

19



OFICINA ESPAÑOLA DE
PATENTES Y MARCAS

ESPAÑA



11 Número de publicación: **2 744 795**

51 Int. Cl.:

H04N 19/30 (2014.01)

G06T 5/00 (2006.01)

H04N 19/46 (2014.01)

H04N 19/98 (2014.01)

12

TRADUCCIÓN DE PATENTE EUROPEA

T3

96 Fecha de presentación y número de la solicitud europea: **19.03.2015 E 17204594 (0)**

97 Fecha y número de publicación de la concesión europea: **19.06.2019 EP 3324629**

54 Título: **Métodos y aparatos para codificar unas imágenes HDR, y métodos y aparatos para usar tales imágenes codificadas**

30 Prioridad:

28.05.2014 EP 14170157

45 Fecha de publicación y mención en BOPI de la traducción de la patente:

26.02.2020

73 Titular/es:

KONINKLIJKE PHILIPS N.V. (100.0%)

**High Tech Campus 52
5656 AG Eindhoven, NL**

72 Inventor/es:

**VAN DER VLEUTEN, RENATUS JOSEPHUS y
MERTENS, MARK JOZEF WILLEM**

74 Agente/Representante:

ISERN JARA, Jorge

ES 2 744 795 T3

Aviso: En el plazo de nueve meses a contar desde la fecha de publicación en el Boletín Europeo de Patentes, de la mención de concesión de la patente europea, cualquier persona podrá oponerse ante la Oficina Europea de Patentes a la patente concedida. La oposición deberá formularse por escrito y estar motivada; sólo se considerará como formulada una vez que se haya realizado el pago de la tasa de oposición (art. 99.1 del Convenio sobre Concesión de Patentes Europeas).

DESCRIPCIÓN

Métodos y aparatos para codificar unas imágenes HDR, y métodos y aparatos para usar tales imágenes codificadas

5 Campo de la invención

La invención se refiere a la codificación de una (es decir, una imagen fija) pero preferiblemente más (es decir, vídeo) imágenes de alto rango dinámico, y los sistemas y métodos técnicos correspondientes para transmitir la información de imagen codificada necesaria a un lado receptor, y decodificadores para decodificar imágenes codificadas y, en
10 última instancia, ponerlas a disposición para su visualización.

Una imagen HDR es una imagen que codifica las texturas de una escena HDR (que normalmente puede contener regiones muy brillantes y oscuras), de tal manera, con información suficiente para la codificación de alta calidad de las texturas de color de los diversos objetos capturados en la escena, que una representación visual de buena calidad de
15 la escena HDR se puede hacer en una pantalla HDR con alto brillo máximo como por ejemplo 5000 nit. En nuestro marco de codificación HDR desarrollado durante los años anteriores, también queremos codificar al mismo tiempo varias subvistas de rango dinámico abarcadas en esa escena, que pueden servir como buenas imágenes de conducción para diversas pantallas de rango dinámico reducido con éxito, al menos una imagen LDR parece adecuada para conducir, por ejemplo, una pantalla de brillo máximo de 100 nit.

Además, una codificación HDR está preferiblemente diseñada técnicamente de modo que no solo coincida relativamente bien con las tecnologías de codificación de vídeo existentes, sino que incluso pueda encajar en los marcos de codificación de imagen o vídeo actuales de la tecnología existente, como por ejemplo almacenamiento en
20 disco blu-ray (y sus limitaciones como la cantidad de memoria de almacenamiento, pero también sustancialmente todos los demás aspectos estandarizados), o conexiones de cable HDMI, u otros sistemas de transmisión o almacenamiento de imágenes, etc. En particular, los sistemas de codificación HDR que usan dos imágenes codificadas por instante de tiempo (por ejemplo, una imagen LDR y una imagen que tiene intensidades de brillo locales para aumentar cada píxel a un píxel de imagen HDR) son vistos por los fabricantes de aparatos como innecesariamente
25 complejos.

La codificación de vídeo HDR (o incluso de imágenes fijas) se ha investigado recientemente de manera tentativa y ha sido una tarea desalentadora hasta ahora, y la creencia típica es que uno necesita ir hacia significativamente más bits, para codificar los brillos por encima del rango LDR de objetos de escena (por ejemplo codificaciones que codifican
30 luminancias de escena directamente), o uno necesita un enfoque de dos capas, en el que por ejemplo Además de una imagen de reflectancia del objeto, hay una imagen de aumento de iluminación o estrategias de descomposición similares. Enfoques alternativos, como por ejemplo el sistema de codificación HDR de Dolby, por lo general, siempre utiliza una tecnología de codificación HDR de dos capas de este tipo (consulte la Patente de Estados Unidos No. 8248486B1). WO2010/105036 es otro ejemplo de dicho sistema de codificación instantánea de 2 imágenes por tiempo de Dolby. Se puede usar un mapeo de tono fijo simple TM (por ejemplo, emular el comportamiento de la fotografía
35 clásica de celuloide analógico) para mapear una imagen LDR a una clasificación HDR correspondiente, o viceversa. Esta puede ser una función de mapeo no muy correcta (porque el clasificador puede haber utilizado decisiones de optimización complicadas, por ejemplo, su clasificación LDR), por lo que puede haber una diferencia considerable entre la predicción de la imagen original de HDR y la propia imagen HDR original, pero eso no es problema porque el sistema puede enviar las diferencias como una segunda imagen de corrección (flujo de bits residual 742).
40

Otro ejemplo de dicho sistema de codificación de dos imágenes es el sistema de codificación HDR Technicolor (documento JCTVC-P0159r1, de la 16ª reunión del Equipo Cooperativo Conjunto sobre codificación de vídeo de ITU-T SG 16 WP3, San José de 9-17 de enero de 2014) que codifica como primera imagen una señal de baja resolución que es la iluminación local, ya que normalmente se generará en una luz de fondo LED modulada en 2D, y la segunda
45 imagen es una imagen de textura de modulaciones de rango dinámico típicamente más bajas en esa señal de baja frecuencia, que normalmente se generará en la pantalla mediante señales de activación de la válvula LCD adecuadas.

Philips ha inventado recientemente un enfoque de imagen única mucho más simple (aún en gran parte inédito, pero algunos aspectos se pueden encontrar en WO2012/147022 y WO2012/153224), que es una nueva dirección en la
50 codificación de imágenes y vídeos, y no solo a priori no fácil de imaginar, pero también cuando realmente lo hace, lo que lleva a muchos problemas técnicos nuevos que deben resolverse, lo que sin embargo funciona muy bien en la práctica.

Con "alto rango dinámico" (HDR) queremos decir que las imágenes capturadas desde el lado de captura tienen una alta relación de contraste de luminancia en comparación con la codificación LDR heredada (es decir, pueden
55 manejarse relaciones de contraste de luminancia de objeto de 10.000:1 o más por la codificación, y todos los componentes de la cadena de manejo de imágenes hasta el representado; y las luminancias de los objetos capturados pueden estar por encima de 1000 nit, o más específicamente, por lo general pueden reproducirse representarse por encima de 1000 nit para, dado el entorno de reproducción, generar un aspecto deseado de una lámpara encendida o un exterior soleado). Y típicamente la representación de tales imágenes es HDR (es decir, las imágenes deben ser
60 adecuadas ya que contienen información suficiente para una representación HDR de alta calidad, y preferiblemente

de una manera técnicamente fácil de usar), lo que significa que las imágenes se representan o pretenden representarse en pantallas con un brillo máximo de al menos 1000 nit (lo que no implica que no puedan y no necesiten estar alternativamente en pantallas LDR de, por ejemplo, un brillo máximo de 100 nits, generalmente después de un mapeo de color adecuado).

5 Antecedentes de la invención

Recientemente se han propuesto varias tecnologías de codificación HDR, como por ejemplo el método de doble capa de Dolby (WO2005/1040035). Sin embargo, la industria todavía está buscando una tecnología pragmática de
 10 codificación de video (/imagen) HDR con ajustes (un equilibrio de) todos los requisitos, como los factores muy importantes como la cantidad de datos, pero también la complejidad computacional (precio de los circuitos integrados), facilidad de presentación, versatilidad para que los artistas creen lo que quieran, etc. En particular, un enfoque de doble capa se considera innecesariamente complejo. Idealmente, a uno le gustaría poder diseñar una tecnología de codificación que se ajuste a la codificación heredada, como por ejemplo Codificación MPEG HEVC basada en DCT.
 15 Un problema es que esto es algo contraintuitivo: ¿cómo se puede codificar una imagen HDR, que por definición debería ser algo diferente de una imagen LDR, que generalmente tiene una mayor cantidad de rangos interesantes de brillo/luminancia, en una tecnología optimizada para contener imágenes LDR? Estos sistemas de codificación/manejo de imágenes LDR heredados fueron diseñados y optimizados para trabajar con escenarios típicos de imágenes LDR, que normalmente están bien iluminados, por ejemplo, una relación de iluminación en estudio 4:1 (o, por ejemplo, 10:1), que proporciona a la mayoría de los objetos (que pueden variar en reflectancia entre, por ejemplo, 85% para blanco y 5% para negro) en la vista una relación de contraste total de aproximadamente 68:1 (resp. 170:1). Si se observa la representación relativa (es decir, el mapeo del blanco de la imagen al blanco de la pantalla disponible) de las luminancias de los objetos a partir de un blanco pico, un monitor LCD temprano típico sin atenuación local habría tenido algo como 100 nit de blanco y 1 nit de negro coincide con la relación de contraste de la imagen, y normalmente se pensaba que, en promedio, los sistemas CRT que podrían haber sido observados también durante el día tendrían algo así como una capacidad de 40:1. Tener una función gamma estándar de asignación de código de luminancia heredada de 2.2 en estos sistemas parecía satisfactoria para la mayoría de los escenarios de contraste de escena aún mayor. Aunque algunos en aquellos días considerados como errores aceptables se cometieron, tales errores de representación de regiones de escena de alta luminosidad mal codificadas (por ejemplo, recorte duro de exteriores brillantes detrás de una ventana) también fueron aceptables porque las pantallas LDR no podían hacer que esas luminancias de objetos fuesen físicamente precisas de todos modos.

Sin embargo, hay escenarios para los que ahora existe el deseo de mejorar la representación, como por ejemplo una escena interior en la que uno puede ver simultáneamente el exterior soleado, en cuyo caso puede haber una relación de iluminación de 100:1 o incluso más. Con el representado relativo lineal (es decir, enfocándose en las regiones codificadas más brillantes en primer lugar, o equivalentemente el gris medio de las regiones de escenas más brillantes, y mapeando la imagen en blanco para mostrar el blanco), ¡el blanco en el interior se mapearía en negro sicovisual para el observador! Entonces, en LDR, esas regiones soleadas generalmente aparecerán como recortadas (suaves) (generalmente ya en la imagen codificada que tiene códigos difíciles de discriminar alrededor del máximo 255 para esos píxeles). Sin embargo, en una pantalla HDR nos gustaría mostrarlos brillantes y coloridos. Eso daría una representación mucho más natural y espectacular de tales escenas (como si realmente estuviera de vacaciones en Italia), pero incluso las escenas en las que el contenido de mayor brillo solo se compone de algunos reflejos especulares ya muestran una mejora importante en la calidad visual. Si aún no, los artefactos como los errores de recorte o cuantificación parecen molestos, por ejemplo, una pantalla de 5000 o 10000 nit, al menos queremos poder manejar tales pantallas HDR con el tipo correcto de imagen, para que las imágenes representadas sean tan hermosas como lo permita la pantalla HDR.

Sin embargo, el pensamiento clásico era que para codificar rangos adicionales de sobrebrillo, uno necesitaría tener (muchos) más bits, que son los bits más altos que codifican las luminancias del objeto por encima de un rango LDR. Eso podría suceder ya sea mediante la codificación nativa en palabras de código más grandes (como OpenEXR con 16 bits de los cuales un bit de signo, exponente de 5 bits y mantisa de 10 bits, o la codificación LogLuv de Ward, que matemáticamente trata de capturar rigurosamente todo el mundo de posibles luminancias de objetos con alta precisión), o mediante el uso de una primera capa con códigos de rango LDR estándar (por ejemplo, una aproximación JPEG clásica de la imagen HDR), y una segunda capa para mejorar tales luminancias de píxeles a un brillo más alto (por ejemplo, una imagen de refuerzo para aumentar cada píxel si es necesario para una luminancia más alta, es decir, una multiplicación de dos de esas imágenes de 8 bits es equivalente a un único código lineal de 16 bits).

Un problema práctico importante que debe resolverse al diseñar una tecnología de codificación HDR práctica, además del hecho de que, por supuesto, debe ser capaz de manejar una gran variedad de imágenes HDR diferentes, es que los fabricantes de hardware desean cantidades menores de bits por palabra de código (canal, es decir, la luma y dos canales cromáticos), sin embargo, y aunque nuestra tecnología propuesta a continuación también puede funcionar con palabras de bits más grandes, venimos con una solución que funciona bien bajo una limitación de 10 bits para al menos un canal de luminancia (o más precisamente una luma) (tenga en cuenta que aunque dilucidamos las realizaciones con un canal de luminancias, nuestros conceptos pueden mutatis mutandis ser incorporados como trabajo en representaciones de color RGB (lineales o no lineales, etc.). Además, desarrollamos un marco que puede hacer con una doble filosofía tanto la codificación de píxeles de color (del aspecto HDR a través de una imagen LDR)

como la conversión del aspecto del color para varios escenarios de representación (es decir, los aspectos óptimos necesarios para representar una escena en varias pantallas con diferente brillo máximo, por ejemplo, PB = 800 nit) de manera funcional, lo que significa que solo las funciones deben cocodificarse al codificar el aspecto de al menos una clasificación adicional, y específicamente un aspecto HDR además de un aspecto LDR, en lugar de para cada foto al menos una segunda foto.

Actualmente tenemos dos categorías de sistemas de codificación HDR, ya que al mercado le gustaría tal versatilidad en un sistema de codificación, dados los diferentes jugadores y sus diferentes necesidades. En el sistema de modo i (o codificado parecido a HDR como una única imagen de definición, por ejemplo, en un disco BD, o un flujo de imágenes AVC o HEVC a través de una conexión de red) utilizamos una imagen parecida a HDR como la única imagen de píxel, que se utiliza para codificar las texturas y formas de color de los objetos (véase en el documento WO2015007505 del solicitante cómo se puede enviar una imagen HDR única a un receptor para definir los colores de píxeles de al menos el aspecto HDR, y cómo con las funciones de reclasificación apropiadas el receptor puede calcular procesando los colores en esa imagen otras imágenes de aspecto). Con esto queremos decir que tomamos la imagen de clasificación maestra HDR original, es decir, una imagen con un color óptimo para que se vea mejor en una pantalla HDR de referencia, como por ejemplo normalmente una pantalla de brillo máximo de 5000 nit, y solo transforma esto mínimamente: básicamente solo aplica una función de asignación de código o una función de transferencia optoelectrónica OETF (tenga en cuenta que aunque este OETF define cómo las luminancias de escenas capturadas, por ejemplo, por una cámara, se transfieren a códigos luma, los ingenieros de televisión en lugar de gustarles especificar la función inversa que es la función de transferencia electroóptica EOTF para pasar de los códigos de luma a las luminancias representadas en la pantalla de referencia) mediante el uso de OETF asigna de manera óptima 10 bits de códigos, disponibles, para el canal luma Y' sobre todos los valores de brillo que uno necesita poder hacer en una pantalla de referencia [0-5000] nit. Luego se pueden realizar otras graduaciones deseadas para pantallas de diferente brillo máximo transformando esta imagen de aspecto HDR. En nuestro marco, permitimos que este aspecto de la pantalla se pueda ajustar haciendo típicamente una segunda clasificación que se encuentra en el otro extremo del rango de posibles pantallas para servir, es decir, un aspecto que sea óptimo o razonable de acuerdo con el creador de contenido/clasificador de color para una pantalla de brillo máximo de 100 nit (que suele ser la pantalla de referencia para la categoría de pantallas LDR). Tenga en cuenta que esta es una cocodificación de un aspecto adicional en lugar de un simple paso de recoloración del lado de la creación. Esta transformación de color requerida se determina aplicando funciones de mapeo como funciones gamma que realizan un reajuste de brillo global (por ejemplo, iluminar los colores más oscuros en la imagen), curvas arbitrarias en forma de S o inversas en forma de S para ajustar el contraste local, funciones de procesamiento de saturación de color para ajustar por ejemplo la saturación al brillo correspondiente de algunos objetos o regiones en la imagen, etc. Podemos codificar libremente esas funciones (cuantas funciones necesitemos siempre que pertenezcan a un conjunto limitado de funciones básicas que el receptor puede entender de manera estandarizada) como metadatos asociados con la imagen pixelada de aspecto HDR, en cuyo caso DEFINE paramétricamente la segunda imagen de clasificación de aspecto LDR a partir de la imagen de aspecto HDR (es decir, ya no necesitamos codificar esa imagen de aspecto LDR como una imagen de píxeles). Observe cuidadosamente la diferencia con los sistemas de codificación de dos capas: en nuestro sistema, todas las funciones de transformación de color están codificadas en el segundo aspecto para poder volver a clasificar el segundo aspecto en el receptor, por lo que en lugar de las funciones aproximadas de tecnologías de 2 imágenes, nuestras funciones contienen el conocimiento inteligente completo de cómo se deben comportar las iluminaciones de los diversos objetos en diversas representaciones de acuerdo con el creador de contenido. Dado este conocimiento de cómo los artistas creadores quieren que el aspecto se transforme de la primera búsqueda de pantallas con un primer nivel de capacidades de reproducción del color a una segunda búsqueda de pantallas con un segundo nivel de capacidades de reproducción del color (en particular el brillo máximo de la pantalla), una pantalla con capacidades intermedias (por ejemplo, brillo máximo de 1200 nit) puede obtener automáticamente una imagen de conducción más óptima para su situación de representación utilizando el conocimiento en las dos graduaciones e interpolando (por ejemplo, la pantalla puede hacer una mezcla asimétrica de las dos imágenes pixeladas del aspecto HDR y la imagen de aspecto LDR derivada de la imagen de aspecto HDR y las transformaciones funcionales, en las que los porcentajes de mezcla multiplicativos se determinan por qué tan cerca de la pantalla HDR o LDR la pantalla real está en una escala no lineal sicovisual), que será mejor que manejar la pantalla con la imagen original HDR o la imagen LDR.

Esta es una definición potente pero simple de no solo un aspecto de imagen (HDR) en una escena (por ejemplo, una representación de 5000 nits), sino un marco completo para obtener representaciones razonables de la escena para varias pantallas posibles en el campo como en el hogar del consumidor (e incluso una posible adaptación al entorno de visualización, por ejemplo, aplicando un modelado postgamma la sensibilidad de contraste cambiada de la visión humana bajo diversas iluminaciones envolventes). Es principalmente útil, por ejemplo, para aplicaciones/escenarios en los que un creador ha hecho una buena versión HDR de su contenido, y quiere tener primero este aspecto HDR en la codificación real enviada a los receptores (por ejemplo, en un disco HDR BD, o al ordenar una película HDR en línea a través del Internet, o una transmisión de televisión HDR, etc.). No es necesario que un cliente que compre esta versión de contenido tenga una pantalla HDR, ya que puede comprarla para más adelante cuando tenga una pantalla HDR y ahora puede usar la conversión HDR-2-LDR, pero sería la opción preferida cuando el cliente quiere contenido para su pantalla HDR.

