each additional original image there will be generated respective disparity data representing the differences between said first relatively high dynamic range original image and that additional original image, and the package will comprise disparity data in respect of that additional original image.

30 The images may be still images or image frames of a video stream.

Viewed from another aspect the invention provides a method of decoding the package to generate the first and second relatively high dynamic range images.

35

A preferred arrangement for generating disparity data representing the difference between the first relatively high dynamic range original images and another one of the relatively high dynamic range original images, comprises the steps of:
generating a first, relatively high frequency, disparity map representing the
5 differences between said first relatively high dynamic range original images and said other of the relatively high dynamic range original images; generating a second, relatively low frequency, disparity map representing the differences between said first relatively high dynamic range original image and said other of the relatively high dynamic range original images; using the second disparity map and
10 said other of the relatively high dynamic range original images to create a restored version of said first relatively high dynamic range original image; comparing said first relatively high dynamic range original images and the restored version of said first relatively high dynamic range original images to create ratio data; and using the first disparity map, the second disparity map and the ratio data to generate a
15 combined disparity map.

Viewed from another aspect of the invention there is provided a method of compressing first and second relatively high dynamic range original video streams, representing different views of the same scene, comprising the steps of generating
20 disparity data representing the disparity between the first relatively high dynamic range original video stream and the second relatively high dynamic range original video stream; tone mapping the first relatively high dynamic range original video stream to create a relatively low dynamic range video stream; creating difference data which represents the difference between the relatively low dynamic range
25 video stream and the first relatively high dynamic range video stream; and creating a package comprising the relatively low dynamic range video stream, the difference data, and the disparity data.

The disparity data may be compressed and included in the package in compressed
30 form. The difference data may be compressed and included in the package in compressed form.

Viewed from another aspect the invention provides a method of decoding the package to generate the first and second relatively high dynamic range video
35 streams.

A preferred arrangement for generating disparity data representing the difference between the first relatively high dynamic range original video stream and another one of the relatively high dynamic range original video streams, comprises the steps of: generating a first, relatively high frequency, disparity map representing the differences between said first relatively high dynamic range original stream and said other of the relatively high dynamic range original video stream; generating a second, relatively low frequency, disparity map representing the differences between said first relatively high dynamic range original video stream and said other of the relatively high dynamic range original video streams; using the second disparity map and said other of the relatively high dynamic range original video streams to create a restored version of said first relatively high dynamic range original video stream; comparing said first relatively high dynamic range original video stream and the restored version of said first relatively high dynamic range original video stream to create ratio data; and using the first disparity map, the second disparity map and the ratio data to generate a combined disparity map.

Viewed from another aspect the invention provides a method of compressing first and second relatively high dynamic range original images, representing different views of the same scene, comprising the steps of : tone mapping the first relatively high dynamic range original image to create a first relatively low dynamic range image; creating first difference data which represents the difference between the first relatively low dynamic range image and the first relatively high dynamic range image; tone mapping the second relatively high dynamic range original image to create a second relatively low dynamic range image; creating second difference data which represents the difference between the relatively low dynamic range image and said one of the relatively high dynamic range images; using video encoding of the first relatively low dynamic range image and the second relatively low dynamic range image to create predicted image data in respect of the second relatively low dynamic range image; and creating a package comprising the first relatively low dynamic range image, the first difference data, the predicted image data, and the second difference data.

Viewed from another aspect the invention provides a method of decoding the package to generate the first and second relatively high dynamic range images.

- 27 -

Viewed from another aspect of the invention there is provided a method of compressing first and second relatively high dynamic range original video streams, representing different views of the same scene, comprising the steps of tone mapping the first relatively high dynamic range original video stream to create a first
5 relatively low dynamic range video stream; creating difference data which represents the difference between the first relatively low dynamic range video stream and the first relatively high dynamic range video stream; tone mapping the second relatively high dynamic range original video stream to create a second relatively low dynamic range video stream; using the first and second relatively low
10 dynamic range video streams to create a tone mapped multiview video encoded stream; and creating a package comprising the multiview video encoded stream and the difference data.

The difference data may be compressed and included in the package in
15 compressed form.

Viewed from another aspect the invention provides a method of decoding the package to generate the first and second relatively high dynamic range video streams.

20

Viewed from another aspect the invention provides a method of producing first and second original high dynamic range video streams each of which represents a different view of the same scene, from a package comprising a tone mapped multiview video encoded stream of relatively low dynamic range video stream data and difference data which represents the difference between one of the multiview
25 streams and the corresponding original relatively high dynamic range video stream data, comprising the steps of decoding the multiview video encoded stream to generate first relatively low dynamic range video stream data representing the first view of the scene, and second relatively low dynamic range video stream data
30 representing the second view of the scene; using the difference data and the first relatively low dynamic range video stream data to generate the first high dynamic range video stream representing the first view of the scene; using the second relatively low dynamic range video stream data and the difference data to create second difference data; and using the second relatively low dynamic range video

stream data and the second difference data to generate the second high dynamic range video stream representing a second view of the scene.

Viewed from another aspect of the invention there is provided a method of
5 compressing first and second relatively high dynamic range original video streams, representing different views of the same scene, comprising the steps of tone mapping the first relatively high dynamic range original video stream to create a first relatively low dynamic range video stream; creating difference data which represents the difference between the first relatively low dynamic range video
10 stream and the first relatively high dynamic range video stream; tone mapping the second relatively high dynamic range original video stream to create a second relatively low dynamic range video stream; using the first and second relatively low dynamic range video streams to create tone mapped video encoded streams; and creating a package comprising the encoded streams and the difference data.

15 In a modified form of this arrangement, there may be created second difference data which represents the difference between the second relatively low dynamic range video stream and the second relatively high dynamic range video stream, and the package comprises the encoded streams, the difference data and the
20 second difference data.

The difference data may be compressed and included in the package in compressed form.

25 Viewed from another aspect the invention provides a method of decoding the package to generate the first and second relatively high dynamic range video streams. The method of encoding corresponds to the previous method save that the package includes the separate stream rather than a multiview stream. In the modified form the second relatively high dynamic range video stream is re-created
30 using the second difference data.

Viewed from another aspect the invention provides a method of compressing first and second relatively high dynamic range original video streams, representing different views of the same scene, comprising the steps of tone mapping the first
35 relatively high dynamic range original video stream to create a first relatively low

dynamic range video stream; creating first difference data which represents the difference between the first relatively low dynamic range video stream and the first relatively high dynamic range video stream; tone mapping the second relatively high dynamic range original video stream to create a second relatively low dynamic range video stream; creating second difference data which represents the difference between the second relatively low dynamic range video stream and the second relatively high dynamic range video stream; appending the second relatively low dynamic range video stream to the first relatively low dynamic range video stream to create combined video data; appending the second difference data to the first difference data to create combined difference data; and creating a package comprising the combined video data and the combined difference data.