Mientras que la forma de aspecto HDR anterior de codificar escenas HDR (como se explicó el modo i es que al menos las imágenes de aspecto HDR codificadas como una imagen de píxeles, pero de hecho también se codifican más

miradas en esa misma escena pero luego paramétricamente con funciones de transformación de color, como, por ejemplo, una realización de recorte, en la que el aspecto LDR aísla un subrango de la imagen HDR y recorta el resto) ya plantea desafíos técnicos significativos para llegar a un nuevo sistema técnico pragmático para la imagen futura, pero principalmente también la codificación de video (teniendo en cuenta tales factores como la simplicidad del diseño de IC para los fabricantes de hardware, pero que permite a los creadores de contenido crear cualquier contenido HDR hermoso como películas de ciencia ficción, programas de televisión espectaculares o documentales sobre la naturaleza, etc. que quieran hacer, con muchos efectos creativos de HDR como lámparas que parecen realmente iluminadas), el mercado deseaba otra capa de complejidad, que enseñaremos en esta descripción de la patente.

Es decir, para algunas aplicaciones (que llamaremos modo-ii), es posible que desee tener una imagen de aspecto LDR como la única imagen pixelada que codifica los objetos de la escena, que está, por ejemplo, escrita como única imagen en un disco blu-ray. Aunque el creador de contenido también se preocupa mucho por la calidad del aspecto HDR, se centra mucho en que el aspecto LDR sea similar al de las tecnologías heredadas. Por lo general, habrá parámetros de función cocodificados en metadatos asociables para derivar una imagen de aspecto HDR actualizando la imagen de aspecto LDR que se comunicó en la señal de imagen S_{im} . Pueden haber diversas razones para elegir esta variante de modo ii (o aspecto LDR), que pueden, por ejemplo, ser para sistemas heredados que no pueden realizar ningún procesamiento (por ejemplo, si uno prefiere codificar la única imagen en una realización particular que codifica los colores como colores $Y'uv$ en lugar de una codificación $YCrCb$, uno podría codificar esto en un marco HEVC heredado pretendiendo que la imagen $Y'uv$ es una imagen $YCrCb$ extrañamente coloreada y utilizando más esquemas de codificación heredados basados en DCT, como estandarizados en uno de los miembros de la familia de códecs MPEG), pero también para aplicaciones que necesitan un aspecto LDR (por ejemplo, ver un película en una pantalla portátil de bajo brillo) y es posible que no desee realizar demasiado procesamiento. O tal vez el creador no quiere invertir demasiado tiempo en crear un aspecto HDR perfecto (pero, por ejemplo, solo lo hace rápidamente al realizar un ajuste fino menor de una autoconversión LDR-2-HDR que, por ejemplo, aísla regiones brillantes y las aumenta de forma no lineal, por ejemplo, para una vieja remasterización de películas de Laurel y Hardy), y considera que su aspecto LDR es la clasificación maestra más importante de los aspectos LDR y HDR, que deben codificarse directamente sin necesidad de ninguna transformación de color, con posibles errores de color., por ejemplo, una emisora de televisión puede elegir esta opción, especialmente para transmisiones de la vida real (por ejemplo, las noticias no necesitan estar en el HDR más espectacular).

Sin embargo, esta codificación de aspecto LDR (modo ii) tiene una complejidad adicional debido a la naturaleza matemática del problema y la codificación de las matemáticas, por un lado, frente a los deseos de clasificación artística liberal por el otro, lo que hace que sea una tarea desalentadora elaborarlo con un buen marco técnico. Para ser más precisos, por un lado, necesitamos funciones que primero descendan de una imagen HDR maestra deseada, y en el receptor con estas funciones recibidas (o las funciones inversas de la declasificación en realidad) el receptor puede actualizar al menos a una aproximación cercana de la imagen HDR original nuevamente, es decir, en los datos de parámetros de la función de metadatos habrá parámetros para funciones (derivadas por el codificador de las funciones que el clasificador utilizó en la declasificación del HDR maestro) que pueden mapear la única imagen LDR a una predicción HDR suficientemente cercana Rec_HDR . Pero, por otro lado, la imagen LDR debería, cuando se representa directamente en una pantalla + -100 nit, es decir, sin una mayor transformación de color, verse lo suficientemente bien de acuerdo con el clasificador de color también. Por lo tanto, habrá un equilibrio entre la selección de las funciones, y cómo influirán en el aspecto LDR y Rec_HDR , y eso también teniendo en cuenta otros problemas, como los fabricantes de IC o aparatos que desearían ver un conjunto limitado de funciones estándar que son útiles para volver a clasificar los aspectos, y a los creadores de contenido les gustan esas funciones para especificar rápidamente cualquier aspecto que deseen, ya que el tiempo de clasificación es costoso y el momento de los lanzamientos de películas puede ser crítico. En la siguiente descripción describiremos un sistema práctico para manejar esta variante del modo ii de la codificación de escenas HDR.

Los siguientes documentos están relacionados tangencialmente con nuestra presente invención, pero no obstante son interesantes de mencionar y distinguir.

Durante la reunión de MPEG2014/M34274 Sapporo, R. van de Vleuten et al. de Philips publicada (julio de 2014, después de la fecha de prioridad de esta solicitud) "Proposed electro-optical transfer function (EOTF) for HDR video delivery". Este documento contiene el EOTF que también se usa en nuestras realizaciones para hacer un mapeo previo al rango dinámico de LDR. Sin embargo, este documento solo enseña que este EOTF tiene buenas propiedades de correlación con el aspecto de ligereza humana y, por lo tanto, podría ser útil como una función de asignación de código luma para la codificación de imágenes HDR. Ergo, claramente, lo que a lo sumo se puede derivar de este documento, es una enseñanza o inspiración de que, si uno desea codificar una única imagen o video HDR maestro, es decir, las imágenes con luma de píxeles codificadas para, por ejemplo, una pantalla de 5000 nit, entonces este EOTF es útil. Codificar únicamente un video HDR significa que solo hay imágenes definidas con referencia a, por ejemplo, brillo máximo de 1000 nit o 5000 nit, que deben comunicarse a los receptores, y no hay nada más, en particular, ningún otro aspecto gradual en las mismas imágenes de la escena, como las imágenes de brillo máximo LDR 100 nit, que también deben comunicarse. Esta única codificación de video HDR es mucho más simple técnicamente, y una visión diferente de la codificación HDR que no es lo que enseña la presente invención. Estas imágenes solo HDR no son adecuadas para la visualización directa en pantallas LDR heredadas, que generalmente tienen un brillo máximo de alrededor de 100 nits, como por ejemplo las regiones oscuras se verán demasiado oscuras

en una pantalla que es 50 veces más oscura que la pantalla de 5000 nit para la que se crearon las imágenes HDR. Como se mencionó claramente anteriormente, la presente solicitud proporciona tecnologías para los escenarios en cuyo caso uno quiere transmitir información HDR de una escena HDR original a un receptor, pero en realidad desea comunicar una versión LDR de dicha imagen (de una manera que sin embargo permite reconstruir la imagen HDR en el receptor), lo que da un aspecto apropiado cuando se procesa directamente en una pantalla heredada, es decir, con las regiones oscuras que no son demasiado oscuras. Es un requisito técnico muy diferente de la codificación de imágenes HDR solamente, y, por lo tanto, esta solicitud enseña factores adicionales bien contemplados.

El documento m34165 de la misma reunión de Sapporo de Tourapis et al. "Report on the XYZ/HDR exploratory experiment 1: EOTFs for XYZ HDR delivery", de manera similar es solo un estudio sobre cómo se desempeñan varias funciones EOTF cuando solo se codifican imágenes HDR con ellas (porque usar el EOTF incorrecto puede conducir a bandas en el lado receptor), es decir, sin ninguna consideración, o incluso mencionar la necesidad de una imagen LDR, y mucho menos la calidad visual o artística de esa imagen LDR.

El documento WO2013/046095 enseña los principios técnicos generales necesarios para poder hacer cualquier decodificación HDR, en particular nuevamente la decodificación de datos de escena HDR que se codificó como imágenes HDR únicas, pero que no necesariamente tienen que mostrarse en una pantalla con el mismo brillo máximo de la pantalla, (PB_D) como la luminancia del punto blanco de la imagen HDR recibida (por ejemplo, considere un decodificador que recibe contenido HDR especificado, es decir, con el brillo de píxeles de objeto de imagen con el color correcto que se representará en una pantalla de referencia PB_D de 5000 nit, sin embargo, tiene una pantalla adjunta de digamos 2000 nit PB_D, o 7000 nit, por lo tanto, tiene que hacer una transformación óptima del rango dinámico para obtener las imágenes que se ven correctamente para enviar a su pantalla conectada real). Necesita saber varias cosas, como cuál era la luminancia de punto blanco del código (por ejemplo, R=G=B=1023 significaba un blanco de 1000 nit, o 5000 nit, ya que eso cambiaría seriamente la transformación necesaria a algo representable blanco en la pantalla conectada), y cuál es el EOTF utilizado que define los códigos RGB o luma, es decir, que vincula los códigos luma a las luminancias que se supone que representan, es decir, que deben representarse en la pantalla conectada. También hay algunas enseñanzas generales de que algunos datos de transformación de rango dinámico se pueden comunicar con la(s) imagen(es)/video(s) HDR comunicada única, o tal vez con imagen(es) LDR comunicadas. Esta enseñanza podría, en el mejor de los casos, inspirar a un lector experto para que cocomunique alguna modificación de aspecto deseada con la imagen (por ejemplo, el HDR podría oscurecerse a propósito por alguna razón, y luego iluminarse nuevamente en el lado del receptor al comunicar una transformación de rango dinámico brillante/mapeo de tonos, o tal vez la transformación sucedería en el lado receptor independientemente de lo que hubiera hecho el lado de codificación con la imagen antes de transmitirla; esta función de mapeo de tonos debería contrastarse con el EOTF, que es una asignación técnica pura de códigos luma a luminancias, pero el mapeo de tonos puede hacer algún cambio adicional de los brillos de los píxeles, por ejemplo, técnicamente realizado como un mapeo input_luma-to-output_luma, por ejemplo, por razones artísticas). Entonces, esa solicitud de patente que enseña sobre el marco de manejo de datos centrales para gestionar la comprensión del significado de algún tipo de codificación de imagen HDR (definida por el brillo máximo al que corresponde el máximo de los códigos, por ejemplo, R=G=B=1023 debe ser representado como 5000 nit blanco; y qué EOTF se usó para asignar todos los códigos de luma RGB a real para ser valores RGB y luminancias lineales representadas en pantalla), lo que no se puede derivar de esto es la construcción de un sistema técnico nuevo y diferente específico que comunica los datos de la imagen HDR como imágenes LDR como en la presente solicitud, y mucho menos los componentes específicos necesarios para hacerlo pragmáticamente como se reivindicó.

El Doc. JVT-S087 de la 19.ª Reunión MPEG de Ginebra: Segall et al. El "Tone mapping SEI message" simplemente enseña genéricamente que si alguien desea que se aplique "algún" mapeo de tonos a las luminancias de píxeles, entonces puede comunicar esto en un mensaje SEI de metadatos dedicado. Es decir, el lado receptor podría recibir por ejemplo una codificación de 10 bits de una imagen HDR, y, por ejemplo, usar la imagen SEI para convertirla en una imagen LDR estándar de 8 bits correspondiente a esta, es decir, con brillos de píxeles de objeto más adecuados para representar en una pantalla PB_C de 100 nit (en lugar de la pantalla de, por ejemplo, 1000 nit o 5000 nit para la que se hizo la imagen HDR). Por ejemplo, se puede usar una forma sigmoïdal. Este es nuevamente un ejemplo de comunicar únicamente los datos de color de píxeles de un video HDR de imágenes, por lo que este documento nuevamente no contiene nada suficiente para inspirar incluso hacia algún componente de un sistema de codificación dual compatible con LDR para imágenes HDR como lo hemos hecho nosotros.

El Doc. JCTVC-P0159r1/m32076 de Lasserre et al. "High dynamic range video coding" es un ejemplo de un sistema de 2 capas (es decir, un flujo de imagen básico y un flujo de imágenes de corrección) (ver Figura 1), por lo tanto, muy alejado de nuestra tecnología, y que no contiene nada útil, o incluso es probable que se le consulte para obtener inspiración útil. Los sistemas de dos capas están en su filosofía técnica, por lo tanto, la construcción y los componentes técnicos utilizados son muy diferentes de las codificaciones de una sola imagen, además de las transformaciones de la función de color para derivar otras imágenes de rango dinámico.

El documento EP2406959 o WO2010/105036 también es solo una codificación HDR de dos capas (ver Figura 7 flujo de bits residual 742), muy diferente e irrelevante para obtener enseñanzas suficientemente relacionadas con los conceptos de la presente solicitud. Este proporciona una vista interesante sobre algunos aspectos de la codificación de imágenes HDR como información de fondo.

El doc. M34355 de la reunión de julio de Sapporo (de nuevo después de la fecha de prioridad) de Stessen et al de Philips "Chromaticitybased color signals" enseña que cuando uno quiere codificar imagen(es) (única) HDR, uno puede desear usar otro espacio de color diferente al Espacio de color Y'CbCr de la codificación LDR heredada. Este menciona nuevamente que el eje de "brillo" de este espacio de color podría definirse con el EOTF de Philips, que sin sorpresa también se usa en algunas de nuestras divulgaciones o enseñanzas, como en la presente solicitud de patente. El documento luego enseña que si uno quiere hacer un procesamiento de color en un espacio que se determina por ej. un EOTF muy doblado y dos crominancias, por ejemplo, si un receptor desea hacer su transformación de rango dinámico u otra optimización de color al final, por ejemplo, lo que se haya decidido automáticamente podría ser una transformación razonable para convertir las imágenes HDR recibidas en imágenes LDR, entonces este espacio puede no ser tan útil y podrían producirse errores de color de una magnitud mayor de lo deseable. Por lo tanto, los autores introducen para los dos canales de color las cromaticidades que son independientes del rango dinámico. De esto se podría derivar poco valor para las enseñanzas de la presente solicitud, y mucho menos que esas enseñanzas caigan en el patrón de la tecnología necesaria para la codificación o decodificación de video HDR basada en LDR.

La propuesta genérica de Philips TM-AVC0671 presentada en el foro de estandarización DVB, es simplemente un sistema de caja negra de alto nivel que enseña a nivel empresarial que hemos desarrollado una posibilidad (modo-ii) para codificar imágenes HDR al comunicar realmente imágenes LDR (+ el tipo correcto de metadatos funcionales, tuvimos que investigar mucho sobre un invento, para que eso funcione de repente). Sin embargo, no proporciona detalles específicos sobre los componentes necesarios, tal como se presenta en esta solicitud de patente. Este solo dice que hay una caja de "convertidor de rango dinámico", que una persona suficientemente capacitada entendería sobre lo que haría en una topología de sistema de este tipo, pero no se revelaron detalles de las partes internas. También se muestra un volcado de pantalla de nuestra herramienta de clasificación, que un clasificador humano en el lado de la creación podría usar para especificar todas las funciones necesarias, como también en la presente invención, pero el solicitante tuvo cuidado de no mostrar ningún detalle técnico que pudiera derivarse de esa simple foto.

Resumen de la invención

Necesitamos tener una codificación mejorada de las imágenes HDR, y en particular, comenzamos con la filosofía de que especialmente en el momento actual cuando todavía hay muchos sistemas LDR heredados en el campo, uno necesita algunos niveles de compatibilidad. Esto significa, por un lado, que nos gustaría seguir utilizando los IC codificadores existentes que implementan funcionalidades como (I) DCT [= compatibilidad de primer nivel con tecnologías de comunicación de imágenes]. Pero, además, debe haber compatibilidad de segundo nivel con pantallas que, debido a su bajo brillo máximo, necesitan imágenes LDR (es decir, un aspecto LDR, no un aspecto HDR con, por ejemplo, colores demasiado oscuros en las partes más oscuras de la imagen, sino más bien con colores más oscuros que se han mejorado para una mejor visibilidad en las pantallas LDR) porque solo pueden representar imágenes LDR (es decir, el aspecto LDR correcto bajo dicha capacidad de rango dinámico de pantalla). Esto se debe a que, además de los televisores heredados actualmente implementados, en el futuro habrá un espectro de pantallas que van desde pequeñas pantallas portátiles con capacidad de bajo brillo como ordenadores portátiles u ordenadores portátiles tableta o incluso teléfonos móviles en los que el consumidor también desea ver algunas representaciones de una película HDR, hasta las pantallas HDR más avanzadas, que en el futuro pueden tener un brillo máximo de, por ejemplo, 10000 nit, y todas las pantallas entre o alrededor de aquellos. Entonces, aunque la pantalla aún puede ser heredada y simple, podría ser servida por un nuevo IC de decodificación y mapeo de color de alta complejidad en, por ejemplo, un futuro de codificador u ordenador que suministre el contenido HDR a través de, por ejemplo, un HDMI u otra conexión, ese decodificador que ofrece cualquier combinación de las opciones que inventamos y describimos. Tenga en cuenta que una imagen LDR heredada necesitaría cierta optimización entre el contraste intraobjetos e interobjetos. Nos gustaría ver las texturas interiores de los objetos con claridad, pero también queremos tener en la imagen LDR una impresión del gran aspecto de contraste HDR de la escena original. Es decir, la diferencia entre una región de alto y bajo brillo puede no ser perfectamente representable con la imagen LDR, sin embargo, aún debe haber un remanente de esto, haciendo que los cambios de iluminación en la escena sean lo óptimos posible transmitidos en LDR por el clasificador humano.

Hemos convertido estos requisitos en un enfoque en el que, en el escenario ideal, se necesitarían (al menos) dos clasificaciones para la misma película o fotos del proveedor de contenido, que simplemente llamaremos una imagen LDR (para usar en escenarios de visualización LDR), por ejemplo, con pantallas con un brillo máximo de alrededor de 100 nit) y una imagen HDR (para las pantallas más brillantes, por ejemplo, una pantalla de referencia de brillo máximo de 5000 nit).

Por lo tanto, para varios escenarios de ejemplo prácticos, tenemos como punto de partida para las nuevas realizaciones de codificación HDR como entrada de una imagen clasificada HDR maestra (digamos que se clasifica a voluntad de acuerdo con el gusto del creador con cualquier software de procesamiento de color y, por ejemplo, codificado en un color de inicio que se codifica como OpenEXR, e incluso puede ser una actualización a un aspecto HDR de una imagen que fue capturada originalmente como LDR, por ejemplo, mediante la adición de efectos gráficos de ordenador). Luego, necesitamos codificar este HDR maestro (M_HDR) de una manera que sea prácticamente utilizable para las tecnologías actuales de codificación de video o imagen (es decir, solo ligeramente modificado de la

forma normal para usar tales tecnologías de codificación que puede implicar una redefinición de los significados del código, es decir, las luminancias respectivas codificadas por varios códigos de luma, pero no es que, por ejemplo, todos los buses deben cambiarse a 12 bits, es decir, nuestros métodos deberían funcionar con hardware de 12 bits, pero también si solo hay 10 bits por componente disponibles, o si uno acepta alguna calidad inferior incluso en sistemas de 8 bits), por ejemplo un nuevo reproductor de discos BD, o un IC de televisión que recibe video transmitido por Internet, o cualquier receptor conectado a cualquier fuente de imagen que cumpla en gran medida con una variante de las tecnologías actuales de codificación de imagen/video.

Hemos llegado a la comprensión importante de que una imagen HDR puede codificarse como una imagen de aspecto LDR (es decir, una imagen que tiene poco o ningún procesamiento colorimétrico, tal vez una conversión a otro espacio de color, pero no cualquiera o mucho mapeo de tono para convertir los brillos de los objetos de imagen para que sean más adecuados para una pantalla con otro rango dinámico de luminancia (se puede usar directamente para una visualización de buena calidad en una pantalla LDR) si solo uno agrega los parámetros de las funciones de mapeo de color que pueden convertir ese aspecto LDR en una imagen de aspecto HDR (nuestro modo ii). El lector debe contemplar que esto no es algo trivial, incluso teóricamente, ciertamente no a priori, pero incluso después de haber establecido la tarea técnica (porque sin el desarrollo posterior correcto parecería algo contradictorio codificar un aspecto a través de otro aspecto que se supone que es diferente). En particular, dado que muchas de nuestras realizaciones comienzan desde un M_HDR existente, las funciones pueden mapear la imagen pixelada de aspecto LDR codificada en una reconstrucción cercana Rec_HDR del M_HDR. Pero, por supuesto, esto no puede hacerse genéricamente de una manera ad hoc en particular, es decir, se necesita una cadena de codificación técnica específica.

Nuestra invención puede realizarse por ejemplo al menos de las siguientes maneras: Un método de codificar una imagen de alto rango dinámico (M_HDR), que comprende los pasos de:

- convertir la imagen de alto rango dinámico en una imagen de rango dinámico de luminancia más baja (LDR_o) aplicando: a) normalización de la imagen de alto rango dinámico a una escala de la luma siendo el eje [0.1] produciendo una imagen normalizada de alto rango dinámico con colores normalizados con luminancias normalizadas (Yn_HDR), b) calcular una función gamma en las luminancias normalizadas produciendo luminancias convertidas en gamma (xg), c) aplicar un primer mapeo de tonos que aumenta esas luminancias de píxeles convertidas a gamma de la imagen normalizada de alto rango dinámico que se encuentran por debajo de 0.1 con una cantidad predeterminada entre 1.5 y 5.0, produciendo lumas (v), y d) aplicar una función de mapeo de tonos arbitrariamente creciente monotónicamente mapeando las lumas resultante de realizar los pasos b y c para generar lumas (Yn_LDR) de la imagen de rango dinámico inferior (LDR_o); y

- emitir en una señal de imagen (S_im) una codificación de los colores de píxeles de la imagen de rango dinámico de luminancia inferior (LDR_o), y

- emitir en la señal de imagen (S_im) valores que codifican las formas de función de las conversiones de color anteriores b a d como metadatos, o valores para sus funciones inversas, cuyos metadatos permiten a un receptor reconstruir una imagen reconstruida de alto rango dinámico (Rec_HDR) a partir de la imagen de menor rango dinámico de luminancia (LDR_o).

Esta combinación especial de funciones de conversión luma ha demostrado ser una muy buena manera de manejar la codificación de imágenes HDR en la mentalidad del modo ii, es decir, en particular en la forma dual de codificar imágenes HDR como imágenes de aspecto LDR derivadas de estas funciones de conversión de un HDR maestro, cuyas imágenes LDR tienen el doble propósito de codificar fielmente el aspecto HDR, además de ser imágenes de aspecto LDR bien utilizables.

Tenga en cuenta que se puede utilizar cualquier codificación de los colores de los píxeles, es decir, representar los colores de los píxeles en algún sistema de codificación de espacio de color, y algunos serán más pragmáticos que otros. Por ejemplo, el LDR_o podría salir como una imagen (R'G'B') [en la que los guiones indican algún mapeo no lineal de componentes RGB lineales]. Aclaramos con un ejemplo que puede codificar la imagen LDR para ser comunicada a un receptor de una manera Yu'v', y luego también el procesamiento como el mapeo de tonos puede realizarse en una representación Yxy con coordenadas de cromaticidad xy como por ejemplo entonces u'v', pero los mismos principios de la invención a continuación también pueden incorporarse en otras representaciones de color, por ejemplo, representaciones lineales RGB (es decir, los cálculos se realizan entonces con componentes RGB directamente en lugar de componentes Y), etc. Además, la persona experta comprenderá cuál de varias maneras se puede usar para cocodificar los parámetros que caracterizan los mapeos funcionales (que pueden ser, por ejemplo, una función multilineal definida por sus puntos de cambio de segmento), que generalmente serán metadatos en o asociable con la señal de imagen S_im (por ejemplo, una estructura de mensaje SEI o similar), lo que significa que en el momento en que un receptor necesita los parámetros para dar sentido a los datos de color de píxeles codificados para transformarlos, por ejemplo una imagen de salida RGB lineal para representar en una pantalla conectada, debe haber obtenido esos parámetros de definición de aspecto mediante alguna tecnología de comunicación de datos, por ejemplo, una ubicación de memoria conectable a través de algún enlace de comunicación.

Esta combinación particular de un mapeo de “sensibilidad” que ilumina al menos los colores de objetos más oscuros en una imagen (en un subrango de los colores de píxeles más oscuros, que en la práctica se puede definir como el 10% de luminancias más bajas de una imagen lineal normalizada, que puede ser, por ejemplo, un canal Y, o los valores RGB lineales más bajos correspondientes) y una función gamma/potencia funciona bien para manejar las características colorimétricas de las imágenes HDR, y en particular su falta de coincidencia típica con la representación LDR. Por supuesto, uno podría imaginar muchas representaciones LDR, por ejemplo, el trivial que simplemente ignora todos los colores que se consideran demasiado brillantes u oscuros por el recorte, pero ese no es necesariamente el mejor aspecto LDR, y ciertamente no es utilizable para una reconstrucción de aspecto HDR de buena calidad.

Como una imagen HDR (directamente en el modo i) y una imagen LDR (en particular, una imagen LDR que realmente está codificando una imagen HDR) pueden codificarse realmente en un contenedor HEVC similar, por ejemplo, de de 3×10 bits, uno podría preguntarse cuál es la diferencia entre una imagen HDR y LDR. Esta diferencia es una diferencia colorimétrica, que es muy importante cuando se trata de una pantalla, un entorno de visualización y un observador. Matemáticamente, se puede medir a partir de los valores típicos de píxeles en la imagen normalizada (es decir, una imagen LDR o HDR normalizada), y en particular el histograma. Si se trata de una imagen LDR normal, normalmente no habrá un contraste tan alto entre los colores de píxeles más oscuros y más brillantes. Por supuesto, también en las imágenes LDR puede haber valores que son blancos, así como valores que son negros (cero), pero estos corresponderán a diferentes luminancias reales para representarlas, porque hay una función de asignación de código diferente que las define. Nuestros nuevos receptores/decodificadores reconocerán la situación y aplicarán en cada caso la decodificación apropiada. Cuando decimos que la imagen LDR debe ser directamente utilizable en una pantalla heredada, queremos decir que el receptor que suministra imágenes decodificadas al receptor comprenderá/decodificará los valores de luma con una función de asignación de código heredado, es decir, típicamente el gamma 2.2 de Rec. 709. Ahora, una imagen HDR maestra (modo i) puede estar codificada por una función de asignación de código totalmente diferente, lo que significa que una luma negra de, digamos, 0.05 corresponde a un negro mucho más oscuro que para una imagen LDR, y 0.96 corresponde a un color mucho más brillante. Además, a ese modo ii, introduce un nuevo concepto adicional, que los códigos luma ahora pueden estar relacionados con el HDR para que se convierta en lumas mediante otra asignación de código, y que incluso puede ser variable dependiendo de las elecciones del clasificador (en particular, la curva personalizada). En particular, una imagen en modo i normalmente no tendrá un histograma relativamente uniforme (bien iluminado) como en la codificación LDR heredada, sino típicamente un histograma bimodal, en el que hay una cantidad de objetos más oscuros y una cantidad de píxeles mucho más brillantes (por ejemplo, los píxeles externos que serían 100 veces más brillantes en una representación de luminancia lineal, pero pueden ser, por ejemplo, 3 veces más brillantes cuando se utiliza una función de asignación de código particular, es decir, en la representación de luma utilizada en última instancia). En algunas imágenes HDR, los píxeles más brillantes también pueden ser pocos en número, por ejemplo, un par de lámparas en una escena nocturna. En una imagen en modo ii, la relación volverá a ser diferente. Todavía habrá alguna diferencia suficiente entre las regiones brillantes y oscuras (suponiendo en la simple aclaración aquí que las imágenes HDR están tan formadas), no solo porque las funciones relativamente simples pueden mapearse en Rec_HDR, sino también porque incluso la representación directa LDR uno puede desear conservar algo del aspecto de contraste. Pero, por otro lado, los dos rangos de luminancia pueden haberse reducido uno hacia el otro en cierta medida debido a las limitaciones de la gama LDR. Pero lo importante en todo esto es que todavía se puede ver alguna firma de si una imagen era de una escena LDR o HDR. No solo los algoritmos matemáticos de análisis de imágenes pueden analizar el aspecto del rango dinámico tal como está codificado en las imágenes (por ejemplo, para la producción de televisión en tiempo real en la que la calidad final del aspecto es menos importante que, por ejemplo, los costos de producción), para lo cual puede ser una unidad 177 de análisis de imagen usada. Pero, en general, nuestras tecnologías de codificación en su formato de más alta calidad se utilizarán con un clasificador de color humano al final de la creación, que puede ver en una pantalla HDR y LDR cómo se comporta el sistema (es decir, cómo se ve el LDR y HDR en realidad), gira los diales de su teclado de clasificación y finalmente codifique una imagen LDR y funciones de reconstrucción HDR con las que esté satisfecho. Tenga en cuenta que, por lo general, los receptores no necesitan hacer un análisis completo de la situación. Ellos no necesitan preocuparse si la imagen normalizada que recibieron fue una imagen HDR o una imagen LDR, y qué variante de imagen LDR. Solo necesitan aplicar “ciegamente” las funciones que reciben. Lo único que necesitan saber normalmente es qué definen las funciones y/o qué define la única imagen. Por lo general, la señal contendrá un indicador (IND) de qué tipo de señal es. Por supuesto, se pueden comunicar muchas cosas sobre la señal, por ejemplo, para televisores que desean realizar su propia mejora de imagen inteligente, pero lo que un receptor necesita saber mínimamente es que esta señal de codificación HDR S_im es del tipo que contiene una imagen directamente utilizable para representación LDR (si su aspecto se puede ajustar a una mejor imagen LDR por receptores que pueden mapear por tonos la imagen LDR recibida LDR_t con otras funciones de mapeo de tonos [parametrizadas con Ff1, Ff2, etc.] o no). Con esta información, el receptor sabe que, si una pantalla LDR conectada se va a servir con las imágenes apropiadas, las imágenes de aspecto LDR se pueden enviar directamente a ella, y si se trata de una pantalla HDR, primero se aplicarán las transformaciones de color para obtener las Imágenes HDR correctas para representación. El lector experto comprenderá que esto puede indicarse de varias maneras, por ejemplo, con una palabra clave como “DIRECTLDR”, o con un conjunto de números “100/5000”, que indican que la única imagen es una imagen destinada a una pantalla de 100 nit (pantalla real o de referencia) y se derivó de y es mapeable a una imagen HDR de 5000 nit (lo que no significa que no se puedan derivar otras imágenes para pantallas de otro brillo máximo de la imagen LDR con la información en los parámetros que definen las funciones de transformación de color), etc.

Si ahora miramos un poco más en detalle lo que normalmente puede ser una imagen HDR (cuando se normaliza y se clasifica a una imagen LDR en modo ii óptimo), uno debería entender cómo se clasificarán diversas escenas típicamente en un ambiente definido por una pantalla de referencia HDR con un brillo máximo de, por ejemplo, 5000 o 10000 nit.

5 Una vez más, tomando el ejemplo esclarecedor de una escena en interiores con un exterior soleado brillante, uno puede querer clasificar los colores exteriores en el M_HDR a aproximadamente gris medio HDR, es decir, alrededor del 20% de 5000 nit, es decir, + -1000 nit. Los colores interiores no se deben representar con las luminosidades interiores reales, ya que estamos viendo la película en otro entorno, por ejemplo, un ambiente tenue típico de ver televisión. Por lo tanto, definitivamente, los colores interiores no se deben representar en 1/100 de las iluminaciones de píxeles exteriores soleadas, ya que tampoco se representan exactamente, solo una copia exacta en cualquier lado receptor de lo que hubiera sido la clasificación maestra de referencia en la pantalla de referencia. Necesitamos tener en cuenta el aspecto del observador promedio adaptado, en particular que en el aspecto HDR, el interior no debe verse irrealmente oscuro. Podemos clasificar esos colores en p. eje., 1/10 de la luminancia promedio de los colores de la región de la imagen "soleado afuera", por lo que alrededor de +- 100 nit. Sin embargo, ahora ingenuamente mapeando esas lumas en una pantalla de referencia LDR de 100 nit (con un mapeo de color que es por decirlo cercano un estiramiento lineal al menos conceptualmente), los colores exteriores soleados en LDR se verían perfectos, alrededor de 20 nit y hacia arriba a blanco, pero los colores interiores se procesarían alrededor de 2 nit, que se verían como negros sicovisuales. Esta es la razón por la que se necesita hacer "algo" de optimización, lo que podría ser bastante complejo dependiendo de la complejidad de una escena HDR particular, por lo que es preferible tener un clasificador de color humano involucrado en el lado de la creación de contenido, para los aspectos de nuestro marco de codificación. Para que los colores de los interiores también se puedan ver razonablemente, podemos ponerlos algo más oscuros que el gris medio (18%), pero no demasiado en la optimización. Por lo tanto, es posible que deseemos aumentar esos colores más oscuros con un factor típicamente entre 5 y 7 (dependiendo de lo que esté en las subregiones oscuras respectivamente brillantes, por supuesto, la optimización puede ser diferente para un sótano oscuro en el que uno apenas debería ver herramientas en la pared y puede recortar la luz de la lámpara iluminándola), manteniendo los colores más brillantes por encima de eso. La figura 5 muestra dos escenarios de ejemplo de nuestra cadena de codificación HDR/LDR. Las curvas 501 y 502 muestran solo las primeras curvas típicas de mapeo de tonos ("sensibilidad"), es decir, antes de la gamma. Ellas están definidas por $v = \log(1 + (RHO - 1) * xg) / \log(RHO)$, con posibles valores normalizados para la entrada xg entre cero y 1.0, y un valor RHO óptimo en caso de que el HDR maestro se haya definido con un brillo máximo de visualización de referencia de 1000 nit para la curva 501 (lo que significa que cualquier contenido estaba en la escena capturada, las iluminancias del el objeto en el M_HDR se definen entre cero y un máximo de 1000 nit, el valor al cual, por ejemplo, se puede clasificar una chispa de soldadura o el sol), y 5000 nit para la curva 502. El valor óptimo de RHO se puede determinar de muchas maneras, como comprenderá el lector experto. P.ej., el clasificador puede elegirlo, apropiado para lo que considera un buen aspecto LDR dada la imagen específica de M_HDR. O bien, un aparato en el lado de la creación puede calcularlo automáticamente, por ejemplo, de acuerdo con la siguiente ecuación:

$$RHO = potencia^{33} \cdot (\log(1 + (33 - 1) * potencia^{PB_{HDR} / 10000.1 / GAM}) / \log(33))$$

40 En esta ecuación, PB_{HDR} es el brillo máximo de la pantalla de referencia asociada con la clasificación M_HDR (es decir, que define el rango de valores posibles, y generalmente corresponde a la PB de una pantalla real en la que el clasificador estudió y creó su aspecto HDR maestro), p. ej 1000 o 5000 nit como en la figura 5, y GAM es un valor gamma, que normalmente puede ser, por ejemplo, 2.4. Por supuesto, el aparato (o clasificador) puede desviarse de estos valores por cualquier otro algoritmo o heurístico, Por ejemplo, en caso de que se requiera un aspecto algo más brillante o más plano, etc.

50 Ahora se puede ver en la figura 5 que si uno determina los factores de aumento (en comparación con la diagonal, la luma HDR normalizada está en el eje x y la luma LDR normalizada en el eje y) para la primera/parte de mapeo de tonos de sensibilidad solo a un valor entre +- 1.5 y 4.0, uno consigue después de aplicar el mapeo gamma con un gamma de 2.4 también, aumentar los factores de alrededor de 6-7 para el 10% más oscuro de los colores (las curvas 503 resp.504 son el mapeo combinado de log y gamma), que es aproximadamente lo que uno necesita (el clasificador puede afinar más tarde de acuerdo con lo desee con su curva de asignación de tonos arbitraria, pero esta es una buena estrategia para, por ejemplo, aparatos de autoconversión, que involucran mínimamente al clasificador solo en caso de que sea necesario o deseable un ajuste fino). En general, a uno le gustaría tener un aumento genérico de +- 4-8 para las operaciones combinadas de mapeo de tonos log/gamma (es decir, unidades 602 y 603), lo que significaría que un valor de refuerzo entre 1.5 y 5.0 sería apropiado para el RHO basada en la parte de sensibilidad solamente (unidad 603). Cualquier función de mapeo de tonos para la unidad 603 que tenga tal comportamiento para los colores más oscuros satisficaría lo que necesitamos para nuestra invención, pero la ecuación basada en el registro anterior es una manera pragmática simple de realizar esto. El comportamiento para los colores más claros arriba será típicamente una compresión suave, es decir, con una forma de función que típicamente mapea de manera no lineal las luminancias más claras por encima del rango tomado por los colores más oscuros potenciados. Ahora, uno puede tener imágenes HDR muy complejas, que podrían desear otros valores, pero tales situaciones extremas pueden ser manejadas por una definición de curva arbitraria apropiada por parte del clasificador (o un algoritmo de clasificación automática). Tenga en cuenta que, en el lado de la decodificación, la cadena de procesamiento necesita ser sustancialmente invertible, para poder calcular Rec_HDR a partir de las únicas imágenes LDR comunicadas.

Sustancialmente invertible significa que no necesariamente tenemos que obtener exactamente los mismos valores de componente de color en Rec_HDR que en el M_HDR original, pero las diferencias de color deben estar dentro de un límite de tolerancia. Por lo tanto, el receptor debería ser capaz de obtener las funciones de transformación de color necesarias para actualizar al aspecto HDR Rec_HDR, ya sea que las calcule invirtiendo las funciones de desclasificación utilizadas originalmente en el lado del receptor al hacer que el LDR_o (o LDR_i) forme M_HDR y reciba el dar forma a la información de esas funciones, o al recibir directamente las funciones inversas necesarias para realizar la actualización a Rec_HDR. Esto significará, entre otras cosas, que para la función de mapeo de tonos arbitraria que el clasificador puede definir para ajustar el aspecto LDR a sus estrictas preferencias, necesitará definir una función monotónicamente creciente que relacione las lumas LDR y HDR normalizadas como la persona experta comprenderá.