The difference data may be compressed and included in the package in compressed form.

Viewed from another aspect the invention provides a method of decoding the package to generate the first and second relatively high dynamic range video streams. In such a method the combined video data is separated into first and second relatively low dynamic range video stream; and the combined difference data is separated into first and second difference data. The first and second relatively high dynamic range video streams can then be created.

Viewed from another aspect the invention provides a method of compressing first and second relatively high dynamic range original video streams, representing different views of the same scene, comprising the steps of resizing the first and second relatively high dynamic range original video streams and combining them to create a combined relatively high dynamic range video stream; tone mapping the combined relatively high dynamic range video stream to create a combined relatively low dynamic range video stream; creating difference data which represents the difference between the combined relatively low dynamic range video stream and the combined relatively high dynamic range video stream; and creating a package comprising the combined relatively low dynamic range video stream and the difference data.

- 30 -

The difference data may be compressed and included in the package in compressed form.

5 Viewed from another aspect the invention provides a method of decoding the package to generate the first and second relatively high dynamic range video streams. In such a method the combined relatively low dynamic range video stream is used with the difference data to re-create the combined relatively high dynamic range video stream, and this is then separated into two streams which are re-sized to create the first and second relatively high dynamic range video streams.

10 For all of the methods disclosed herein, the invention provides a method of encoding; a method of decoding; an encoder; a decoder; data processing apparatus configured to carry out the method of encoding; data processing apparatus configured to carry out the method of decoding; computer software for
15 configuring data processing apparatus to carry out the method of encoding; and computer software for configuring data processing apparatus to carry out the method of encoding. The computer program may be in tangible form such as data encoded on a DVD or a solid state memory device, or may be intangible in the form of signals transmitted over a network over the network.

20 In all embodiments of the inventions disclosed herein, the encoding process involves not only tone mapping of an HDR image or a video stream, but also compression of the tone mapped image or video stream. Similarly, decoding involves de-compression of an image or a video stream, and not just reversal of the
25 tone mapping process to provide an HDR image or video stream. Techniques can be applied to still images, video stream or frames of video streams, as appropriate for the method used. Whilst backwards compatibility is an advantage of embodiments described, it is not an absolute requirement of the inventions in their broadest sense.

CLAIMS

1. An encoding method for first and second relatively high dynamic range original images, representing different views of the same scene, comprising the steps of tone mapping the first relatively high dynamic range original image to create a first relatively low dynamic range image; creating difference data which represents the difference between the first relatively low dynamic range image and the first relatively high dynamic range image; tone mapping the second relatively high dynamic range original image to create a second relatively low dynamic range image; using the first and second relatively low dynamic range images to create encoded image data representing the first and second relatively low dynamic range images; and creating a package comprising the encoded image data and the difference data.
2. An encoding method as claimed in claim 1, wherein the method is carried out on a series of first original images constituting a first relatively high dynamic range original video stream, and a series of second original images constituting a second relatively high dynamic range original video stream, the first and second video streams representing different views of the same scene.
3. An encoding method as claimed in claim 2, wherein the method comprises compressing the first and second relatively high dynamic range original video streams, representing different views of the same scene, by the steps of tone mapping the first relatively high dynamic range original video stream to create a first relatively low dynamic range video stream; creating difference data which represents the difference between the first relatively low dynamic range video stream and the first relatively high dynamic range video stream; tone mapping the second relatively high dynamic range original video stream to create a second relatively low dynamic range video stream; using the first and second relatively low dynamic range video streams to create encoded video data representing the first and second relatively low dynamic range video streams; and creating a package comprising the encoded video data and the difference data.
4. An encoding method as claimed in claim 3, wherein the first and second relatively low dynamic range video streams are used to create a tone mapped

- 32 -

multiview video encoded stream; and the package comprises the multiview video encoded stream and the difference data.

- 5 5. An encoding method as claimed in claim 4, wherein the multiview video encoded stream is created by using the first and second relatively low dynamic range video streams together to create a single encoded video stream using MVC (Multiview Video Coding).
- 10 6. An encoding method as claimed in claim 1, wherein the first and second images represent stereoscopic views of the same scene.
7. An encoding method as claimed in any of claims 2 to 5, wherein the first and second video streams represent stereoscopic views of the same scene.
- 15 8. An encoding method as claimed in any preceding claim, wherein the difference data is compressed.
- 20 9. A decoding method for producing first and second original high dynamic range images each of which represents a different view of the same scene, from a package comprising encoded image data representing first and second relatively low dynamic range tone mapped images and difference data which represents the difference between the first relatively low dynamic range image and the first original relatively high dynamic range image, comprising the steps of decoding the encoded video data to generate first relatively low dynamic range image data representing a first view of the scene, and second relatively low dynamic range image data representing a second view of the scene; using the difference data and the first relatively low dynamic range image data to generate the first high dynamic range image data representing the first view of the scene; using the second relatively low dynamic image data and the difference data to create second difference data; and
25 30 using the second relatively low dynamic range image data and the second difference data to generate the second high dynamic range image representing a second view of the scene.
- 35 10. A decoding method as claimed in claim 9, wherein the method is carried out on a package comprising encoded video stream data representing first and second

relatively low dynamic range tone mapped video streams and difference data which represents the difference between the first relatively low dynamic range video stream and a first original relatively high dynamic range video stream, to produce a series of first original images constituting the first relatively high dynamic range original video stream, and a series of second original images constituting a second relatively high dynamic range original video stream, the first and second video streams representing different views of the same scene.

11. A decoding method as claimed in claim 10, wherein the method comprises producing first and second original high dynamic range video streams each of which represents a different view of the same scene, from a package comprising a tone mapped encoded stream of relatively low dynamic range video stream data and difference data which represents the difference between one of the streams and the corresponding original relatively high dynamic range video stream data, comprising the steps of decoding the encoded stream to generate first relatively low dynamic range video stream data representing the first view of the scene, and second relatively low dynamic range video stream data representing the second view of the scene; using the difference data and the first relatively low dynamic range video stream data to generate the first high dynamic range video stream representing the first view of the scene; using the second relatively low dynamic range video stream data and the difference data to create second difference data; and using the second relatively low dynamic range video stream data and the second difference data to generate the second high dynamic range video stream representing a second view of the scene.

12. An decoding method as claimed in claim 11, wherein the encoded stream of relatively low dynamic range video stream data is a multiview video encoded stream.

13. A decoding method as claimed in claim 12, wherein multiview video encoded stream to be decoded has been created by using the first and second relatively low dynamic range video streams together to create a single encoded video stream using MVC (Multiview Video Coding).

14. A decoding method as claimed in claim 9, wherein the first and second

images represent stereoscopic views of the same scene.