La cadena técnica del modo básico ii puede funcionar de manera simple. Por ejemplo, Para algunas escenas menos críticas, el clasificador puede llenar la función arbitraria con valores predeterminados como una transformación de identidad. Tenga en cuenta también que, aunque describimos los componentes técnicos básicos necesarios en la cadena, en las realizaciones prácticas uno o más de estos bloques pueden agruparse en unidades reales que realizan las funciones. Por ejemplo, En algunas aplicaciones, puede ser deseable enviar una LUT total de todas las funciones de mapeo de color juntas, mientras que en otras aplicaciones puede ser ventajoso enviar las funciones separadas, porque la televisión (automáticamente, por ejemplo, después del análisis de la escena, o bajo el control de la interfaz de usuario por el observador) puede, por ejemplo, desea sintonizar aún más, por ejemplo, La primera función que ilumina la imagen en comparación con la sensibilidad o el valor RHO recibido a través de la tecnología de comunicación de imagen/video. Las versiones más avanzadas pueden usar algunos pasos de procesamiento adicionales, por ejemplo, el método de codificación puede determinar un valor de ganancia (gai) para mapear la luma máxima de la imagen de rango dinámico inferior (LDR_o) a un valor específico de los valores posibles en la imagen reconstruida de alto rango dinámico (Rec_HDR) y codificar ese valor de ganancia en la señal de imagen (S_im), que no debe confundirse con la escala final de colores normalizados al brillo máximo de la pantalla conectada (por ejemplo Lm = 5000 nit). Esta ganancia permite una clasificación y/o codificación más versátil.

Un método mejorado muy útil para codificar una imagen de alto rango dinámico (M_HDR) comprende: después de aplicar cualquiera de los mapeos de color anteriores para determinar la imagen de rango dinámico inferior (LDR_o), aplicar una asignación de tono técnico adicional (301) para determinar una segunda menor imagen de rango dinámico (LDR_i) que se puede usar para controlar las pantallas LDR como una imagen de conducción alternativa a la imagen de rango dinámico de luminancia inferior (LDR_o), cuyo mapeo técnico de tonos se determina al: a) determinar una primera región geométrica de la luminancia inferior imagen de rango dinámico (LDR_o) para la cual la visibilidad de las bandas en la imagen de alto rango dinámico reconstruida correspondiente (Rec_HDR) está por encima de un nivel aceptable, b) determinar un rango de lumas (L_u) para esa región, c) determinar un segundo rango de píxel lumas (L_uu) adyacente en el eje luma al rango de lumas (L_u), en el que el segundo rango se identifica para cumplir las condiciones de que tiene un número de lumas por encima de un número mínimo (MIN), y corresponde a una segunda región de imagen geométrica que contiene una textura que puede representarse utilizando menos del número mínimo de códigos en una imagen LDR (LDR_i) sobre la cual aplicar las funciones que producen una imagen reconstruida de alto rango dinámico (Rec_HDR) de calidad visual suficiente para esa segunda región, y d) determinar una función de mapeo de redistribución que redistribuye las lumas del primer y segundo rango de lumas, de modo que hay códigos adicionales disponibles para el primer rango, y emitir en la señal de imagen (S_im) valores que codifican la función forma de la función de mapeo de redistribución o preferiblemente su inversa.

Uno es algo limitado en la compensación entre la reconstrucción completa o suficientemente precisa de Rec_HDR, y el aspecto de la imagen LDR LDR_o, en particular si el hardware (y el costo de clasificación) dicta que se debe usar una cantidad relativamente limitada de funciones de clasificación. Algunas escenas HDR pueden no ser tan difíciles (por ejemplo, el observador promedio puede no ser demasiado crítico sobre si las sombras de un lado oscuro de una calle iluminada por el sol son un poco más oscuras o un poco más grises claras, siempre que las desviaciones de lo óptimo el aspecto no es demasiado excesivo), pero algunas escenas HDR pueden ser más críticas (por ejemplo, en algún lugar del rango de luminancia HDR puede haber un tipo parcialmente oculto en la niebla luminosa, y si el contraste local es demasiado alto, puede ser demasiado visible, pero si el contraste es demasiado bajo, puede ser invisible, cambiando la historia). Sería ventajoso tener otra dimensión de clasificación posible, al menos para receptores que no son heredados (y no saben cómo hacer ningún procesamiento HDR), y que puedan hacer más mapeo de tonos. Una pantalla heredada podría obtener la imagen LDR de "mejor esfuerzo", que será la única imagen transmitida, pero los receptores futuros inteligentes podrían hacer algunos trucos técnicos inteligentes para optimizar aún más el aspecto LDR, de modo que se acerque a lo que desea el clasificador (tal vez incluso recortar algunos valores en el aspecto LDR definitivo, lo que sería incongruente con la reconstrucción HDR si eso sucediera en la única imagen LDR transmitida). Teniendo tal posibilidad, algunos métodos de codificación o codificadores podrían atender esto. Exprimir una imagen HDR de muy alta relación de contraste muy compleja en una única imagen LDR (por ejemplo, una imagen HDR que tiene varias regiones importantes con muchos valores grises, por ejemplo, una habitación oscura sin luz, una segunda habitación relativamente bien iluminada y, al mismo tiempo, una soleada colorida afuera, y con estas 3 regiones que también contienen gradientes importantes que cubren muchos valores grises, por ejemplo, de una mesa blanca debajo de la lámpara en la habitación bien iluminada), podría suceder que una o más regiones se vuelvan inaceptables, debido a la palabra limitada longitud (por ejemplo, 10 bits) para los componentes de color, en algún lugar (dependiendo de las formas de las funciones de mapeo de color) hay bandas

que se consideran demasiado severas. Esta región se puede identificar, por ejemplo, por el clasificador que lo detecta (y puede que lo ayude señalando áreas potencialmente críticas mediante el software de análisis de imágenes HDR en el aparato de clasificación). Los detectores de bandas pueden calcular, por ejemplo, que para una región extendida (potencialmente también teniendo en cuenta qué luminancias tiene esta región y los JND estimados) hay saltos de cada vez una cantidad de colores sucesivamente iguales, y pueden definir un nivel aceptable basado en los valores de dicho cálculo (y típico en experimentos de fábrica). El aparato de clasificación después de haber encontrado dicha región (por ejemplo, al segmentar más finamente lo que el clasificador ha seleccionado aproximadamente), puede determinar aproximadamente el rango L_u de luminancias correspondiente a este. Por ejemplo, puede haber bandas en un cielo azul, cuyos colores tienen luminancias entre L_{sky_low} y L_{sky_high} . El problema se mitigaría si la codificación LDR tuviera más valores para codificar la imagen, donde deberíamos entender que en el lado de la codificación, el M_HDR y cualquier transformación pueden ser de muy alta precisión. Pero estos códigos no existen: solo tenemos los 10 bits disponibles para todas las luminancias necesarias, y también necesitamos codificar suficientemente todas las demás regiones de imagen de iluminación diferente. Pero se puede usar un truco, si se pueden tomar prestados algunos códigos de regiones de la imagen que tienen luminancias adyacentes a L_u , especialmente si la calidad visual de esas regiones se degrada poco al tomar algunos códigos de su rango de códigos (lo que normalmente hará el clasificador) juzgue, con una simple operación de aceptar el resultado, o en desacuerdo en cuyo caso se ensayará otro intento, que es más agresivo en caso de que las bandas sigan siendo demasiado altas para el área con bandas originales y la región adyacente aún pueda deteriorarse más, o un poco menos agresiva si el clasificador indica que la región adyacente comienza a deteriorarse demasiado). Una manera simple de redistribuir códigos es, por ejemplo, Una modificación lineal o no lineal de la parte de la función local. Ahora el problema con la única imagen transmitida LDR_o , es que el cielo puede, por ejemplo, haberse vuelto un poco demasiado oscuro y quizás demasiado contrastado por esta operación (y también las regiones adyacentes pueden estar algo oscuras y su aspecto de textura puede haber cambiado, etc.). Esto puede no ser demasiado problemático en caso de pequeños cambios y escenas menos críticas, y un poco más inconveniente para escenas difíciles. Es el precio que los sistemas heredados pueden tener que pagar porque no pueden hacer absolutamente nada con ninguno de los datos recibidos, excepto representar directamente el LDR_o , pero los nuevos receptores pueden aplicar el inverso de las transformaciones utilizadas para redistribuir las lumas, para crear un aspecto LDR muy cercana de la original (es decir, con las luminancias del cielo apropiadas, etc.), pero ahora con menos bandas. Un receptor no necesita hacer mucho análisis inteligente, solo necesita ver que dicha función técnica de mapeo de tonos esté disponible y aplicarla a la reconstrucción de la única imagen LDR transmitida LDR_t para obtener la mejor imagen de aspecto LDR LDR_{ul} . También se pueden aplicar varios métodos en los aparatos de clasificación para llegar a buenas sugerencias para la región adyacente, por ejemplo, se puede determinar una región con una cantidad suficiente de lumas (por ejemplo, igual a la cantidad en el cielo) y con cierta textura compleja. Realizaciones simples pueden, por ejemplo, usar todos los códigos por debajo del rango de las regiones con bandas, hasta el negro más negro.

La cantidad de códigos adicionales para el primer rango se determina en función de un criterio de visibilidad de bandas para la primera región geométrica. Un algoritmo automático puede venir con una propuesta, por ejemplo, 20% de códigos adicionales, y típicamente el clasificador humano lo reconocerá. El algoritmo puede también resaltar las regiones que esta tuvo que deteriorar, por ejemplo, parpadeando un color de esas regiones, para que el clasificador pueda verificar rápidamente si son de calidad visual suficiente también en el HDR Rec_HDR reconstruido.

En la mayoría de las realizaciones prácticas, la identificación de la primera región geométrica que muestra las bandas excesivas es realizada en última instancia por un clasificador humano a través de una unidad de interfaz de usuario (105), por ejemplo, garabateando una línea ondulada a lo largo de la región con bandas, y la cantidad de bandas de la primera región geométrica en la imagen reconstruida de alto rango dinámico (Rec_HDR), y la calidad visual de la reconstrucción de la segunda región geométrica en la imagen reconstruida de alto rango dinámico (Rec_HDR) son juzgados por el clasificador humano como aceptables o inaceptables, en donde en el caso del juicio aceptable, los valores que codifican la forma de la función de mapeo de redistribución o su inverso se codifican en la señal de imagen, o en el caso del juicio inaceptable, los pasos se realizan nuevamente con diferentes parámetros para llegar a una función de mapeo de redistribución alternativa. Por ejemplo, se puede asignar un 10% más de códigos a la región en bandas, tal vez a expensas de un rango de luma adyacente ampliado L_{uu} .

Una realización interesante del método de codificación de una imagen de alto rango dinámico (M_HDR) tiene los colores de píxeles de la imagen de menor rango dinámico de luminancia (LDR_o) están codificados como un canal de luma y coordenadas de color u' y v' que se calculan como $u' = 4X/(X+15Y+3Z)$ y $v' = 9Y/(X+15Y+3Z)$, siendo X , Y y Z las coordenadas de color CIE de 1931 independientes del dispositivo que se pueden derivar para cualquier representación RGB (es decir, la representación de cromaticidad y CIE 1976 (u' , v')). Normalmente de acuerdo con la filosofía heredada, las imágenes (especialmente las imágenes LDR) se codificarían como imágenes YCrCb. Pero si uno cambia cualquier códec (por ejemplo, para la transmisión por Internet), también podría codificar los componentes de color como planos de componentes Yuv, lo que tiene algunas ventajas tanto en la calidad de imagen de las imágenes transmitidas como en la facilidad de aplicar las diversas transformaciones de color de nuestro sistema (por supuesto, los televisores heredados no podrán hacer buenas fotos con esto).

Hemos encontrado una definición de luma elegida genéricamente (definida por la estrategia de mapeo de tono completo elegida de los pasos anteriores, obteniendo finalmente lumas en LDR_o o LDR_i) que, junto con dos coordenadas de cromaticidad independientes de luma, en particular las coordenadas u' , v' estandarizadas por la CIE,

será una buena codificación para ser utilizada en tecnologías estándar de compresión de imagen o video. Tecnologías de compresión similares a, por ejemplo, HEVC generalmente aplicará al menos compresión espacial al hacer DCT de bloques de muestras, pero para video también pueden hacer compresión basada en estimación de movimiento, etc.

5 En realizaciones simples, la codificación del comportamiento de transformación de color funcional de las conversiones de color se puede comunicar con solo unos pocos parámetros simples almacenando en los metadatos asociados o asociables con la(s) única(s) imagen(es): a) un valor de sensibilidad (por ejemplo RHO, o un parámetro equivalente que define RHO llamado SENS y se define a continuación, o cualquier función o correlato de RHO que permita determinar un valor RHO), b) un valor gamma (GAM), y c) una serie de valores que caracterizan la forma funcional de la función arbitraria mapeando las lumas.

10 Un método para codificar una imagen de alto rango dinámico (M_HDR) que comprende determinar un valor de ganancia para mapear la luma máxima de la imagen de rango dinámico inferior (LDR_o) a un valor específico de los valores posibles en la imagen reconstruida de alto rango dinámico (Rec_HDR), y codificar ese valor de ganancia en la señal de imagen (S_im). Esto es útil para una mayor ampliación de la imagen Rec_HDR obtenible de la imagen LDR, por ejemplo, si la imagen LDR de una toma relativamente oscura se representa con un brillo relativamente alto, es decir, qué lumas hasta un valor relativamente alto en el rango de lumas LDR posibles, sin embargo, la imagen HDR debe representarse de manera no muy brillante, lo que se maneja mejor al decodificarla con una luma máxima no demasiado alta.

15 Un método para codificar una imagen de alto rango dinámico (M_HDR) que comprende determinar una estrategia de modificación de saturación, ya sea de los colores de la imagen de alto rango dinámico (M_HDR) a los colores en la imagen de rango dinámico inferior (LDR_o), o al revés, y codificando esta estrategia de modificación de saturación como valores paramétricos en metadatos en la señal (S_im). Por lo general, los clasificadores también querrán afectar la saturación de una imagen, por ejemplo, pueden cambiar la saturación del LDR_o obtenido de M_HDR con alguna estrategia de mapeo de saturación, y/o la saturación de Rec_HDR de LDR_o (por ejemplo, primero un mapeo de tonos que deja las cromaticidades u, v de los colores Rec_HDR obtenidos en los valores que tenían en LDR_o, y luego cambiando las saturaciones de esos colores Rec_HDR).

20 Algunas realizaciones del método de codificación de una imagen de alto rango dinámico (M_HDR) comprenden: después de aplicar cualquiera de los mapeos de color anteriores para determinar la imagen de rango dinámico inferior (LDR_o), aplicar un mapeo (301) de tono técnico adicional para determinar una segunda imagen de rango dinámico inferior con lumas redistribuidas, es decir, lumas de valores típicamente ligeramente modificadas para al menos una región geométrica y un subrango de luma de la imagen (es decir, una imagen de bajo rango dinámico luma redistribuido LDR_i), lo que garantiza que al menos en el nivelador sea más importante regiones de la segunda imagen de rango dinámico inferior (LDR_i), por ejemplo, que son observados cuidadosamente por el observador típico, porque son, por ejemplo, grandes y brillantes, y propensos a las bandas, se pueden asignar suficientes códigos luma para codificar las texturas en esas regiones con suficiente precisión para permitir la reconstrucción de la imagen reconstruida de alto rango dinámico (Rec_HDR) con errores por debajo de un criterio de error predeterminado (una cantidad mínima de bandas).

25 Es importante para algunas realizaciones que uno no elija ninguna estrategia de mapeo de tonos extraño. En particular, si uno quiere poder obtener una Rec_HDR de buena calidad, es decir, un valor de color de píxel matemático cercano a M_HDR, entonces debe asegurarse de que no haya texturas submuestreadas en LDR_i, lo que sucede si se asegura que el mapeo final antes de la cuantificación uniforme no es demasiado plano.

30 Por lo general, esto se puede hacer mediante estrategias de conteo de luma en la imagen LDR y/o estrategias de conteo de luma, como por ejemplo un detector de bandas en la imagen HDR, o cualquier criterio de error de reconstrucción HDR predeterminado. En algunas realizaciones, el criterio puede ser realizado por un clasificador humano. Su presencia se puede ver al tener una estrategia de remapeo técnico cocodificado en S_im, para ser aplicado por los receptores más inteligentes de la generación futura.

35 El método puede realizarse en un codificador (100) de imagen dispuesto para codificar una imagen de alto rango dinámico (M_HDR), que comprende:

40 - Una unidad (104) de conversión de rango dinámico dispuesta para convertir la imagen de alto rango dinámico en una imagen de menor rango dinámico de luminancia (LDR_o), la unidad (104) de conversión de rango dinámico comprende que conectada en orden de procesamiento: a) un normalizador (601) dispuesto para normalizar la imagen de alto rango dinámico a un eje luma que se extiende sobre [0.1] y para emitir luminancias normalizadas (Yn_HDR), b) una unidad (602) de conversión gamma dispuesta para aplicar una función gamma a las luminancias normalizadas y para emitir luminancias (xg) convertidas a gamma, c) una primera unidad (603) de mapeo de tonos dispuesta para aplicar un primer mapeo de tonos que aumenta esas luminancias convertidas en gamma que están por debajo de 0.1 con una cantidad predeterminada que descansa entre 1.5 y 5.0, produciendo lumas (v), d) una unidad (604) de mapeo de tonos arbitraria dispuesta para d) aplicar una función arbitraria que mapea las lumas (v) para generar lumas (Yn_LDR) de la imagen de rango dinámico inferior (LDR_o); y el codificador (100) de imagen que comprende además:

5 - un compresor (108) de imagen dispuesto para aplicar una transformación de reducción de datos a los colores de la imagen de rango dinámico inferior (LDR_o) cuyos colores están organizados en imágenes componentes, y cuya transformación de reducción implica aplicar al menos una transformación DCT a bloques de valores de componentes de color adyacentes, produciendo una codificación comprimida (LDR_c) de los colores de píxeles de la imagen de rango dinámico de luminancia inferior; y

10 - un formateador (110) dispuesto para emitir en una señal de imagen (S_im) la codificación comprimida (LDR_c), y dispuesto además para emitir en la señal de imagen (S_im) los valores que codifican la forma de la función de las conversiones de color como metadatos, o valores para sus funciones inversas, cuyos metadatos permiten a un receptor reconstruir una imagen de alto rango dinámico (Rec_HDR) basada en la imagen de rango dinámico de luminancia más baja (LDR_o).

15 Una realización pragmática de dicho codificador es aquella en la que la unidad (602) de conversión gamma usa un valor gamma igual a $1/(2.4)$, y/o la primera unidad (603) de mapeo de tonos usa un mapeo de tonos definido por la ecuación $v = \log(1 + (RHO1) * xg) / \log(RHO)$, con RHO que tiene un valor predeterminado, cuyo valor es típicamente una función del brillo máximo de la pantalla de servicio prevista y/o la pantalla de referencia asociada con la codificación HDR maestra M_HDR.

20 Un codificador (100) de imagen dispuesto para especificar una ganancia que permite mapear el máximo de los códigos luma en la imagen de rango dinámico inferior (LDR_o) a un valor luma elegido de la imagen reconstruida de alto rango dinámico (Rec_HDR), y que tiene el formateador (110) dispuesto para generar esta ganancia como un valor en metadatos en la señal de imagen (S_im).

25 Un codificador (100) de imagen de acuerdo con cualquiera de las reivindicaciones de codificador anteriores que comprende una unidad (106) técnica de mapeo de tonos, dispuesta para determinar automáticamente o guiada por humanos la textura y la información estadística de la imagen de rango dinámico inferior (LDR_o), y en particular al menos una región geométrica crítica que es propensa a errores de reconstrucción en particular en bandas en Rec_HDR, y sobre la base de ello calcular un segundo mapeo de tonos (Ff1, Ff2, ...) para ser aplicado como una transformación a la imagen de rango dinámico inferior (LDR_o) para obtener una segunda imagen de rango dinámico inferior (LDR_i) que tiene un número mínimo de códigos luma (por ejemplo $1.3 * L_u$) que caracteriza las texturas de al menos algunas regiones importantes y propensas a errores de la segunda imagen de rango dinámico inferior (LDR_i), lo que permite reconstruir la imagen reconstruida de alto rango dinámico (Rec_HDR) con errores por debajo de un criterio de error predeterminado. Para posibilitar la comunicación de la información necesaria que permite, después de codificar un receptor, implementar a manera de espejo nuestro sistema a modo ii, es útil transmitir (o almacenar para una transmisión posterior) una señal de imagen de alto rango dinámico (S_im) que comprende:

- una imagen pixelada de rango dinámico inferior (LDR_o) con colores de píxeles codificados; y además:

40 - un valor de sensibilidad (RHO); y

- un valor gamma (GAM); y

- un valor de ganancia (GAD); y

45 - un conjunto de valores que especifican una forma arbitraria de función de mapeo de tonos (P_CC).