15. An decoding method as claimed in any of claims 10 to 13, wherein the first and second video streams represent stereoscopic views of the same scene.

5

16. Data processing apparatus configured to carry out the encoding method of any of claims 1 to 8.

10 17. Computer software carrying instructions which when carried out on data processing apparatus will cause the data processing apparatus to carry out the encoding method of any of claims 1 to 8.

18. Data processing apparatus configured to carry out the decoding method of any of claims 9 to 14.

15

19. Computer software carrying instructions which when carried out on data processing apparatus will cause the data processing apparatus to carry out the decoding method of any of claims 9 to 14.

20 20. A method of compressing first and second relatively high dynamic range original images, representing different views of the same scene, comprising the steps of generating disparity data representing the disparity between the first relatively high dynamic range original image and the second relatively high dynamic range original image; tone mapping the first relatively high dynamic range original image to create a relatively low dynamic range image; creating difference data which represents the difference between the relatively low dynamic range image and the first relatively high dynamic range image; and creating a package comprising the relatively low dynamic range image, the difference data, and the disparity data.

30

21. A method as claimed in claim 19, wherein the generation of disparity data representing the difference between the first relatively high dynamic range original images and another one of the relatively high dynamic range original images, comprises the steps of: generating a first, relatively high frequency, disparity map representing the differences between said first relatively high dynamic range

35

original images and said other of the relatively high dynamic range original images; generating a second, relatively low frequency, disparity map representing the differences between said first relatively high dynamic range original image and said other of the relatively high dynamic range original images; using the second
5 disparity map and said other of the relatively high dynamic range original images to create a restored version of said first relatively high dynamic range original image; comparing said first relatively high dynamic range original images and the restored version of said first relatively high dynamic range original images to create ratio data; and using the first disparity map, the second disparity map and the ratio data
10 to generate a combined disparity map.

22. A method of compressing first and second relatively high dynamic range original video streams, representing different views of the same scene, comprising the steps of generating disparity data representing the disparity between the first
15 relatively high dynamic range original video stream and the second relatively high dynamic range original video stream; tone mapping the first relatively high dynamic range original video stream to create a relatively low dynamic range video stream; creating difference data which represents the difference between the relatively low dynamic range video stream and the first relatively high dynamic range video
20 stream; and creating a package comprising the relatively low dynamic range video stream, the difference data, and the disparity data.

23. A method as claimed in claim 22, wherein the step of generating disparity data representing the difference between the first relatively high dynamic range original video stream and another one of the relatively high dynamic range original
25 video streams, comprises the steps of: generating a first, relatively high frequency, disparity map representing the differences between said first relatively high dynamic range original stream and said other of the relatively high dynamic range original video stream; generating a second, relatively low frequency, disparity map
30 representing the differences between said first relatively high dynamic range original video stream and said other of the relatively high dynamic range original video streams; using the second disparity map and said other of the relatively high dynamic range original video streams to create a restored version of said first relatively high dynamic range original video stream; comparing said first relatively
35 high dynamic range original video stream and the restored version of said first

relatively high dynamic range original video stream to create ratio data; and using the first disparity map, the second disparity map and the ratio data to generate a combined disparity map.

5 24. A method of compressing first and second relatively high dynamic range original images, representing different views of the same scene, comprising the steps of : tone mapping the first relatively high dynamic range original image to create a first relatively low dynamic range image; creating first difference data which represents the difference between the first relatively low dynamic range image and
10 the first relatively high dynamic range image; tone mapping the second relatively high dynamic range original image to create a second relatively low dynamic range image; creating second difference data which represents the difference between the relatively low dynamic range image and said one of the relatively high dynamic range images; using video encoding of the first relatively low dynamic range image
15 and the second relatively low dynamic range image to create predicted image data in respect of the second relatively low dynamic range image; and creating a package comprising the first relatively low dynamic range image, the first difference data, the predicted image data, and the second difference data.

20 25. A method of compressing first and second relatively high dynamic range original video streams, representing different views of the same scene, comprising the steps of tone mapping the first relatively high dynamic range original video stream to create a first relatively low dynamic range video stream; creating difference data which represents the difference between the first relatively low
25 dynamic range video stream and the first relatively high dynamic range video stream; tone mapping the second relatively high dynamic range original video stream to create a second relatively low dynamic range video stream; using the first and second relatively low dynamic range video streams to create a tone mapped multiview video encoded stream; and creating a package comprising the multiview
30 video encoded stream and the difference data.

26. A method of producing first and second original high dynamic range video streams each of which represents a different view of the same scene, from a package comprising a tone mapped multiview video encoded stream of relatively
35 low dynamic range video stream data and difference data which represents the

difference between one of the multiview streams and the corresponding original relatively high dynamic range video stream data, comprising the steps of decoding the multiview video encoded stream to generate first relatively low dynamic range video stream data representing the first view of the scene, and second relatively low dynamic range video stream data representing the second view of the scene; using the difference data and the first relatively low dynamic range video stream data to generate the first high dynamic range video stream representing the first view of the scene; using the second relatively low dynamic range video stream data and the difference data to create second difference data; and using the second relatively low dynamic range video stream data and the second difference data to generate the second high dynamic range video stream representing a second view of the scene.

27. A method of compressing first and second relatively high dynamic range original video streams, representing different views of the same scene, comprising the steps of tone mapping the first relatively high dynamic range original video stream to create a first relatively low dynamic range video stream; creating difference data which represents the difference between the first relatively low dynamic range video stream and the first relatively high dynamic range video stream; tone mapping the second relatively high dynamic range original video stream to create a second relatively low dynamic range video stream; using the first and second relatively low dynamic range video streams to create tone mapped video encoded streams; and creating a package comprising the encoded streams and the difference data.

28. A method as claimed in claim 27, wherein there is created second difference data which represents the difference between the second relatively low dynamic range video stream and the second relatively high dynamic range video stream, and the package comprises the encoded streams, the difference data and the second difference data.

29. A method of compressing first and second relatively high dynamic range original video streams, representing different views of the same scene, comprising the steps of tone mapping the first relatively high dynamic range original video stream to create a first relatively low dynamic range video stream; creating first

difference data which represents the difference between the first relatively low dynamic range video stream and the first relatively high dynamic range video stream; tone mapping the second relatively high dynamic range original video stream to create a second relatively low dynamic range video stream; creating
5 second difference data which represents the difference between the second relatively low dynamic range video stream and the second relatively high dynamic range video stream; appending the second relatively low dynamic range video stream to the first relatively low dynamic range video stream to create combined video data; appending the second difference data to the first difference data to
10 create combined difference data; and creating a package comprising the combined video data and the combined difference data.