50 A partir de estos valores, el receptor puede determinar las formas de función de todas las funciones que se aplicarán a la única imagen LDR comunicada (LDR_o o LDR_i), en caso de que se requiera y calcule cualquier imagen de rango dinámico superior a la imagen LDR de 100 nit.

55 En particular, el S_im también puede comprender valores 207 que codifican una estrategia de remapeo técnico (Ff1, Ff2,...) Para el mapeo entre una clasificación artística LDR, de acuerdo con lo desee el creador/clasificador humano del contenido, y una LDR técnica que, cuando se muestrea, tiene suficientes lumas para todas las regiones de la imagen para una buena reconstrucción de Rec_HDR, o al menos aquellas regiones determinadas como más críticas por una unidad de análisis de imagen automática y/o un humano.

60 En particular, es útil, porque es muy pragmático para los receptores determinar rápidamente cuál de los ahora varios (muy) diferentes posibles mecanismos de codificación de imagen HDR se utilizan, en particular al comprender en la señal de imagen S_im un indicador (IND) que especifica que una imagen de alto rango dinámico ha sido codificada en esta, y con un método que lo codifica como una imagen de bajo rango dinámico, que se puede usar directamente, sin necesidad de más mapeo de tonos, para representar en una pantalla LDR. Se pueden dividir y acordar diversas formas de codificación, siempre que cualquier receptor lo entienda.

65 Un producto de memoria como un disco blu-ray que almacena cualquier realización de nuestra señal de imagen de alto rango dinámico (S_im).

Para tener una cadena de comunicación de imagen, en el extremo receptor se pueden tener varias realizaciones de aparatos que son o comprenden un decodificador (150) de imagen dispuesto para recibir una señal de imagen de alto rango dinámico (S_im) y que comprende:

5 - un desformateador (151) dispuesto para obtener una imagen comprimida pixelada de rango dinámico inferior (LDR_c) y datos de parámetros (P) fuera de la señal de imagen (S_im); y

10 - un descompresor (152) dispuesto para aplicar al menos una transformación DCT inversa a la imagen comprimida de rango dinámico inferior pixelada (LDR_c) para obtener una imagen pixelada de rango dinámico inferior (LDR_t); y una unidad (153) de conversión de rango dinámico dispuesta para transformar la imagen de rango dinámico inferior (LDR_t) en una imagen reconstruida de alto rango dinámico (Rec_HDR), en donde la unidad (153) de conversión de rango dinámico comprende en orden de procesamiento:

15 a) una unidad de mapeo (402) de tono arbitrario dispuesta para aplicar un mapeo de tonos arbitrario, los parámetros que lo definen (P_CC) se reciben en los datos de parámetros (P), b) una primera unidad (403) de mapeo de tonos dispuesta para aplicar un mapeo como se define en al menos un parámetro recibido (RHO) que define el primer mapeo de tonos que fue determinado previamente por cualquiera de nuestras realizaciones de codificador o método de codificación, y c) una unidad (404) de conversión gamma dispuesta para aplicar un mapeo gamma con un valor gamma recibido (GAM).

20 Este decodificador primero desharrá todo el legado típico, por ejemplo, HEVC o codificaciones de compresión similares, y luego aplicará los diversos mapeos en orden inverso (tenga en cuenta que no todo lo que necesita en todas las realizaciones debe ser exactamente en el orden inverso; por ejemplo, en Yu'v' uno puede elegir hacer el procesamiento de la luma ortogonal y la saturación en orden inverso, ya sea con funciones matemáticas ligeramente diferentes, siempre que el resultado final sea exactamente o aproximadamente el color deseado). Tenga en cuenta también que puede haber pasos de procesamiento adicionales, que solo pueden existir en el extremo receptor (por ejemplo, una imagen puede estar codificada en alguna representación RGB como la Rec.2020, pero puede necesitar convertirse a otro formato tal como lo entiende un televisor, por ejemplo, DCI-P3, y además convertido a las primarias reales de la TV).

25 Por lo tanto, el decodificador (150) de imagen comprenderá una unidad (153) de conversión de rango dinámico dispuesta para transformar la imagen de rango dinámico inferior (LDR_t) en una imagen reconstruida de alto rango dinámico (Rec_HDR), y normalmente puede haber típicamente unidades lógicas y funciones de procesamiento de color adicional que definen al menos cuándo hacer qué (por ejemplo, dependiendo de qué pantalla o pantallas están conectadas y servidas actualmente).

30 Una realización pragmática del decodificador de imagen tiene la primera unidad (403) de mapeo de tonos dispuesta para aplicar una función de la forma: $xg=(potencia(RHO,v)-1)/(RHO-1)$, en el que v es un pixel luma, y RHO es un parámetro de valor real o entero recibido en los datos del parámetro (P).

35 Una realización útil del decodificador (150) de imágenes comprende una unidad (159) de remapeo de tonos dispuesta para aplicar un mapeo de tonos adicional (Ff1, Ff2, ...) recibido en la señal de imagen (S_im) a la imagen de rango dinámico inferior (LDR_t) para obtener una segunda imagen de rango dinámico inferior (LDR_ul) que invierte una acción de redistribución de código, aplicada por cualquiera de los métodos de codificador 5 a 7, produciendo una segunda imagen de rango dinámico bajo (LDR_i) con lumas redistribuidas para obtener una banda reducida en al menos una región de la imagen reconstruida de alto rango dinámico (Rec_HDR). De hecho, el codificador no necesariamente necesita saber exactamente cómo un codificador llegó a una función de transformación particular redistribuyendo las lumas, solo necesita aplicar las funciones inversas para obtener sustancialmente el aspecto LDR deseado (LDR_ul).

40 Otra realización útil del decodificador puede entender las codificaciones Yu'v' de la imagen LDR, y comprende una unidad (155) de transformación de color dispuesta para convertir una representación de color Yu'v' en una representación de color RGB. El mapeo de tonos se puede hacer antes de que se realice la conversión, por lo tanto, dejar la conversión a RGB a la última parte de la cadena de procesamiento, o alternativamente, la conversión se puede hacer primero y se puede hacer un procesamiento de color equivalente en las señales RGB.

45 Correspondiendo a cualquiera de los decodificadores corresponden métodos de decodificar una señal de imagen de alto rango dinámico (S_im) que comprende obtener una imagen reconstruida de alto rango dinámico (Rec_HDR) aplicando las conversiones de color codificadas en los datos de parámetros (P) a la imagen de rango dinámico inferior (LDR_t), en particular un método de decodificación de una señal de imagen de alto rango dinámico (S_im) que comprende:

50 - obtener una imagen comprimida pixelada de menor rango dinámico (LDR_c) y datos de parámetros (P) fuera de la señal de imagen (S_im); descomprimir la imagen comprimida de rango dinámico inferior pixelada (LDR_c) aplicando al menos una transformación DCT inversa a la imagen comprimida de rango dinámico inferior pixelada (LDR_c) para obtener una imagen pixelada de rango dinámico inferior (LDR_t); y transformar la imagen de rango dinámico inferior

(LDR_t) en una imagen reconstruida de alto rango dinámico (Rec_HDR), mediante: a) aplicar un mapeo de tonos arbitrario, los parámetros que lo definen (P_{CC}) se reciben en los datos de parámetros (P), b) aplicar un mapeo de acuerdo con lo definido por al menos un parámetro recibido (RHO) que define el primer mapeo de tonos que fue determinado previamente por cualquiera de nuestras realizaciones de codificador o método de codificación, y c) aplicar un mapeo gamma con un valor gamma recibido (GAM), que es preferiblemente igual a 2.4. Describimos un sistema que permite al clasificador optimizar de manera simple pero poderosa el aspecto de un aspecto LDR en una imagen HDR de una escena HDR. Preferiblemente, se hace el menor sacrificio de calidad visual posible, pero dado que LDR puede necesitar alguna optimización debido a sus limitaciones de rango dinámico, el sistema permite al clasificador ajustar microcontrastes de objetos de escena interesantes, es decir, objetos característicos típicamente importantes en esta escena y, por lo tanto, si es necesario sacrificar alguna calidad del brillo, sacrificar el aspecto preciso de algunos objetos menos importantes, como una pared en el fondo, en lugar del objeto principal de la escena. La invención se puede realizar de muchas otras maneras (parciales) como con intermedios que contienen los requisitos técnicos centrales de las diversas realizaciones, como los parámetros de definición incorporados en las señales, y muchas aplicaciones de la misma pueden resultar, como varias formas de comunicarse, usar, transformar el color, etc., las diversas señales posibles y diversas formas de incorporar los diversos componentes de hardware, o utilizar los diversos métodos, en sistemas de consumo o profesionales. Por supuesto, cualquier componente puede realizarse en o como un componente pequeño, o viceversa, como un núcleo clave de un aparato o sistema grande que funciona principalmente debido a este componente.

20 Breve descripción de los dibujos

Estos y otros aspectos del método y aparato de acuerdo con la invención serán evidentes y se aclararán con referencia a las implementaciones y realizaciones descritas a continuación, y con referencia a los dibujos adjuntos, que sirven meramente como ilustraciones específicas no limitantes que ejemplifican el concepto más general.

25 La figura 1 muestra esquemáticamente un ejemplo de una realización de un codificador y un decodificador de acuerdo con nuestra invención en una tecnología de comunicación de imágenes;

30 La figura 2 muestra esquemáticamente una realización de cómo puede ser una señal de imagen HDR S_{im} de acuerdo con nuestra invención;

35 La figura 3 aclara esquemáticamente por una especie cómo uno puede obtener genéricamente una clasificación técnica de LDR, que incluso podría ocurrir bajo la caperuza automáticamente sin molestar al clasificador o al creador de contenido en algunas realizaciones, lo que permite un mejor muestreo de las lumas del objeto y, por lo tanto, una mejor calidad de reconstrucción de Rec_HDR;

La figura 4 es una elucidación esquemática simple de un posible decodificador de acuerdo con nuestra invención;

40 La figura 5 muestra dos posibles curvas de realización dos x para realizar el mapeo de sensibilidad que ilumina significativamente los colores, o la clasificación LDR inicial combinada, además del comportamiento gamma; y

La figura 6 es una aclaración esquemática simple de una posible unidad de conversión de rango dinámico posible de un codificador.

45 Descripción detallada de las realizaciones

La figura 1 describe un sistema típico ejemplar que incorpora nuestra invención, con un codificador 100 de imagen (o video) en un lado de creación, y un decodificador 150 de imagen. Suponemos que hay una memoria 101 en un sistema de clasificación que contiene una imagen de aspecto HDR con clasificación maestra (M_HDR) que ha sido clasificada de acuerdo con lo deseado por el creador de contenido de acuerdo con las técnicas de clasificación de color conocidas actualmente por decir una película en un software de clasificación de color como por ejemplo Da Vinci (otros sistemas similares pueden beneficiarse de las enseñanzas de nuestra presente solicitud, por ejemplo, M_HDR puede provenir directamente de una cámara, después de, por ejemplo, una sintonización de una curva de aspecto de la cámara en los diales de la cámara, etc.). En este M_HDR, por ejemplo, los brillos de luz que brillan a través de las ventanas pueden haber sido elegidos para dar un aspecto más agradable en la pantalla de referencia de [0.5000] nit al dar a esos píxeles una luminancia L_{out} prevista para ser emitida y el código v HDR luma correspondiente, y muchos efectos de luz adicionales pueden haber sido diseñados, así como otras optimizaciones de color. M_HDR se ingresa a través de la entrada 115 de imagen en nuestro codificador 100, y también puede verse en una pantalla 102 de referencia HDR (que exactamente las características de la pantalla de referencia teórica por ejemplo [0-5000] nit proponemos para la codificación HDR). Esto significa que cuando el clasificador desea hacer un aspecto LDR (que no solo debe codificar las texturas de los objetos con la suficiente precisión para que en el lado receptor se pueda obtener una reconstrucción razonablemente precisa Rec_HDR del M_HDR, sino que también este aspecto LDR debería ser adecuado para representar de manera óptima la escena HDR codificada en una pantalla LDR), el clasificador puede comparar al mismo tiempo cuánto se ve el LDR dadas las limitaciones técnicas en la pantalla 103 LDR similar al M_HDR, y optimizar cambiando las funciones de mapeo de color para obtenerlo de M_HDR de acuerdo con lo deseado de acuerdo con su gusto. Las dos pantallas pueden estar en sus diferentes entornos óptimos de visualización y el

clasificador puede mirar a ambos separados por, por ejemplo, una pared (por ejemplo, en dos entornos de referencia incluidos con sus respectivas ventanas abiertas para mirar simultáneamente hacia ellos, y con cortinas que se pueden cerrar si el clasificador quiere ver solo uno de ellos durante un intervalo de tiempo). El clasificador también puede verificar la clasificación reconstruida del aspecto HDR en la pantalla 102 HDR (por ejemplo, alternar Rec_HDR y M_HDR alternativamente).

Por medio de una unidad 105 de interfaz de usuario que ofrece al clasificador controles clásicos como por ejemplo girando ruedas o deslizadores similares para establecer valores como un valor gamma o de sensibilidad, el clasificador puede hacer transformaciones colorimétricas que definen cómo se debe mapear el M_HDR a la imagen de aspecto LDR, con los parámetros de las transformaciones que se emitirán en una señal de imagen S_im a través de un salida 116 del codificador que puede conectarse a cualquier medio 140 de transmisión de imágenes por ejemplo una red de comunicación, o una memoria portadora física como un BD o memoria de estado sólido, etc.

El aspecto LDR se genera a través de una unidad 104 de conversión de rango dinámico, que está dispuesta para aplicar transformaciones colorimétricas en al menos las lumas de los colores de píxeles, pero también típicamente en las coordenadas de cromaticidad. Por lumas nos referimos a cualquier codificación que finalmente sea convertible en una luminancia física, o incluso a través de modelos sicovisuales, un brillo (que es el último aspecto que verá el observador cuando la imagen se muestre en una pantalla). Tenga en cuenta que, por matemática equivalente, las transformaciones de luma pueden aplicarse directamente como transformaciones correspondientes en componentes RGB. Aunque el objetivo final es el brillo correcto de los objetos (aspectos) en el aspecto, podemos limitar nuestra discusión técnica a la determinación de luminancias en la referencia, por ejemplo, rango [0-5000], o un espacio de color independiente del dispositivo como XYZ definido por este rango. Además, supondremos que cualquiera de las transformaciones cromática de los colores se realiza en el plano UCS del espacio CIE Luv de 1976, sin embargo, la persona experta puede entender cómo de manera similar se pueden usar otros componentes de segundo y tercer color, siendo generalmente los componentes básicos de nuestra invención aplicable

CIELuv define u y v desde XYZ (de manera similar, uno puede transformarse desde algunos RGB) como: $u' = 4X/(X+15Y+3Z)$ y $v' = 9Y/(X+15Y+3Z)$.

Suponemos por simplicidad que las gamas HDR y LDR (es decir, las gamas de pantallas teóricas asociadas con la codificación matemática de las dos imágenes) tienen los mismos tres (o más) R, G, B, primarios y, por lo tanto, pueden escalar los máximos respectivos de digamos 5000 y 100 nit a 1.0, se colocarán como superpuestos exactamente. Entonces, un mapeo de tonos de HDR a LDR se convierte en una transformación relativa a lo largo de la dirección de luma normalizada dentro de esta gama RGB dependiente de un solo dispositivo. Por ejemplo, si se quiere que los colores más oscuros en el HDR se vean iguales en una pantalla LDR y HDR, esto se convierte en una transformación relativa en la misma gama de lo siguiente: porque en una definición de color definida de 5000 nit, tales colores en la imagen HDR tendrán códigos pequeños (por ejemplo, debajo de 0.1) necesitamos aclararlos para que sean lo suficientemente visibles en una pantalla LDR de 100 nit, por ejemplo, con valores alrededor de 0.3. El mapeo exacto dependerá de la definición de las lumas tanto para la imagen LDR como para la imagen HDR, porque como una generalización de las definiciones "gamma 2.2" de la codificación de imagen y video LDR heredado, ahora podemos definir el mapeo de funciones de asignación de código arbitrario a partir de luminancias físicas a los códigos de luma (o al revés, porque típicamente los ingenieros de televisión comienzan definiendo una pantalla de referencia que además de un rango de referencia [0-5000] nit tiene un comportamiento EOTF de pantalla de referencia que indica cómo, por ejemplo, el mapa de 1024 lumas a luminancias representables a lo largo de ese rango de referencia). No solo podríamos usar una potencia $1/(7.0)$ gamma como OETF, sino que incluso podríamos usar funciones de asignación de código discontinuas si en una captura de imágenes no hay luminancias presentes entre un rango inferior de luminancias y un rango superior de luminancias. También tenga en cuenta que trabajar en una representación $Y'uv$ con cromaticidades independientes de luma (u , v) nos permite trabajar de manera totalmente independiente y libre en las direcciones acromáticas y cromáticas del espacio de color.

Limitando nuestra aclaración para el lector experto a los mapeos acromáticos de HDR-2-LDR solamente, estos pueden formularse genéricamente como, en principio, una función de mapeo de tonos arbitraria desde las [0.1] lumas de la imagen de aspecto HDR a [0.1] lumas de la imagen de aspecto LDR, como se puede ver con un ejemplo en la figura 2a.

Al especificar dicha función, asumiremos que la asignación de todos los colores (Y_M_HDR , u , v) se realiza de modo que para un color no acromático ($u < > u_wp$, $v < > v_wp$) donde (u_wp , v_wp) son las coordenadas de cromaticidad de un punto blanco elegido como D65, la función de mapeo de tonos determinada 210 se escala linealmente a una luminancia máxima L_max (u , v) alcanzable para ese color, como se enseña con más detalle en el documento WO2014056679. El lector experto puede comprender cómo dicho procesamiento en lugar de ser aplicado en una codificación de color $Y'uv$ también se puede hacer de manera similar en una codificación de color RGB.

Una vez que el clasificador especifica tal comportamiento de mapeo de tonos, los codificadores tienen información suficiente para aplicar una transformación de rango dinámico de brillo en cualquier color posible en M_HDR, produciendo un aspecto LDR original (sin comprimir, posiblemente aún sin cuantificar en una representación flotante) LDR_o. A partir de esto, cualquier transformación matemática exacta o aproximada, se puede determinar por el

codificador que permite que un receptor haga la predicción al revés, desde LDR_o hasta Rec_HDR. El clasificador puede verificar a través de una salida 111 de imagen cómo dicha imagen (después de formatearse suficientemente en una señal de imagen que se puede comunicar a través de un enlace de comunicación de imagen, como por ejemplo HDMI) buscaría en una pantalla LDR de referencia (digamos 100 nit, o en el futuro tal vez 500 nit) 103.