30. A method of compressing first and second relatively high dynamic range original video streams, representing different views of the same scene, comprising
15 the steps of resizing the first and second relatively high dynamic range original video streams and combining them to create a combined relatively high dynamic range video stream; tone mapping the combined relatively high dynamic range video stream to create a combined relatively low dynamic range video stream; creating difference data which represents the difference between the combined
20 relatively low dynamic range video stream and the combined relatively high dynamic range video stream; and creating a package comprising the combined relatively low dynamic range video stream and the difference data.

31. A method of decoding a package created by a method as claimed in any of
25 claims 20 to 30.

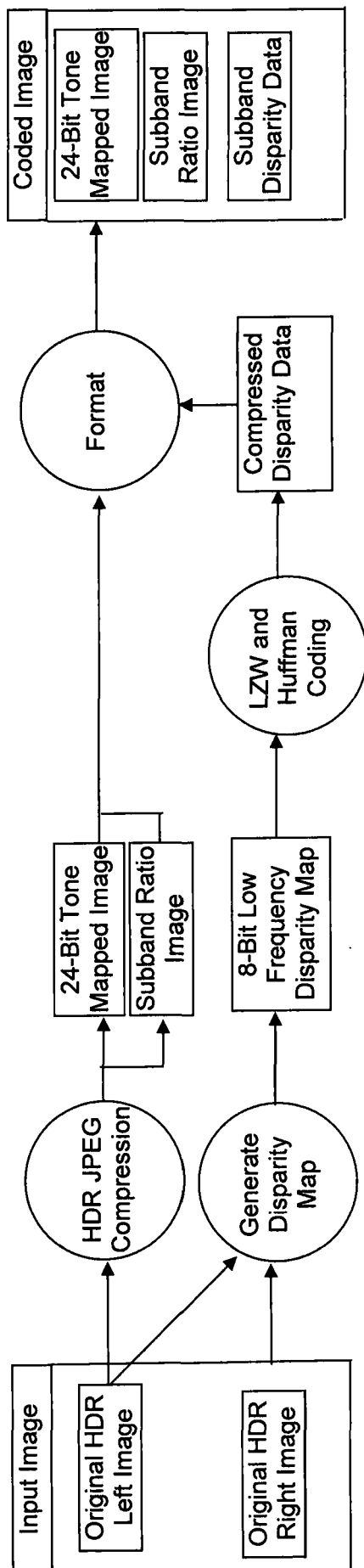


Figure 1: IPD encoding

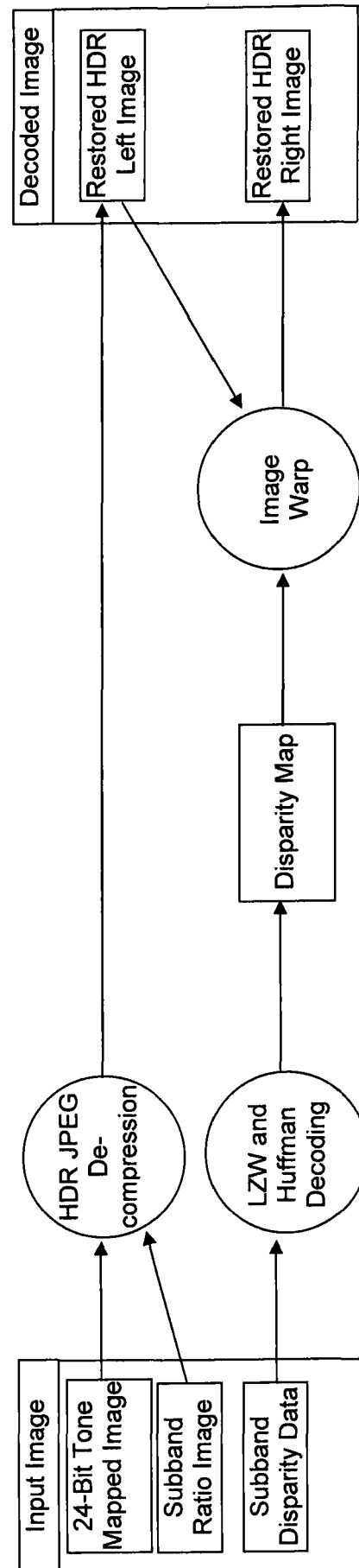


Figure 2: IPD decoding

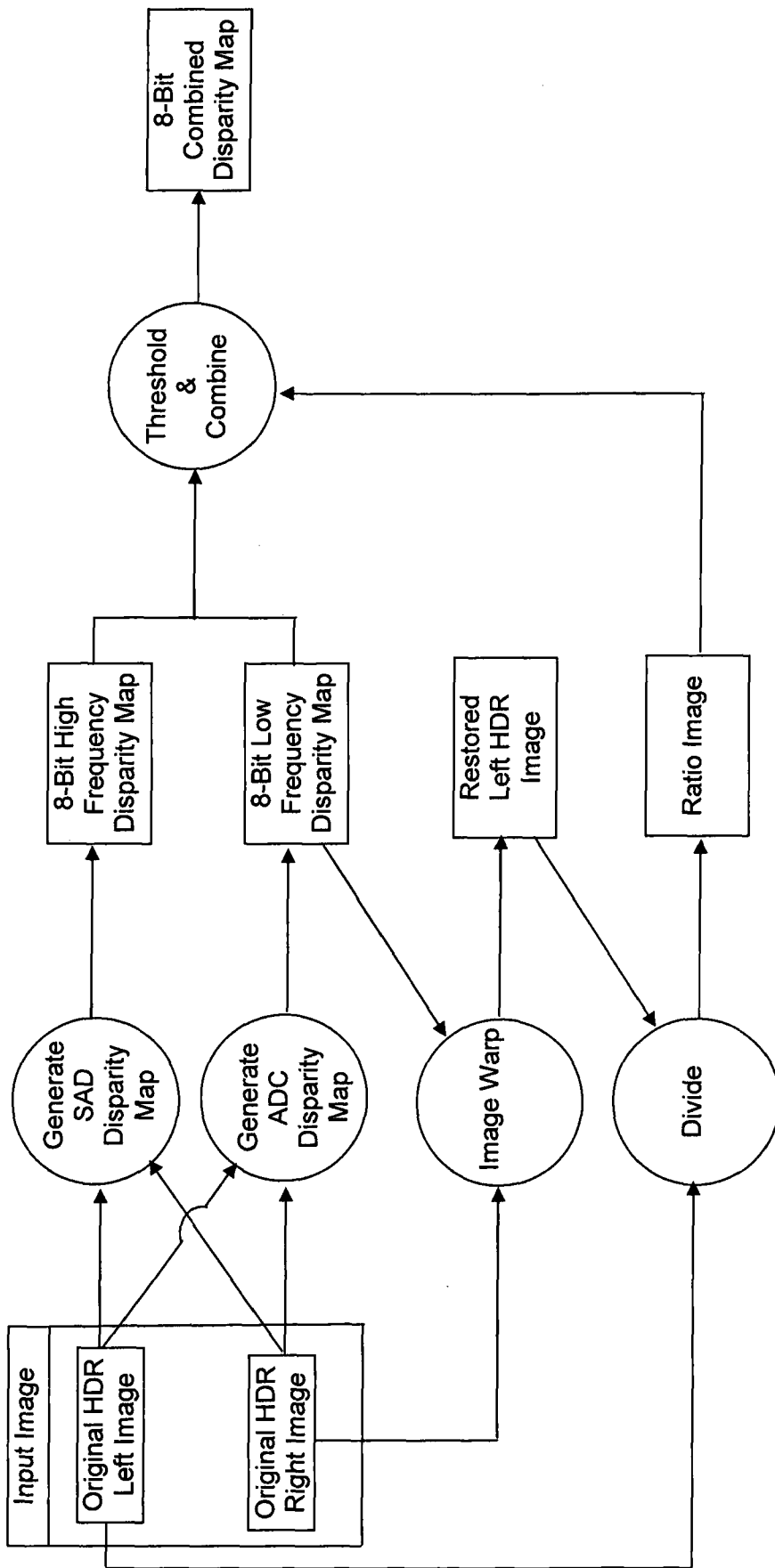


Figure 3: IPDC disparity generation

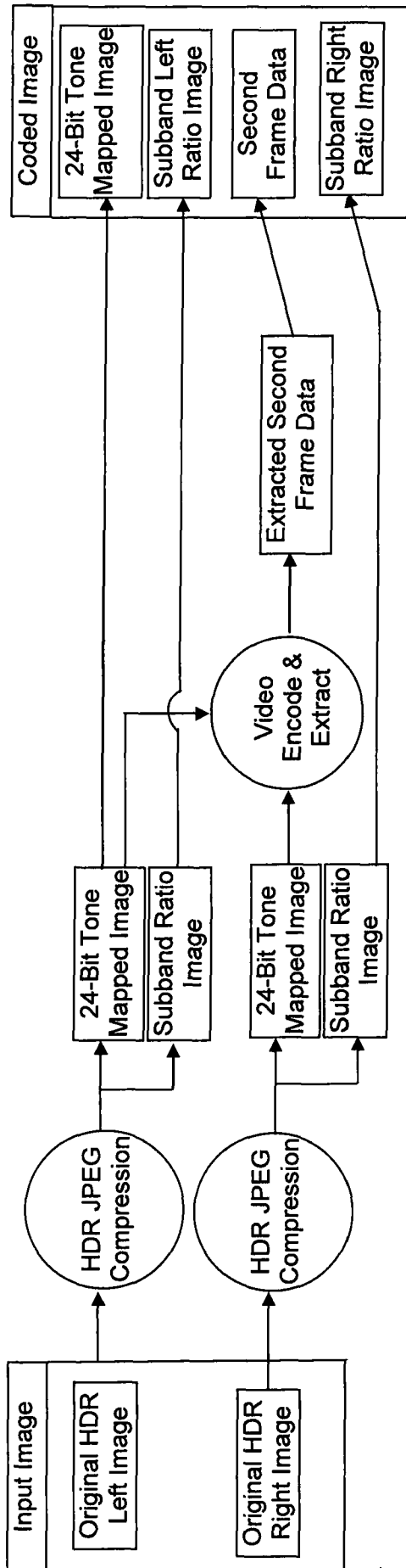


Figure 4: MC Encoding

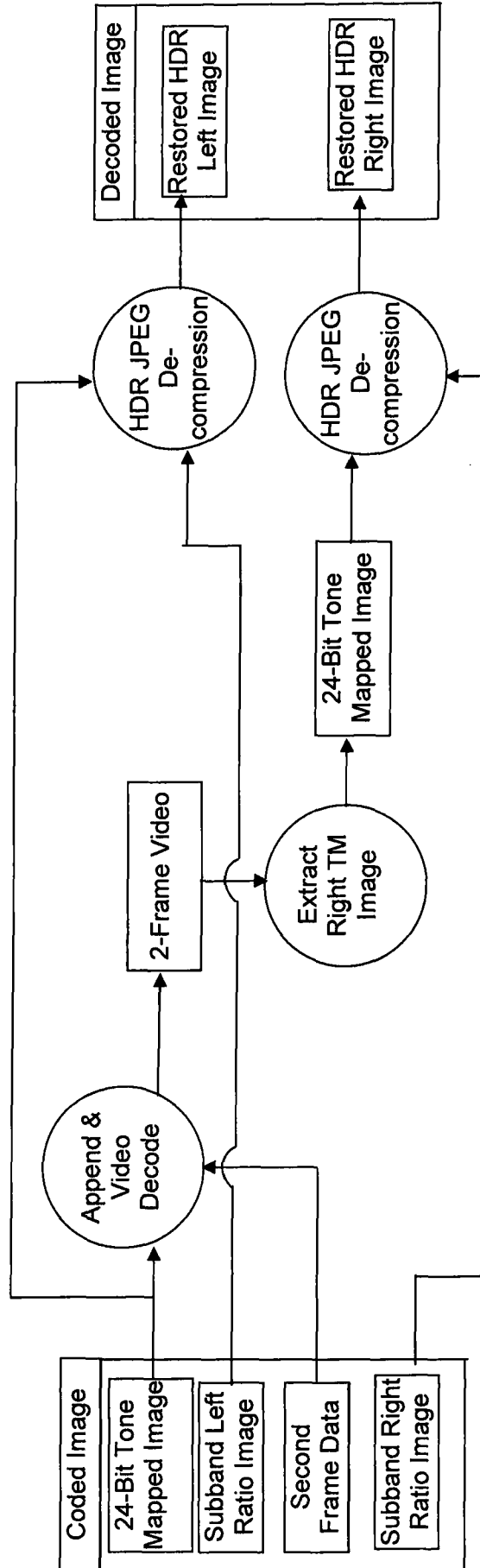


Figure 5: MC Decoding

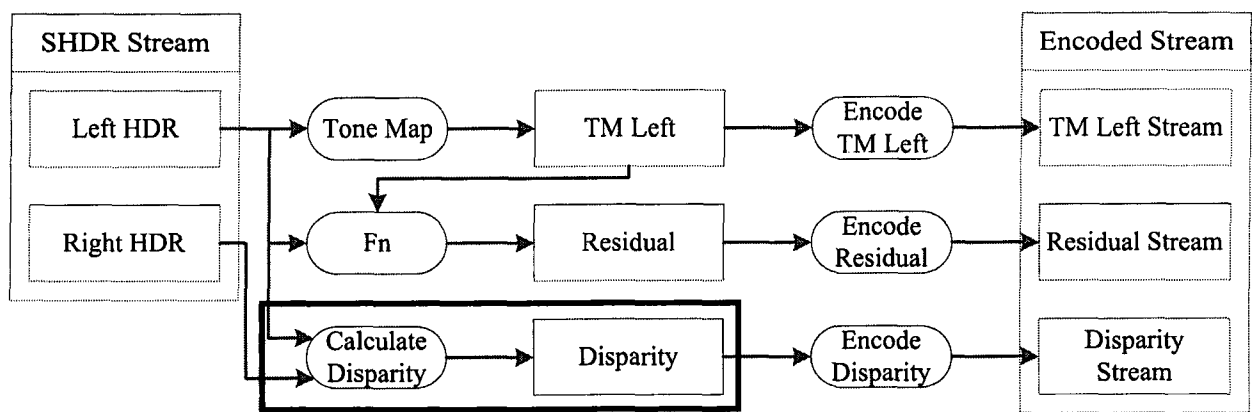


Figure 6

5/9

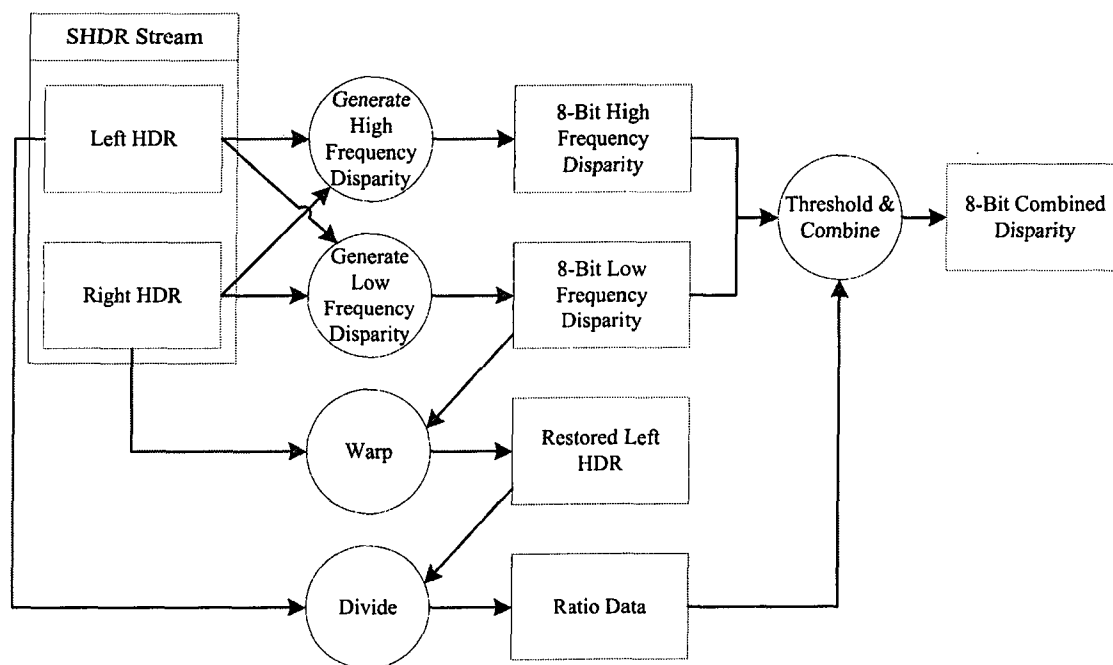


Figure 7

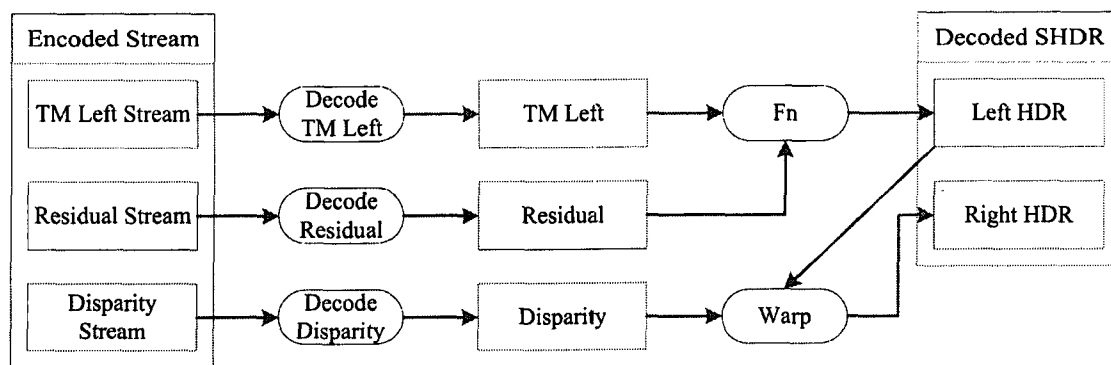


Figure 8

6/9

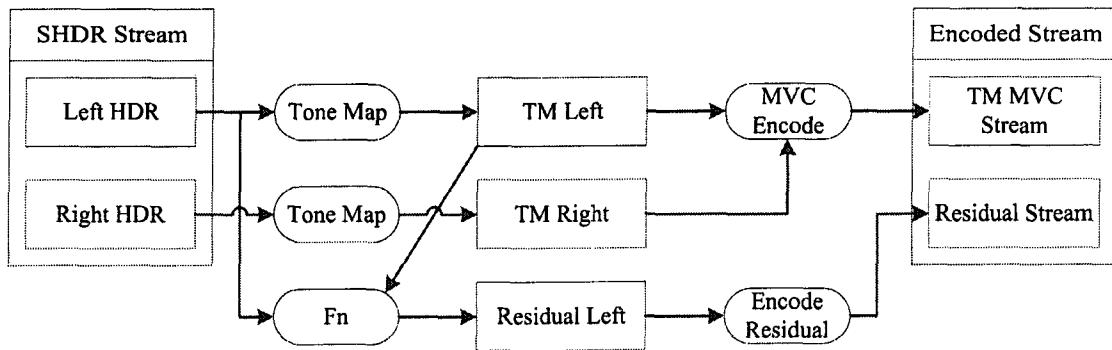


Figure 9

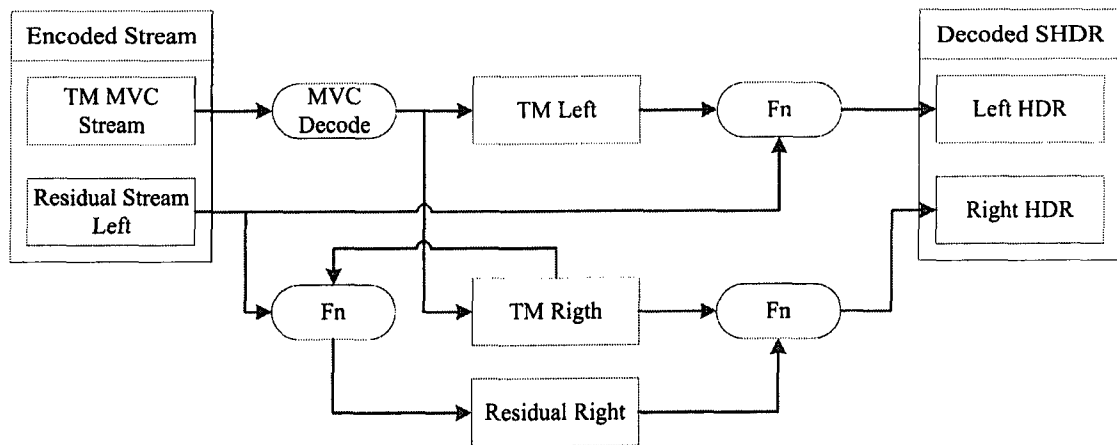


Figure 10

7/9

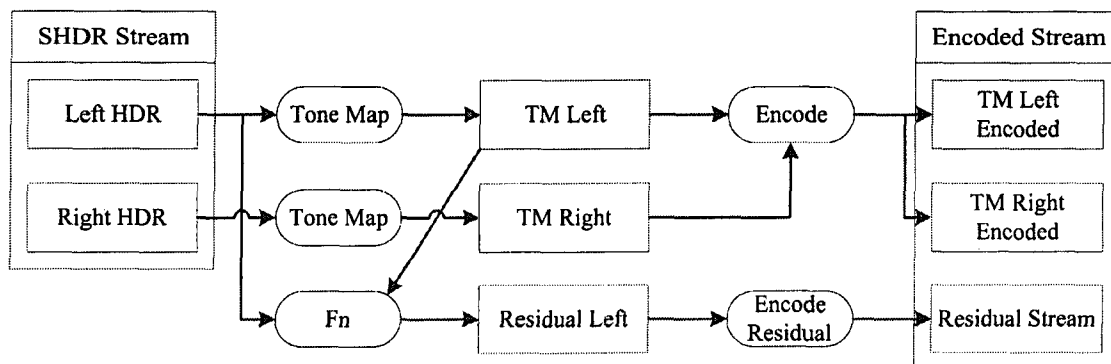


Figure 11

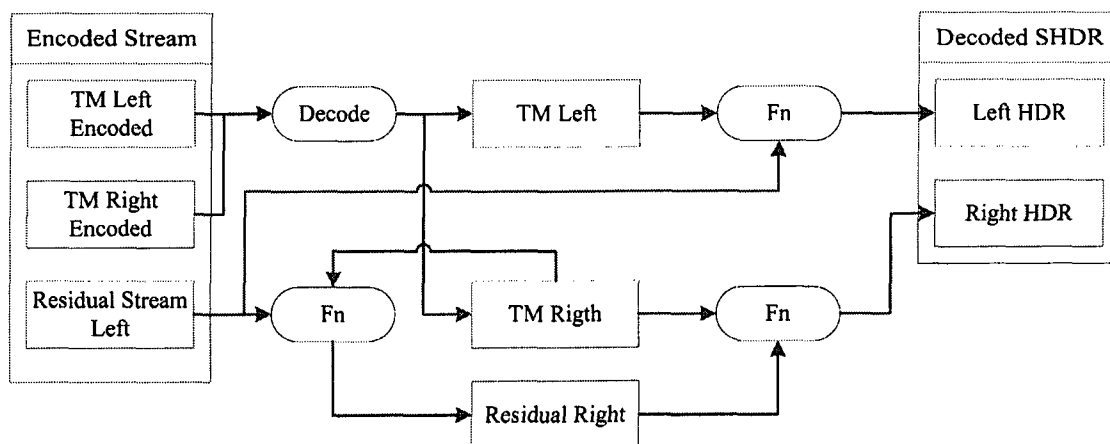


Figure 12

8/9

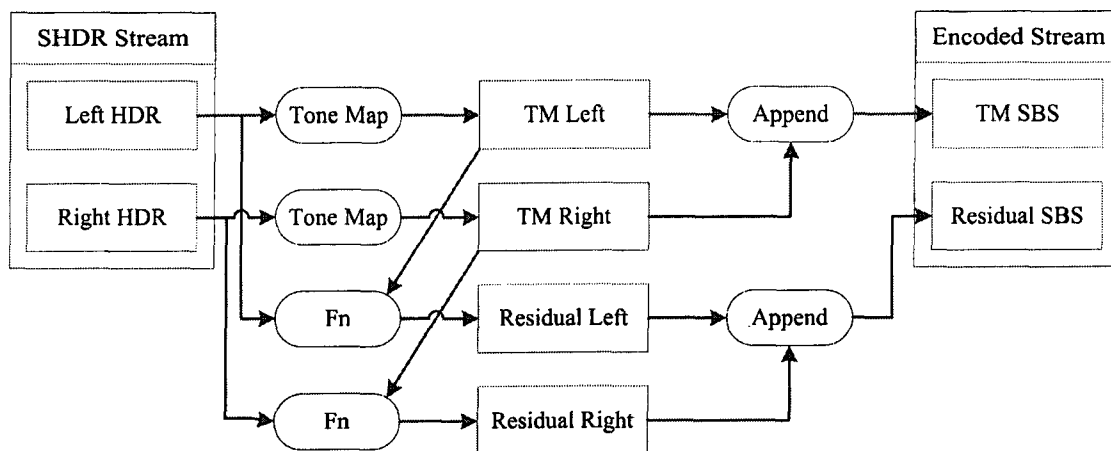


Figure 13

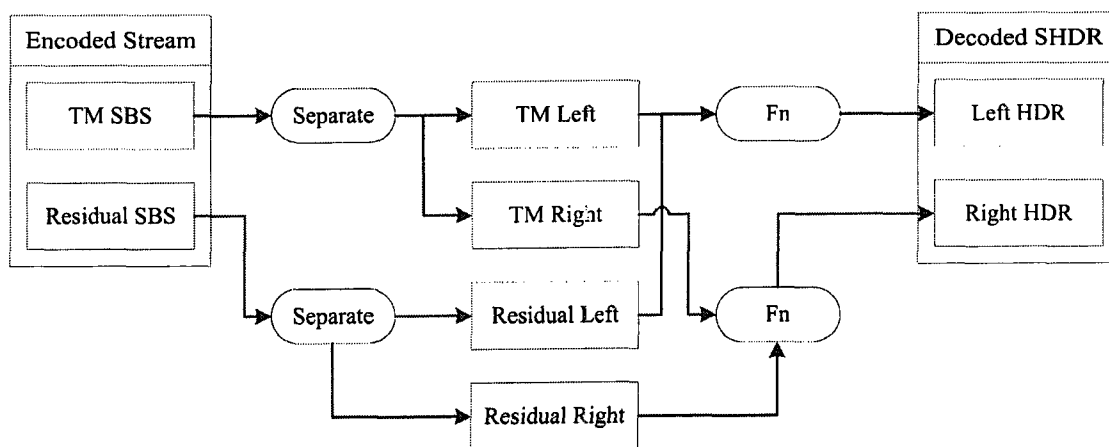


Figure 14

9/9

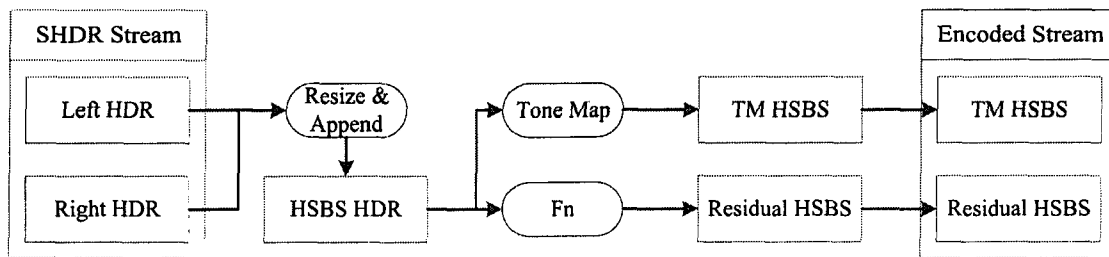


Figure 15

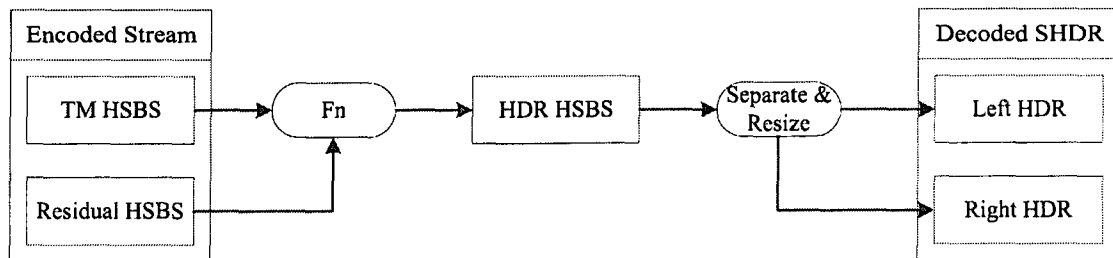


Figure 16

INTERNATIONAL SEARCH REPORT

International application No
PCT/GB2013/052390

A. CLASSIFICATION OF SUBJECT MATTER

INV. H04N7/26 H04N7/32 H04N7/46