Sin embargo, enseñaremos en la presente invención que es útil cuando el mapeo de tonos no solo se construye de manera genérica, sino de una manera particular, y los (pocos) parámetros correspondientes se codifican útilmente como metadatos separados en la señal de imagen S_im, porque se pueden usar ventajosamente en el lado receptor, por ejemplo, durante la capacidad de ajuste para obtener una imagen de conducción óptima para una pantalla X nit particular.

Como primer parámetro, el clasificador elegirá, por ejemplo, un parámetro de sensibilidad SENS, o RHO directamente. Este será un valor que es intuitivamente similar a los valores ASA o ISO conocidos de la fotografía, y generalmente determina qué tan brillante aparecerá la imagen LDR (entre otras cosas, cuánto se elevan los colores de objetos oscuros de M_HDR).

Como una realización preferida, el codificador puede usar una función EOTF/OETF que ya proporciona un buen aspecto inicial de LDR,

cuya función EOTF se define como sigue:

$$L=Lm*potencia(((potencia(RHO,v)-1)/(RHO-1)),GAM)$$

Esta ecuación define las luminancias HDR L a ser representadas que corresponden a los códigos v luma en [0.1] distribuidos equidistantemente en función de la cantidad de bits disponibles para la palabra de código luma de los colores de píxeles, como por ejemplo 1024 valores posibles. Lm es una variable seleccionable que indica el brillo máximo de la pantalla de referencia de la representación lineal de color/luminancia M_HDR o Rec-HDR, que puede, por ejemplo, ser fijado como 5000. Por ejemplo, el clasificador tendrá diales para elegir la sensibilidad que generalmente puede estar relacionada con RHO como:

$$RHO=potencia((SENS/8*SQRT(2)-1) ,2$$

Junto con el valor SENS (RHO) que determina el comportamiento de los colores oscuros y un aspecto de brillo general, el clasificador puede cosintonizar gamma (GAM) como algún parámetro de flexión que reasigna los brillos de objeto/región a lo largo del rango de posibles lumas LDR. Por supuesto, al mapear desde luminancias L en una representación espacial de referencia XYZ de la clasificación M_HDR (que puede ser una representación intermedia útil), a los valores v luma del aspecto LDR, el clasificador definirá la función inversa.

Al hacer cálculos matemáticos elementales en la división RHO, se puede ver que la función inversa (OETF) es: primero aplicar un rendimiento $1/(GAM)$ $xg=potencia(L/Lm,1/GAM)$, y luego calcular: $v=log(1+(RHO-1)*xg)/log(RHO)$.

Típicamente en el codificador puede haber una de diversas formas de realización posibles de una unidad 177 de análisis de imagen. Esta unidad puede disponerse con inteligencia artificial para analizar regiones en la imagen, y cuál de estas regiones podría generar problemas particulares en la codificación HDR, en particular del tipo de modo ii. En particular, puede identificar regiones que podrían ser propensas a la formación de bandas, y regiones que están suficientemente texturizadas, de modo que puedan codificarse con una menor cantidad de códigos de componentes de color y/o luma. En algunas aplicaciones, esta unidad puede llegar automáticamente a una propuesta de codificación final (por ejemplo, Un transcodificador) sin ninguna participación del clasificador humano, pero en otras aplicaciones puede por ejemplo llevar regiones bajo la atención del clasificador para que pueda examinarlas. Por supuesto, puede haber una interacción con la interfaz de usuario, por ejemplo, el clasificador podría indicar que quiere mitigar las bandas con una región particular, o con una textura particular, y luego la unidad 177 puede extraer dicha región y su rango de luma, etc.

Como podemos ver en la figura 2, aunque uno puede decidir codificar una función de mapeo de tonos final en típicamente un espacio LUT reservado en los metadatos 205, típicamente uno codificará un parámetro de sensibilidad (por ejemplo 200 ISO) o un valor RHO y un valor gamma (por ejemplo 2.8) en el campo 202 de metadatos de sensibilidad y el campo 203 de metadatos gamma respectivamente. La figura 2 muestra esquemáticamente cómo se ve una señal de imagen o video S_im (200), y la persona experta sabrá, por supuesto, que se puede definir en la práctica en las muchas variantes digitales, dados los contenedores de datos de imágenes existentes, etc. Nuestras realizaciones de codificador usan una codificación clásica de 3 componentes de la imagen (201) en color de píxeles, que será nuestra imagen de aspecto LDR optimizada por el clasificador. Esta imagen LDR LDR_o generalmente tendrá una codificación DCT clásica, codificación de longitud de ejecución, formateada, etc. de acuerdo con un formato estandarizado de codificación de imagen como JPEG, o un formato de codificación de video estandarizado como MPEG-HEVC, VP1, etc. El lector experto comprenderá que reformatear colorimétricamente para poder reutilizar tecnologías de codificación heredadas (o futuras similares) como un concepto genérico es parte de nuestra invención, pero no es tan importante cuál de esas codificaciones se utiliza realmente. Y otra parte de nuestra invención son los

metadatos necesarios para dar sentido a los datos, por ejemplo, al menos cuando se recupera el aspecto Rec_HDR de la escena (porque el aspecto LDR en teoría se puede usar directamente para conducir una pantalla LDR, sin más procesamiento de rango dinámico sino solo mapeo de redefinición colorimétrica de Y'uv a alguna codificación de espacio RGB dependiente del dispositivo).

5 Además, el clasificador puede usar un valor GAIN (cocodificado en un campo 204 de metadatos de ganancia) para que las funciones no necesiten per se el mapa 1.0 a 1.0. Por ejemplo, la ganancia puede indicar cómo una imagen LDR que se define en el rango completo [0.1] debe ser mapeada para decir solo un subrango [0.1500] del rango [0.5000] de la pantalla HDR. En principio, también es posible, por el contrario, limitar, el rango LDR utilizado, aunque es menos probable que se utilice. Esta ganancia se puede utilizar para hacer que algunas imágenes no sean demasiado brillantes, como se puede imaginar si la escena es, por ejemplo, una escena brumosa o una imagen oscura que está razonablemente iluminada en LDR, pero necesita permanecer oscura en HDR.

15 Estos tres parámetros (RHO, GAM, GAI) ya brindan un primer mapeo muy útil de una imagen M_HDR a una imagen de aspecto LDR correspondiente, con un ajuste de brillo o iluminación más o menos global. Esto puede, por ejemplo, ser suficiente para transmitir espectáculos de la vida real, donde los parámetros óptimos se determinan justo antes del inicio de la transmisión. Los usuarios más críticos, como los productores de películas, pueden querer un control más preciso sobre el aspecto. Es posible que quieran especificar una función de mapeo de tonos más general que la anterior "loggamma", con flexiones finamente posicionadas en la curva que pueden elevar, por ejemplo, el brillo o contraste local promedio de un objeto en particular (por ejemplo, una cara) a un subrango deseado de todas las luminancias LDR representables (o más precisamente sus lumas correspondientes). O una especificación de una pendiente local puede especificar el contraste deseado en algún subrango BL interesante de una región importante en la imagen, a costa de las posiciones de brillo y los contrastes de otras regiones/objetos en la imagen de aspecto LDR.

25 Ahora, una cosa importante a entender es que con nuestro sistema de modo i (de aspecto HDR) el clasificador puede definir tales mapeos arbitrarios, porque solo necesitamos derivar una imagen de aspecto LDR (que no es una reconstrucción, pero se pueden hacer datos destructivamente si así lo desea el clasificador), porque en ese enfoque de codificación tenemos la imagen de aspecto HDR ya codificada como única imagen en la señal de imagen S-im. Sin embargo, en los sistemas en modo ii necesitamos cumplir un criterio doble: por un lado, debemos poder reconstruir la imagen Rec_HDR con buena calidad, pero, por otro lado, queremos la libertad suficiente para crear la mayoría, si no todos, los aspectos LDR que un clasificador puede desear (y luego puede ser bastante creativo a veces, como se puede ver, por ejemplo, en la película Sin City 2).

35 Pero uno debe entender que cualquiera que sea la clasificación LDR_o que haya hecho el clasificador con su mapeo 210 de tonos preferido, en una codificación heredada, estas lumas LDR de salida pasarán por la cuantificación uniforme clásica (e incluso DCT-eo). Por lo tanto, debemos tener cuidado de no crear mapeos que sean demasiado planos en algunas partes de su rango (es decir, la derivada local $\Delta_{LDR_out}/\Delta_{HDR_in}$ no debe ser demasiado pequeña, de modo que se asigne una cantidad mínima requerida de códigos luma LDR a ese rango Δ_{HDR_in} o el correspondiente Δ_{LDR_out}), porque de lo contrario al aumentar ese rango en el mapeo de tonos LDR-2-HDR, veremos artefactos como bandas o artefactos DCT excesivamente contrastantes y visibles.

45 Podríamos tener un mecanismo de control con una rigidez de los puntos de control locales que el usuario usa para cambiar la forma del mapeo de tonos arbitrario, pero eso es desagradable para el usuario, especialmente si se implementa con dureza (por supuesto, el sistema puede advertir si el clasificador desea hacer curvas de mapeo realmente extrañas, por ejemplo, no se deben hacer inversiones como una curva N).

50 Una realización útil se muestra en la figura 3, que aclara el comportamiento de una unidad 106 técnica de mapeo de tonos, que puede usarse para determinar un segundo aspecto LDR, alternativamente utilizable a LDR_o por receptores más inteligentes que necesitan dar servicio a una pantalla LDR. Asumimos que el clasificador ha elegido su curva deseada que le da el aspecto LDR apropiado, que es la curva sólida en la figura 3. Si la curva de mapeo de tonos no es buena, esto significará que hay al menos un rango que es demasiado plano, que asumimos aquí es la parte R-u del HDR más brillante y los píxeles LDR, digamos el cielo de la escena. Necesitamos ser capaces de estirar ese rango L_u en LDR, de modo que se puedan asignar un poco más de códigos luma y de una manera tan no destructiva como sea posible (cambiando poco su aspecto) para el clasificador.

Esto se puede hacer cuando hay un rango adyacente L_uu que contiene más objetos texturizados.

60 Esta es una forma de salir del enigma de que nuestra curva de aspecto para obtener un aspecto LDR deseado al mismo tiempo determina la cuantificación o el número de códigos luma disponibles para codificar fielmente las diversas texturas de la región HDR (la caracterización fiel suficiente de todas las texturas que se encuentran en la escena es el objetivo principal de la calidad de codificación en la codificación HDR). Tener 1024 niveles diferentes de luma/gris (y millones de códigos) debería ser suficiente para codificar bien todas las texturas para la visión humana, si está bien hecho. Los objetos complejos se pueden codificar con relativamente menos códigos, ya que el ojo primero ve el patrón de textura gruesa, y luego no tanto los valores precisos de los colores de los píxeles. Solo en situaciones desfavorables particulares podemos tener un problema si tenemos gradientes de brillo para los que hemos usado muy pocos códigos.

Por lo tanto, hay dos cosas cuando se adapta una curva: la unidad 106 de mapeo de tonos técnicos generalmente mantiene la adaptación cuando se necesita lo suficientemente local en el eje luma, para que no perturben las lumas de demasiados colores de objeto (por ejemplo, evite oscurecer regiones oscuras críticas demasiado de nuevo). Un criterio de calidad para esta escena de ejemplo puede ser que necesitemos iluminar los colores oscuros para obtener un buen aspecto de LDR, por lo que un cambio local en los colores brillantes no perturbará eso de ninguna manera. Por lo tanto, la unidad 106 de mapeo de tonos típicamente redistribuirá los códigos en algún subrango local de luma alrededor del área problemática, y determinará una curva de adaptación correspondiente para esto, que es la línea de puntos (esta curva puede seguir algo la forma de la curva original, en sus dos partes de codificación de la región de la imagen, es decir, si había una forma local de flexión parabólica para las lumas del cielo, normalmente puede usar un segmento parabólico de flexión similar más grande y escalado para el aire, pero eso no es absolutamente necesario, ya que solo la precisión de la codificación es el criterio).

Por lo tanto, necesitamos estirar un poco el rango de brillo de la región del cielo, para tener suficientes códigos para codificar fielmente un gradiente de cielo azul Rec_HDR. Pero ¿cuánto necesitamos hacer eso y hasta qué punto deberíamos extender el rango de ajuste R_Adj?

Eso depende de varias cosas. Por supuesto, R_adj debería cubrir la región donde hay un problema, que típicamente será una región relativamente simple visualmente, como regiones relativamente uniformes, como un gradiente en el cielo (este gradiente azul existirá en algún lugar a lo largo del rango de luma LDR). Por otro lado, debemos necesitar una región adyacente que este suficientemente texturizada. En la situación poco probable de que la región adyacente sea otro gradiente más suave (que podría ocurrir en imágenes sintéticas como imágenes de prueba de gradiente artificial, en ese caso tendremos que estar satisfechos con cualquier asignación óptima de luma que podamos obtener, pero esto no ocurre típicamente en imágenes naturales), R_adj puede llegar a ser relativamente grande. En la situación normal donde pronto nos encontramos con un rango texturizado, podemos extender L_u con un rango L_uu de un tamaño que depende de cuántos códigos tengamos que agregar y la complejidad del patrón de textura. Si necesitamos agregar solo 3 códigos al cielo, necesitamos guardar 3 códigos de luma en L_uu, y si tenemos suficiente textura podríamos hacerlo en un rango de ay 10-15 lumas, dependiendo de lo que encuentre/pueda el clasificador u observador encontrar aceptable.

El aparato puede contener tablas para eso.

Entonces, el desagradable problema con la codificación de luma dependiente de la curva de aspecto ahora está resuelto en gran medida. Por un lado, no oscurecemos demasiado los objetos oscuros adyacentes, ya que solo cambiamos un poco los colores de L_uu en el rango superior al expandir nuestro rango de cielo L_u, pero principalmente mantenemos la parte inferior de L_uu igual, solo muestreo un poco menos, lo cual no es un problema visualmente visible de todos modos, porque las texturas no necesitan tantos códigos de todos modos. El rango extendido del cielo puede ser un poco subóptimo, pero normalmente no debería ser un problema, y a cambio obtenemos una calidad mejorada Rec_HDR. Pero todo esto sigue siendo solo si no tomamos ninguna medida en el extremo receptor, por ejemplo, por un receptor que no puede hacer ningún procesamiento. Porque en el decodificador podemos hacer una estrategia de precompensación en la unidad 159 de remapeo de tonos. Esto hará que la asignación de luma sea una cuestión puramente técnica fuera de las preocupaciones de los intentos artísticos del clasificador. Debido a que la unidad 159 de remapeo de tono aplicará nuevamente la corrección para el estiramiento local en una compresión, antes de usar el aspecto LDR deseado resultante (LDR_ul), por ejemplo, manejando una pantalla LDR. Entonces, en el ejemplo del cielo, donde estiramos el límite inferior del cielo de L_u hacia abajo hacia los brillos de objetos en el rango adyacente L_uu (oscureciendo así esos objetos), la unidad 159 de remapeo de tono de un decodificador 150 aplicará el mapeo inverso de 301 como una corrección. Esto significa que visualmente el rango del cielo tendrá su rango de luma original L_u nuevamente, y cuando se visualiza en un LDR muestra el rango de luminancia correcto, sin embargo, tiene más precisión porque se le asignaron más códigos de luma que codifican la textura. De manera similar, en el aspecto LDR_ul, el objeto con brillo adyacente en L_uu también tendrá el brillo correcto no atenuado, y solo diferirá en precisión debido a la cantidad reducida de códigos. Y la persona experta puede comprender cómo esta técnica puede siempre, en diversas otras posibles situaciones, mejorar la precisión de codificación en aquellas regiones de una imagen donde sea necesario, mientras mantiene el aspecto LDR_ul deseado del clasificador. Lo único que debe poder hacer la unidad 159 de remapeo de tonos es aplicar una estrategia de mapeo de tonos al LDR_t técnico decodificado, por ejemplo, mediante una LUT, que puede ser cocodificada en la señal S_im (o codificarse parcialmente si el mapeo de tonos puede derivarse, por ejemplo, de un conjunto limitado de puntos de control, por ejemplo, delimitación de segmentos lineales) y, por lo tanto, debe quedar claro por qué es ventajoso codificar esta función de ajuste técnico por separado (Ff1, Ff2,...) en S_im, porque puede ser utilizado por el decodificador incluso para obtener un aspecto LDR más deseable LDR_ul, una vez que se ha determinado en el lado de la creación y aceptado por el clasificador y comunicado a un lado receptor.

Habrá en gran medida dos categorías de realizaciones de codificador que permitirán lo anterior. El primero en gran medida realiza todo el procesamiento automáticamente y no necesita involucrar al usuario. Los detectores de suavidad y textura categorizarán automáticamente las diversas regiones y, por lo tanto, identificarán el patrón de gradiente en el cielo y los objetos con textura ubicados adyacentes (es decir, en el rango de luma ubicado debajo y/o arriba de L_u). Se pueden incorporar diversos caracterizadores de textura para determinar la complejidad de la textura (por ejemplo,

Grano fino, cantidad de valores de grises entrelazados, etc.) y determinar a partir de eso cuán visibles serán las perturbaciones visualmente conspicuas y de esta el rango L_{uu} resultante necesario. Como se dijo, estas preferencias pueden estar preconstruidas en fórmulas que determinan el L_{uu} funcionalmente, o con LUT. También en algunas realizaciones, DCT u otros emuladores de compresión pueden estar presentes, por ejemplo, que calculan las imágenes LDR descomprimidas resultantes LDR_d bajo varias opciones para R_{adj} y la forma 301 de perturbación de mapeo de tonos funcional, y calculan una medida de gravedad para la visibilidad típica (en el rango de visualización normal, tamaño de pantalla, brillo envolvente, etc.) de la banda y/u otros artefactos de compresión. La unidad 117 de análisis de textura puede estar presente para esto, la cual generalmente está dispuesta para analizar texturas, y en particular su impacto visual, tanto en el LDR_c original como en el LDR_c codificado, o de hecho la decodificación del mismo LDR_d que finalmente estará presente en el extremo receptor. En particular, los remapeos a HDR por la unidad 118 de mapeo de color LDR-2-HDR se pueden usar para permitir que el clasificador verifique el impacto visual si es necesario. Si el clasificador desea verificar la reconstructibilidad de este M_HDR como Rec_HDR , puede, por ejemplo, alternarlos a tiempo en su pantalla 102 HDR, a través de la salida 119 de imagen HDR. De hecho, el decodificador puede tener varias salidas (que hemos mostrado por separado, pero, por supuesto, pueden enrutarse internamente a una sola salida) 111, 112, 113, 114 para poder verificar las diversas versiones de LDR.

Una segunda categoría de codificadores con reclasificación técnica puede involucrar directamente al clasificador humano. Si ya está verificando la calidad de los algoritmos automáticos, puede tener una opción para influir en los resultados (es decir, generalmente de forma semiautomática). Esto debería ser simple para el clasificador, ya que puede querer involucrarse más con la determinación artística del aspecto, es decir, la colocación de las lumas del objeto, en lugar de problemas técnicos como los artefactos de compresión (si ya quiere ver eso, y aunque él verificará uno o más escenarios típicos y aprobados, en la línea de comunicación de la imagen, por supuesto, puede haber más compresiones que podrían tener artefactos más graves).