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.
X,P	SELMANOVIC E ET AL: "Backwards Compatible JPEG Stereoscopic High Dynamic Range Imaging", PREPRINTS OF CONFERENCE PROCEEDINGS. EG UK THEORY AND PRACTICE OF COMPUTER GRAPHICS, RUTHERFORD APPLETON LABORATORY, DIDCOT, UK, SEPTEMBER 13-14, 2012,, 13 September 2012 (2012-09-13), pages 1-8, XP008165726, the whole document	1-31
X	----- WO 2011/037933 A1 (PANASONIC CORP [JP]; KASHIWAGI YOSHIICHIRO [JP]; CHEN TAO [US]; NISHI) 31 March 2011 (2011-03-31)	1-19, 25-28
Y	the whole document ----- -/-	29,30



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

14 November 2013

Date of mailing of the international search report

21/11/2013

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

Santos Luque, Rocio

INTERNATIONAL SEARCH REPORT

International application No

PCT/GB2013/052390

C(Continuation). DOCUMENTS CONSIDERED TO BE RELEVANT		
Category*	Citation of document, with indication, where appropriate, of the relevant passages	Relevant to claim No.
X	WO 2012/004741 A1 (KONINKL PHILIPS ELECTRONICS NV [NL]; BRULS WILHELMUS HENDRIKUS ALFONSU) 12 January 2012 (2012-01-12) the whole document	1-8, 20-25, 27,28,31
Y	----- ANTHONY VETRO ET AL: "Overview of the Stereo and Multiview Video Coding Extensions of the H.264/MPEG-4 AVC Standard", PROCEEDINGS OF THE IEEE, IEEE. NEW YORK, US, vol. 99, no. 4, 1 April 2011 (2011-04-01), pages 626-642, XP011363626, ISSN: 0018-9219, DOI: 10.1109/JPROC.2010.2098830 page 626 - page 628 page 637 - page 638	29,30
A	----- EP 1 827 024 A1 (SHARP KK [JP]) 29 August 2007 (2007-08-29) abstract figures 1-9	1-31
A	----- WARD GREG ET AL: "JPEG-HDR: A BACKWARDS-COMPATIBLE, HIGH DYNAMIC RANGE EXTENSION TO JPEG", PROCEEDINGS OF THE ANNUAL ACM SYMPOSIUM ON THE THEORY OF COMPUTING, XX, XX, 1 January 2005 (2005-01-01), pages 283-290, XP008079630, the whole document	1-31
A	----- WO 97/43863 A1 (DEUTSCHE TELEKOM AG [DE]; BREIDE STEPHAN [DE]; DAHMEN KLAUS NORBERT [D]) 20 November 1997 (1997-11-20) the whole document -----	1-31

INTERNATIONAL SEARCH REPORT

Information on patent family members

International application No

PCT/GB2013/052390

Patent document cited in search report	Publication date	Patent family member(s)	Publication date
WO 2011037933 A1	31-03-2011	CN 102224731 A EP 2481214 A1 JP 2013505647 A KR 20120089561 A RU 2011120192 A US 2011222605 A1 WO 2011037933 A1	19-10-2011 01-08-2012 14-02-2013 13-08-2012 27-10-2013 15-09-2011 31-03-2011
WO 2012004741 A1	12-01-2012	AU 2011275436 A1 CA 2804345 A1 CN 102959957 A CN 102986214 A EP 2591602 A1 EP 2591606 A1 TW 201215160 A US 2013107956 A1 US 2013108183 A1 WO 2012004709 A1 WO 2012004741 A1	28-02-2013 12-01-2012 06-03-2013 20-03-2013 15-05-2013 15-05-2013 01-04-2012 02-05-2013 02-05-2013 12-01-2012 12-01-2012
EP 1827024 A1	29-08-2007	EP 1827024 A1 JP 4607136 B2 JP 2007243942 A JP 2010213360 A US 2007201560 A1	29-08-2007 05-01-2011 20-09-2007 24-09-2010 30-08-2007
WO 9743863 A1	20-11-1997	DE 19619598 A1 WO 9743863 A1	20-11-1997 20-11-1997