En estas realizaciones de codificador, la unidad 105 de interfaz de usuario permitirá típicamente al clasificador especificar áreas de imagen geométrica que de acuerdo con él son áreas particularmente problemáticas. Por ejemplo, puede garabatear a través del cielo, y las unidades de análisis de histograma y análisis de textura se centrarán en esta parte de la imagen al hacer su análisis y la actualización técnica de la determinación de la curva de mapeo de tono parcial. Por ejemplo, pueden proponer sucesivamente una estrategia que agregue algunos más códigos de luma a la vez al cielo, hasta que el clasificador esté satisfecho. Por ejemplo, un algoritmo de realización de la unidad 106 de mapeo de tono puede multiplicar este rango del objeto de gradiente (sensible a las bandas) por $k =$ por ejemplo. 1.5, y seleccionar un rango vecino de una región de imagen texturizada y comprimir ese a $L_{uu} - 1.5 * L_u$. Es decir, se puede usar cualquier redistribución lineal o curvilínea de los códigos en las dos regiones. El L_{uu} puede seleccionarse para ser al menos por ejemplo $3 * L_u$, cuyos valores son típicamente optimizados por un diseñador de aparatos sobre la base de un conjunto de imágenes representativas. Si la propuesta del aparato es buena, el clasificador la acepta, haciendo que el codificador almacene los parámetros correspondientes en S_{im} , o de lo contrario se inicia una nueva iteración, por ejemplo, con $k = 1.1 * 1.5$.

La perturbación 301 conducirá a un mapeo de tonos final, con el que corresponde una clasificación técnica final LDR_i , que será el aspecto LDR que se envía al sistema de comunicación después de un formateo adicional de acuerdo con nuestro sistema de codificación HDR en modo ii, y que corresponde en gran medida a lo que el clasificador desea como aspecto LDR. La ventaja de la participación del clasificador es que puede indicar, al menos con un mínimo de participación, qué regiones son semánticamente más relevantes. El analizador estadístico de textura puede determinar que realmente existen pocas lumas (es decir, pocos píxeles) en una región entre, por ejemplo, las lumas oscuras de una habitación en el interior y las lumas brillantes de los soleados exteriores y, por lo tanto, deciden aplicar una estrategia de remapeo que aplique pocos códigos allí (en caso de que el remapeador 159 del decodificador pueda reconstruir arbitrariamente el aspecto LDR deseado, incluso podríamos usar una curva de deformación técnica fuerte que casi corta todo el subrango escasamente utilizado fuera de la codificación LDR_i , lo que hace inmediatamente adyacente el valor luma LDR_i los subrangos interiores y exteriores). Sin embargo, si en esta pequeña región hay un objeto importante como la cara de alguien o un objeto que se destacó de alguna manera como un objeto que aparece, el clasificador puede contrarrestar esto. Son posibles varias realizaciones prácticas, por ejemplo, él puede garabatear en nuestro dibujo un rectángulo alrededor de esta región, y luego girar un dial que aumenta la cantidad de códigos luma que se utilizarán para esa región. El lector experto comprenderá que hay varias otras formas de interfaz de usuario para seleccionar una región u objeto crítico en la imagen o toma, e indicar cómo debe codificarse con lumas, incluso hasta que el clasificador dibuje o influya en la forma de la curva de modificación 301 en sí mismo.

El resto de nuestro sistema mode-ii es el siguiente:

Opcionalmente, la unidad de conversión de rango dinámico puede realizar un procesamiento de saturación de color (por ejemplo, dado que el color disminuye con el oscurecimiento y viceversa, el clasificador puede querer compensar la saturación que se ha vuelto algo inapropiada debido al mapeo de tono luma). Una buena realización ejemplar práctica funciona con una función de saturación general del tipo no destructivo de información. Con esto queremos decir que esta función de saturación tampoco es demasiado plana, por lo que también se puede revertir. Pero en algunas realizaciones, la función de saturación solo puede aplicarse en la actualización LDR-2-HDR, y luego puede ser más liberal. En la Figura 3 hemos mostrado una saturación suave de s_{in} a s_{out} , que puede codificarse con varios valores $S1, S2, S3$ en un LUT en la señal S_{im} . Estos pueden ser los valores s_{out} para valores equidistantes

s_{in} (una cantidad suficiente para que la curva deseada pueda recuperarse razonablemente sin problemas en el decodificador), pero esto también podría ser, por ejemplo, puntos de control de forma de función. Una función de desaturación puede, por ejemplo, ser codificada como una línea con pendiente menor a 45 grados (en un diagrama de s_{in} vs. s_{out}). En tal caso de desaturación, la señal de imagen podría tener un valor entero o flotante para el multiplicador en los metadatos. Asumimos en el ejemplo que aclara que s_{out} será la saturación de la imagen HDR, y necesitamos aumentar la saturación de los colores oscuros ahora más oscuros de la escena para aumentar el colorido, pero la persona experta puede entender que puede haber diferentes variantes de procesamiento en la misma filosofía de codificación estructural. Por simplicidad de aclaración, supondremos que la saturación se realiza en el espacio uv, por ejemplo, independientemente de la luma, podemos realizar la operación $s_{out} = s_{in} + MS(s_{in}) * s_{in}$. La MS (s_{in}) es entonces el valor multiplicativo recuperable de la función como se ve en la figura 2b y codificado en LUT 206, que estira el vector de saturación en una dirección de tono en comparación con algún punto blanco. Suponemos por simplicidad que hemos definido nuestro espacio uv en uno cilíndrico con la máxima saturación en la periferia (y codificado como 1.0). Por supuesto, la persona experta comprenderá que ambos podemos codificar nuestra estrategia de saturación en otra definición colorimétrica, o dado que la definición es, por ejemplo, en el espacio cilíndrico $Y'uv$, el diseñador del hardware o software del decodificador puede elegir realizarlo de manera equivalente en otro espacio de color, como el espacio YCrCb basado en RGB, etc. El clasificador también puede determinar y codificar en S_{im} , estrategias de saturación dependientes de luma, es decir, funciones que cambian la saturación, cuyo multiplicador varía con la luminancia del color procesado. Básicamente, una realización más avanzada de S_{im} tendrá una estructura de codificación de saturación. Esto puede ser, por ejemplo, una definición basada en la web que tiene para varios tonos clave (por ejemplo, 6: RGBCYM) una función multiplicadora definida sobre luma: MS (Y'). A partir de esto, que puede codificarse como 6 LUT de valores similares a 206, en el extremo receptor el decodificador puede determinar una estrategia de saturación para todos los colores en la gama por interpolación. Una estrategia más compleja puede incluso introducir variabilidad de la saturación en la dirección radial. Esto puede codificarse fácilmente determinando estas funciones (similar a lo que se ve en la figura 2b, pero ahora variable sobre la altura de la luma en la gama) simplemente paramétricamente, por ejemplo, como funciones de desface, gamma, ganancia. En este caso, uno tendría: $s_{out} = s_{in} + F(s_{in}, Y')$ para los tonos clave, y en el caso de, por ejemplo, un control de forma de función de tres parámetros, uno puede codificar esto en S_{im} como 3x6 LUT especificando el comportamiento de luma de, por ejemplo, el parámetro `saturation_gamma` como variando de acuerdo con Y' o 6 LUT para los tonos, pero donde no se codifica un solo valor multiplicativo en cada posición, sino un triplete [`sat_offset` (Y'_i), `sat_gain` (Y'_i), `sat_gamma` (Y'_i)] _ LUT_of _yellow, consecutivamente en varias posiciones, muestreando las posibles lumas en la gama.

Ahora, en algunas realizaciones de un codificador (y el decodificador correspondiente) hay una transformación opcional a u'v' para las características de color de los píxeles, que ahora aclararemos (pero otras realizaciones pueden codificar alternativa o adicionalmente en, por ejemplo, R'G'B' o YCrCb, etc. directamente, y ni siquiera tienen la unidad 107 opcional adentro; tenga en cuenta también que algunos procesamientos $Yu'v'$ pueden reescribirse matemáticamente como procesamiento RGB lineal equivalente).

Habiendo aplicado la transformación de rango dinámico para crear el aspecto LDR correcto (por ejemplo, en el espacio RGB, o XYZ, etc.), suponiendo que no hayamos hecho el mapeo en el espacio $Y'uv$, la unidad 107 de transformación de color de la realización de aclaración ejemplar hará la conversión a nuestra representación u'v', con las lumas Y' en esa representación de color que se determina por nuestra función de mapeo de tonos totales (es decir, las lumas de la imagen LDR intermedia LDR_i), y u, v de acuerdo con las ecuaciones anteriores. También podríamos hacer transformaciones colorimétricas en la unidad 107, que condicionan los colores cuando se prevé un RGB dependiente de un dispositivo diferente o un espacio multiprimario. Por ejemplo, Si nuestro M_HDR fue codificado con un triángulo RGB más pequeño, pero el LDR es para una pantalla de gama amplia, el clasificador ya puede predefinir una estrategia de aumento de saturación, aunque las cosas a menudo serán al revés, en cuyo caso la unidad 107 puede implementar un mapeo de gama cromático.

Finalmente, el LDR_{uv} resultante se codifica con una imagen LDR clásica o un compresor 108 de video, es decir, típicamente DCT o wavelet transformado, etc.

Esta imagen comprimida LDR_c se envía a un formateador 116, que agrega los metadatos en la función de mapeo aplicada de acuerdo con un formato estandarizado, para que esté adecuadamente disponible en el lado receptor. Es decir, este formateador agrega el valor de sensibilidad (RHO o, alternativamente, SENS), el mapeo de tonos adicional para ajustar el aspecto LDR de acuerdo con lo determinado típicamente por el clasificador humano (aunque en el futuro además algunos codificadores pueden ser lo suficientemente inteligentes como para hacer algunos ajustes por sí mismos) con función que define los parámetros 205 típicamente como una LUT de valores ($F1, F2, \dots$), la saturación que codifica 206, por ejemplo también un conjunto de parámetros que definen una función multilínea, etc.

El mapeo de tono adicional por razones técnicas generalmente se almacena por separado en la imagen o señal de video S_{im} , preferiblemente como un conjunto de valores 207 enteros o reales, que pueden usarse para almacenar, por ejemplo, un LUT de 256 puntos o 1024 puntos.

El LDR_c codificado se puede decodificar nuevamente a LDR_d , y luego se puede actualizar mediante la unidad 118 de mapeo de color para que el clasificador pueda ver a través de la salida de imagen 119 cómo se vería el HDR

Rec_HDR reconstruido en un extremo receptor. Si lo desea, incluso podría probar la influencia de algunas configuraciones de compresión típicas, por ejemplo, fuerte compresión. El decodificador descrito en este documento también podría usarse en una estrategia de recodificación, donde el aspecto de clasificación ya puede haberse preparado previamente, pero ahora, por ejemplo, una versión LDR altamente comprimida de baja calidad se redetermina para alguna solicitud particular de comunicación de imagen/video. Ese clasificador secundario puede incluso reajustar los parámetros. Dependiendo de si tiene el M_HDR original disponible, puede, por ejemplo, redeterminar las funciones de subclasificación para lograr un nuevo aspecto LDR más ajustado de manera apropiada (por ejemplo, servir a los observadores de teléfonos móviles), y de hecho incluso puede hacerlo cuando solo tiene disponible el buen Rec_HDR en lugar de M_HDR. La división de una parte de clasificación técnica para asignar más adecuadamente los códigos luma es muy útil para tales escenarios. Debido a que las funciones que mapean a LDR_o (y la correspondiente reconstrucción cercana LDR_ul del mismo) determinan el aspecto artístico real de LDR, y pueden haber sido determinadas de una vez por todas por el clasificador primario en el momento de la producción inicial del contenido. Pero el codificador aún puede, de forma automática o semiautomática, con la participación del clasificador secundario determinar el mapeo técnico con las pequeñas modificaciones como 301, y el correspondiente LDR_i (o LDR_t), y los metadatos codificados Ff1, Ff2, en conjunto de valores 207 reales o enteros en S_im, que, por supuesto pueden ser diferentes para diferentes limitaciones tecnológicas, como la cantidad de bits (por ejemplo, solo 8 bits para el canal luma).

El decodificador 150 puede ser un IC en, por ejemplo, como en esta aclaración, un decodificador u ordenador conectable a una pantalla 160 o televisión (por lo tanto, cuando decimos decodificador, pretendemos cubrir cualquier pequeña realización de esto, como un "decodificador en una memoria USB" o cualquier aparato grande que se dé cuenta y se beneficie de nuestra invención, como un decodificador con disco duro y funciones de lectura de disco óptico, y el codificador puede ser cualquier cosa, desde un dispositivo pequeño hasta un sistema de clasificación grande, etc.), pero, por supuesto, la televisión puede no ser un monitor tonto, pero comprende todo esta tecnología de decodificación en su propio IC. La pantalla 160 puede ser tanto una pantalla LDR como una pantalla HDR, o básicamente cualquier pantalla conectada a través de cualquier tecnología de comunicación de imagen a través de la salida 157 de imagen, como por ejemplo transmisión inalámbrica a un dispositivo multimedia portátil o un proyector de cine profesional.

El decodificador obtiene nuestro S_im formateado a través de la entrada 158 de imagen, y un deformador 151 lo dividirá en una imagen LDR_c (IMG en la figura 2) para descomprimirlo mediante un descompresor 152 clásico de tipo JPEG o MPEG, y los parámetros P de metadatos (por ejemplo, una configuración 1000 de sensibilidad y algunos valores que pueden usarse para reconstruir un mapeo de tonos o una forma funcional de mapeo de saturación). Opcionalmente, en el decodificador se encuentra la unidad 159 de remapeo de tono, ya que este remapeo técnico generalmente no es una deformación severa del aspecto LDR_ul destinado al clasificador, algunos decodificadores pueden permitirse ignorarlo. Sin embargo, los decodificadores totalmente compatibles con HDR deben usar esta unidad 159 para aplicar una estrategia de corrección técnica codificada en los valores Ff de 207, para llegar al aspecto LDR correcto LDR_ul (que es una aproximación cercana de LDR_o). Esta imagen LDR corregida (LDR_ul) va a otra unidad 154 de ajuste de color de pantalla. Esta unidad 154 puede aplicar la optimización necesaria para una pantalla particular de gama amplia por decir 1300 nit (por ejemplo, capacidad de ajuste). Aunque las variantes son posibles, hemos dibujado un decodificador típico para nuestra filosofía de codificación HDR, que tiene una ruta de procesamiento de imagen para recuperar LDR_ul (o si 159 no está presente su aproximación LDR_t), pero también tiene una segunda ruta de procesamiento de imagen para determinar Rec_HDR. Esto se realiza en la unidad 153 de conversión de rango dinámico, que normalmente aplica los mapeos inversos aplicados en el codificador (en realidad, en la señal uno codificará típicamente los parámetros de este mapeo inverso, es decir, una actualización). La unidad de sintonización del color de la pantalla 154 generalmente se organizará para combinar la información en las dos graduaciones, lo que podría hacerse basándose en el uso de una sola imagen y los parámetros P de mapeo del color, pero asumimos en esta realización aclarada que esta obtiene una imagen Rec_HDR y LDR_ul como entrada y luego las interpola, de acuerdo con la pantalla con la que se conecta el brillo máximo y que se suministrará con las imágenes graduadas adecuadamente.

Además del mapeo de tonos para obtener el aspecto de brillo correcto, una unidad 155 de transformación de color puede estar típicamente dispuesta para hacer adaptaciones cromáticas para optimizar una gama de colores diferente a la gama de codificación (por ejemplo Rec. 2020 a DCI-P3 o Rec. 709, etc.).

Lo que se emitirá a través de la salida de imagen 157 y, por lo tanto, calculado por la unidad 154 dependerá, por supuesto, de la pantalla conectada. Si se trata de una pantalla LDR, la unidad 154 puede enviar, por ejemplo, LDR_ul, después de, por supuesto, el correcto remapeo de color (por la unidad 155) de Y'uv a una codificación R'G'B' dependiente del dispositivo particular, por ejemplo, si la pantalla 160 conectada está cerca de una pantalla de brillo máximo de 5000 nit (vea también cómo el aparato de decodificación puede preguntarle a un t.v. sus capacidades en el documento WO 2013/046096; un controlador 161 puede hacer tal comunicación con la pantalla e incluso con el observador para obtener sus preferencias, y puede organizarse para configurar cómo debe comportarse la unidad de sintonización de pantalla 154 y qué tipo de imagen debería calcular y emitir) la imagen de aspecto Rec_HDR puede emitirse, nuevamente después de un formateo adecuado de acuerdo con lo que el televisor desea recibir (es decir, esto todavía puede ser una codificación Y'uv, por ejemplo, nuestro formato S_im con ahora una imagen de aspecto HDR almacenada en 201/IMG, y también se pueden transmitir algunos metadatos funcionales para que el televisor

pueda hacer un ajuste fino colorimétrico basado en la información sobre cómo cambian las clasificaciones sobre el espectro de posibilidades de representación codificadas en estos metadatos, o ya puede ser una imagen de conducción de pantalla R'G'B' HDR). Para pantallas de brillo pico intermedio, la unidad 154 puede generar una imagen de manejo adecuado, nuevamente en nuestro formato Y'uv u otro formato.

5 Finalmente, el creador de contenido puede prescribir en la señal si desea que no se omita el mapeo de compensación de la unidad 159, por ejemplo, porque el creador de contenido piensa que LDR_t se desvía seriamente de LDR_ul. Esto se puede hacer codificando un Booleano 209 en un campo IGNORE_TECHNICAL_MAPPING de los metadatos.

10 Debe quedar claro para el lector que cuando hemos aclarado solo el mínimo de un conjunto de parámetros, por supuesto, a lo largo del mismo razonamiento, se pueden codificar varios conjuntos de metadatos funcionales de mapeo de colores en S_im, por ejemplo, un conjunto para pasar de la única imagen IMG (que es una imagen LDR) a una referencia, por ejemplo, imagen de aspecto nit de [0-5000] HDR, y se puede agregar un segundo conjunto para ir a, por ejemplo, un aspecto MDR de 1500 nit. Y aunque hacer una descomposición específica de una forma de sensibilidad, gamma, ganancia y función de ajuste fino adicional es ventajoso, y al menos bueno para la aclaración técnica, cualquiera de los mapeos, por ejemplo, el mapeo LDR-2-MDR podría estar codificado en S_im en forma condensada, por ejemplo, solo llenando el LUT de mapeo de tono o un conjunto de valores 205, que codifican la función de mapeo final (es decir, todo lo relacionado con sensibilidad, ajuste fino y mapeo técnico juntos).

20 La figura 4 muestra esquemáticamente una realización típica de nuestra unidad 400 central de decodificador (en este ejemplo, la parte mínima del modo ii, sin reclasificación técnica, o conversión Yu'v', etc.). Después de que un descompresor 401 realiza la longitud de ejecución o la decodificación aritmética, y DCT inverso, etc., obtenemos una imagen LDR_t, que asumiremos que está en representación de gamma 2.2. (es decir, con lumas o componentes R'G'B' definidos de acuerdo con la Rec. 709) y normalizados. Puede haber una primera unidad 420 de control, que
 25 puede enviar directamente esta imagen a un TV LDR 410 conectado (lo que significa directamente que puede haber algún formato heredado involucrado; en principio, LDR_t también podría ser, por ejemplo, una imagen lineal, en cuyo caso habrá necesidad de mapear este regamma-2.2 antes de enviarlo a la pantalla LDR, pero puede ser ventajoso si no es necesario; las funciones adicionales de mapeo de tono generalmente serán diferentes dependiendo del tipo de LDR_t, que también puede ser indicado con un indicador IND_2 en S_im). Entonces, una primera unidad 402 de mapeo de tono realiza el mapeo inverso del mapeo de tonos arbitrario, recibiendo los parámetros de definición de esa forma de función P_CC en los metadatos MET(F). Luego, una segunda unidad 403 de mapeo de tono realiza un mapeo de tono reoscureciendo los colores más oscuros relativamente a los más brillantes, por ej. aplicando la ecuación rho anterior, con un valor RHO recibido. La unidad también podría calcular el valor RHO de un brillo de pico de pantalla recibido PB_HDR, recibido de la pantalla 411 HDR conectada. Luego, una tercera unidad 404 de mapeo de tono
 35 realiza una función de potencia gamma con un valor GAM recibido, por ejemplo, 2.4 preferiblemente. Entonces, un multiplicador 405 puede hacer una multiplicación con GAI, que por defecto puede ser 1.0. Opcionalmente, un procesador 406 de saturación de color puede realizar un procesamiento de saturación. Finalmente, la unidad 421 de control puede enviar la imagen a la pantalla HDR 411, y puede realizar algún procesamiento adicional, por ejemplo, para formatear correctamente la imagen de acuerdo con un estándar que la pantalla entiende, por ejemplo, a través de una conexión HDMI.

La figura 6 muestra una realización de unidad de conversión de rango dinámico de codificador simple. Ésta comprende una unidad 601 de normalización para normalizar todos los componentes de color a 1 (es decir, si, por ejemplo, R, G y B están normalizados a 1.0, por lo que la luminancia máxima se normalizará a 1.0 y viceversa). Las luminancias normalizadas Yn_HDR de los píxeles de la imagen HDR (o en realizaciones equivalentes, por ejemplo, los componentes RGB lineales normalizados) van a un primer mapeador 602 de tono que realiza una operación gamma, con una gamma deseada por el clasificador (o unidad de clasificación automática), pero generalmente fija a 1/(2.4). Luego, un segundo mapeador 603 de tono realiza la transformación que ilumina adecuadamente los colores oscuros HDR, por ejemplo, con $v = \log(1 + (RHO - 1) * xg) / \log(RHO)$, con un factor RHO apropiado propuesto por el sistema de clasificación dependiendo de la diferencia de rango dinámico entre (el brillo máximo de) M_HDR y el LDR típicamente de 100 nit, y típicamente finalmente aceptado por el clasificador, quien puede o no cambiar este valor RHO propuesto inicialmente. Luego, al usar el tercer mapeador 604 de tono el clasificador comienza a afinar mirando varios objetos en la imagen y, en última instancia, define una curva de mapeo de tono personalizado CC, cambiando diversas lumas de aquellos diversos de acuerdo con los objetos de imagen importantes del clasificador. Esto produce las lumas
 55 Yn_LDR de la imagen LDR_o, con todos los datos listos para ser codificados.

Los componentes algorítmicos divulgados en este texto pueden (en su totalidad o en parte) realizarse en la práctica como hardware (por ejemplo, partes de un CI específico de la solicitud) o como software que se ejecuta en un procesador de señal digital especial, o un procesador genérico, etc.

60 Debe ser comprensible para la persona experta desde nuestra presentación qué componentes pueden ser mejoras opcionales y se pueden realizar en combinación con otros componentes, y cómo los pasos (opcionales) de los métodos corresponden a los medios respectivos de los aparatos, y viceversa. La palabra "aparato" en esta solicitud se utiliza en su sentido más amplio, es decir, un grupo de medios que permiten la realización de un objetivo particular y, por lo tanto, por ejemplo, ser (una pequeña parte de) un IC, o un aparato dedicado (como un aparato con pantalla), o parte de un sistema en red, etc. "Arreglo" también está destinado a ser utilizado en el sentido más amplio, por lo que puede

comprender, entre otros, un solo aparato, una parte de un aparato, una colección de (partes de) aparatos cooperantes, etc.

5 Debe entenderse que una versión de producto de programa de ordenador de las presentes realizaciones como denotación abarca cualquier realización física de una colección de comandos que permite un procesador genérico o de propósito especial, después de una serie de pasos de carga (que pueden incluir pasos de conversión intermedios, como la traducción a un lenguaje intermedio y un lenguaje de procesador final) para ingresar los comandos en el procesador y ejecutar cualquiera de las funciones características de una invención. En particular, el producto del programa informático puede realizarse como datos en un operador como, por ejemplo, un disco o una cinta, los datos
10 presentes en una memoria, los datos que viajan a través de una conexión de red, por cable o inalámbrica, o el código del programa en papel. Además del código del programa, los datos característicos requeridos para el programa también pueden incorporarse como un producto de programa de ordenador. Debe quedar claro que con ordenador nos referimos a cualquier dispositivo capaz de hacer los cálculos de datos, es decir, también puede ser, por ejemplo, un teléfono móvil. También las reivindicaciones de aparatos pueden cubrir versiones implementadas por ordenador
15 de las realizaciones.

Es posible que algunos de los pasos necesarios para el funcionamiento del método ya estén presentes en la funcionalidad del procesador en lugar de estar descritos en el producto del programa informático, como los pasos de
20 entrada y salida de datos.

Debe observarse que las realizaciones mencionadas anteriormente ilustran en lugar de limitar la invención. Cuando la persona experta puede realizar fácilmente un mapeo de los ejemplos presentados a otras regiones de las reivindicaciones, tenemos por concisión que no mencionamos todas estas opciones en profundidad. Aparte de las combinaciones de elementos de la invención como se combinan en las reivindicaciones, son posibles otras
25 combinaciones de los elementos. Cualquier combinación de elementos se puede realizar en un solo elemento dedicado.

Cualquier signo de referencia entre paréntesis en la reivindicación no pretende limitar la reivindicación. La palabra "que comprende" no excluye la presencia de elementos o aspectos que no figuran en una reivindicación. La palabra
30 "un" o "una" que precede a un elemento no excluye la presencia de una pluralidad de tales elementos.

REIVINDICACIONES

1. Un método para codificar una imagen de alto rango dinámico (M_HDR), que comprende los pasos de:

5 - convertir la imagen de alto rango dinámico en una imagen de menor rango dinámico de luminancia (LDR_o) mediante al aplicar:

a) normalización de la imagen de alto rango dinámico a una escala del eje luma siendo [0.1] produciendo una imagen normalizada de alto rango dinámico con colores normalizados que tienen luminancias normalizadas (Yn_HDR), b) calcular una función gamma en las luminancias normalizadas que producen luminancias convertidas en gamma (xg), c) aplicar un primer mapeo de tonos que produce lumas (v) que se define como $v = \log(1 + (RHO - 1) * xg) / \log(RHO)$, con RHO que tiene un valor predeterminado, y d) aplicar una función de mapeo de tonos arbitrariamente creciente monotónicamente mapeando las lumas para producir lumas (Yn_LDR) de la imagen de rango dinámico inferior (LDR_o); y

15 - emitir en una señal de imagen (S_im) una codificación de los colores de píxeles de la imagen de rango dinámico de luminancia inferior (LDR_o), y

20 - emitir en la señal de imagen (S_im) en valores de metadatos que codifican las formas de la función de mapeo de tono arbitrariamente creciente monotónicamente, cuyos metadatos permiten a un receptor reconstruir una imagen reconstruida de alto rango dinámico (Rec_HDR) a partir de la imagen de rango dinámico de luminancia inferior (LDR_o).

2. Un codificador (100) de imagen dispuesto para codificar una imagen de alto rango dinámico (M_HDR), que comprende:

25 - una unidad (104) de conversión de rango dinámico dispuesta para convertir la imagen de alto rango dinámico en una imagen de rango dinámico de luminancia más baja (LDR_o), comprendiendo la unidad (104) de conversión de rango dinámico: a) un normalizador (601) dispuesto para normalizar la imagen de alto rango dinámico a un eje luma que se extiende sobre [0.1] y para emitir luminancias normalizadas (Yn_HDR), b) una unidad (602) de conversión gamma dispuesta para aplicar una función gamma a las luminancias normalizadas y para emitir luminancias (xg) convertidas a gamma, c) una primera unidad (603) de mapeo de tono dispuesta para aplicar un primer mapeo de tonos que produce lumas (v) que se define como $v = \log(1 + (RHO - 1) * xg) / \log(RHO)$, con RHO teniendo un valor predeterminado, d) una unidad (604) de mapeo de tono arbitraria dispuesta para aplicar una función arbitrariamente creciente monotónicamente que mapea las lumas (v) para producir lumas (Yn_LDR) de la imagen de rango dinámico inferior (LDR_o); y el codificador (100) de imagen que comprende además:

30 - un compresor (108) de imagen dispuesto para aplicar una transformación de reducción de datos a los colores de la imagen de rango dinámico inferior (LDR_o) que colores están organizados en imágenes componentes, y que transformación de reducción involucra al menos aplicar una transformación DCT a bloques de valores de componentes de color adyacentes, produciendo una codificación comprimida (LDR_c) de los colores de píxeles de la imagen de rango dinámico de luminancia inferior; y

35 - un formateador (110) dispuesto para emitir en una señal de imagen (S_im) la codificación comprimida (LDR_c), y dispuesto además para emitir en la señal de imagen (S_im) los valores que codifican la forma de la función de la función de mapeo de tonos arbitrariamente creciente monotónicamente como metadatos, que permiten que un receptor reconstruya una imagen de alto rango dinámico (Rec_HDR) basada en la imagen de rango dinámico de luminancia más baja (LDR_o).

3. Un decodificador (150) de imagen dispuesto para recibir una señal de imagen de alto rango dinámico (S_im) y que comprende:

40 - un desformateador (151) dispuesto para obtener una imagen comprimida pixelada de rango dinámico inferior (LDR_c) y datos (P) de parámetros fuera de la señal de imagen (S_im); y

45 - un descompresor (152) dispuesto para aplicar al menos una transformación DCT inversa a la imagen comprimida de rango dinámico inferior pixelada (LDR_c) para obtener una imagen pixelada de rango dinámico inferior (LDR_t); y una unidad (153) de conversión de rango dinámico dispuesta para transformar la imagen de rango dinámico inferior (LDR_t) en una imagen reconstruida de alto rango dinámico (Rec_HDR), en donde la unidad (153) de conversión de rango dinámico comprende: a) una unidad (402) de mapeo de tono arbitrario dispuesta para aplicar un mapeo de tono monotónicamente creciente arbitrario, los parámetros que lo definen (P_CC) se reciben en los datos de parámetros (P), b) una primera unidad (403) de mapeo de tono dispuesta para aplicar un mapeo como lo define una función de la forma: $xg = (\text{potencia}(RHO, v) - 1) / (RHO - 1)$, siendo RHO una constante dependiendo de la recepción en los metadatos de un brillo máximo (PB_HDR) de una pantalla de referencia de la imagen de alto rango dinámico y c) una unidad (404) de conversión gamma dispuesta para aplicar un mapeo gamma con un valor gamma recibido (GAM).

4. Un método para decodificar una señal de imagen de alto rango dinámico (S_{im}) que comprende:

- 5 - obtener una imagen comprimida pixelada de rango dinámico inferior (LDR_c) y datos de parámetros (P) fuera de la señal de imagen (S_{im}); descomprimir la imagen comprimida de rango dinámico inferior pixelada (LDR_c) aplicando al menos una transformación DCT inversa a la imagen comprimida de rango dinámico inferior pixelada (LDR_c) para obtener una imagen pixelada de rango dinámico inferior (LDR_t); y transformar la imagen de rango dinámico inferior (LDR_t) en una imagen reconstruida de alto rango dinámico (Rec_HDR), mediante: a) aplicar un mapeo de tono monotónicamente creciente arbitrario, recibiendo los parámetros que lo definen (P_CC) en los datos de parámetros (P), b) aplicar un mapeo de acuerdo con lo definido por un parámetro (RHO) que depende de un brillo máximo (PB_HDR) de una pantalla de referencia de la imagen de alto rango dinámico que se recibe en los datos del parámetro, el mapeo está definido por una función del forma: $xg=(potencia(RHO,v)-1)/(RHO-1)$, y c) aplicar un mapeo gamma con un valor gamma recibido (GAM), que preferiblemente es igual a 2.4.
- 10

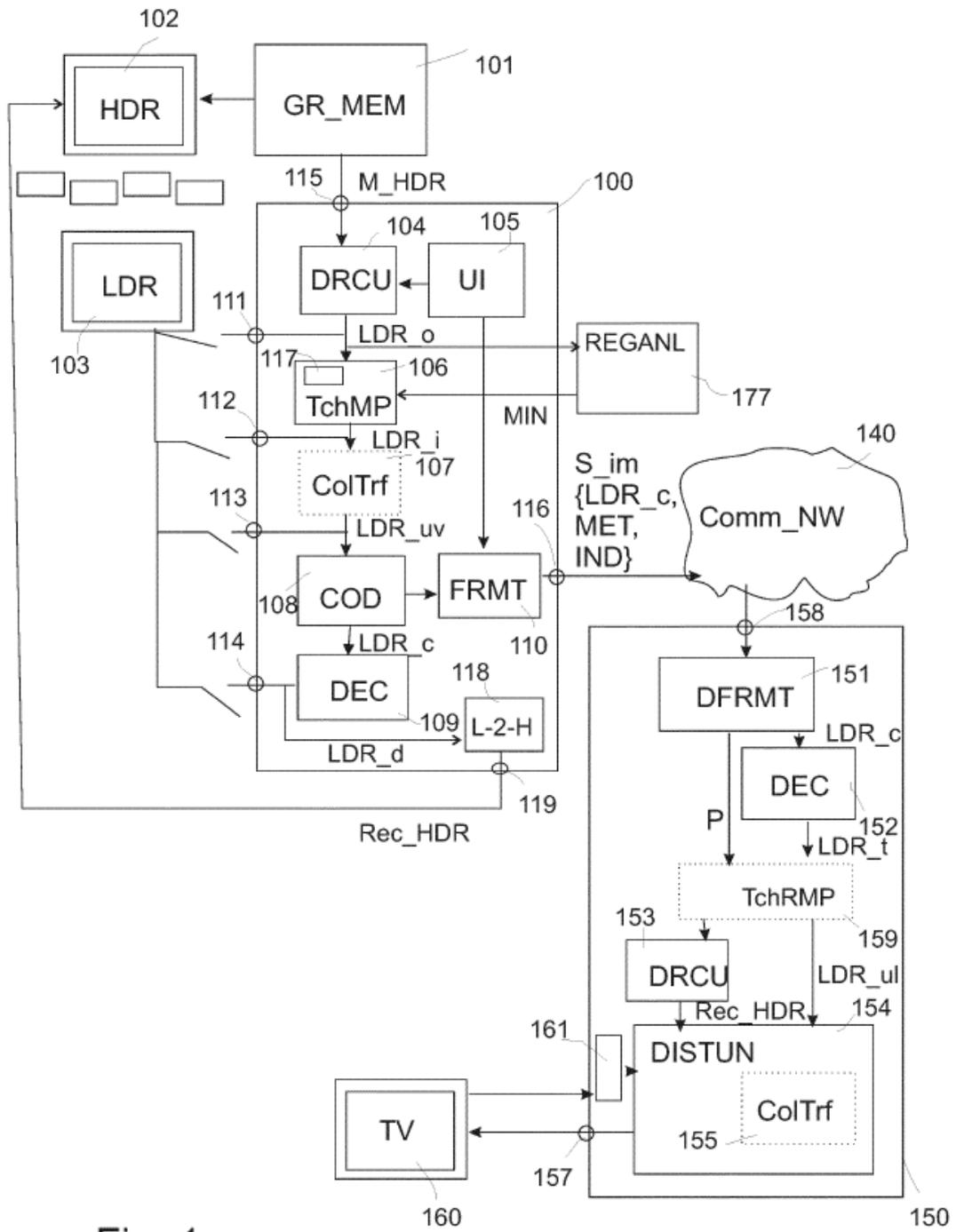


Fig. 1

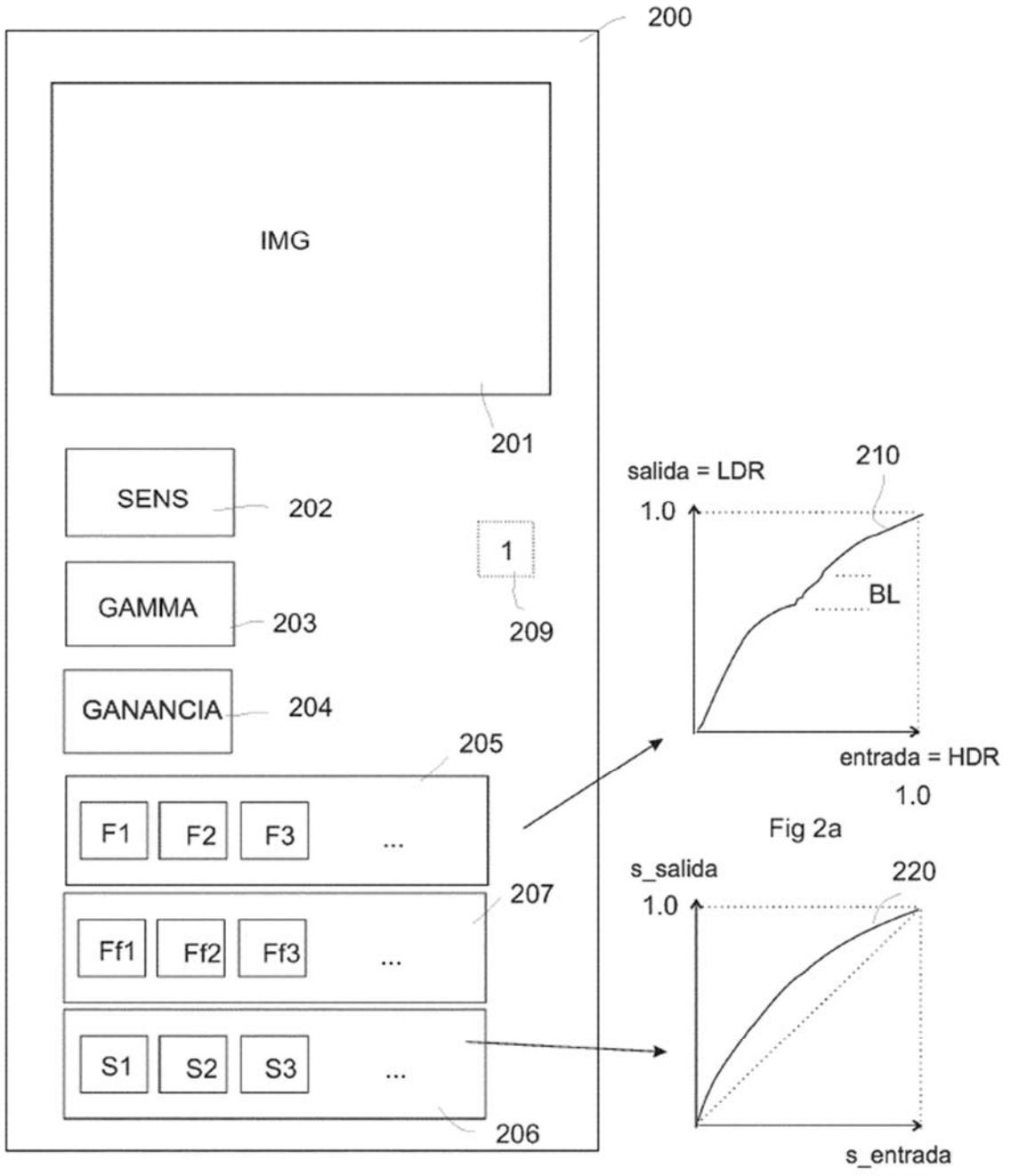


Fig. 2

Fig. 2b

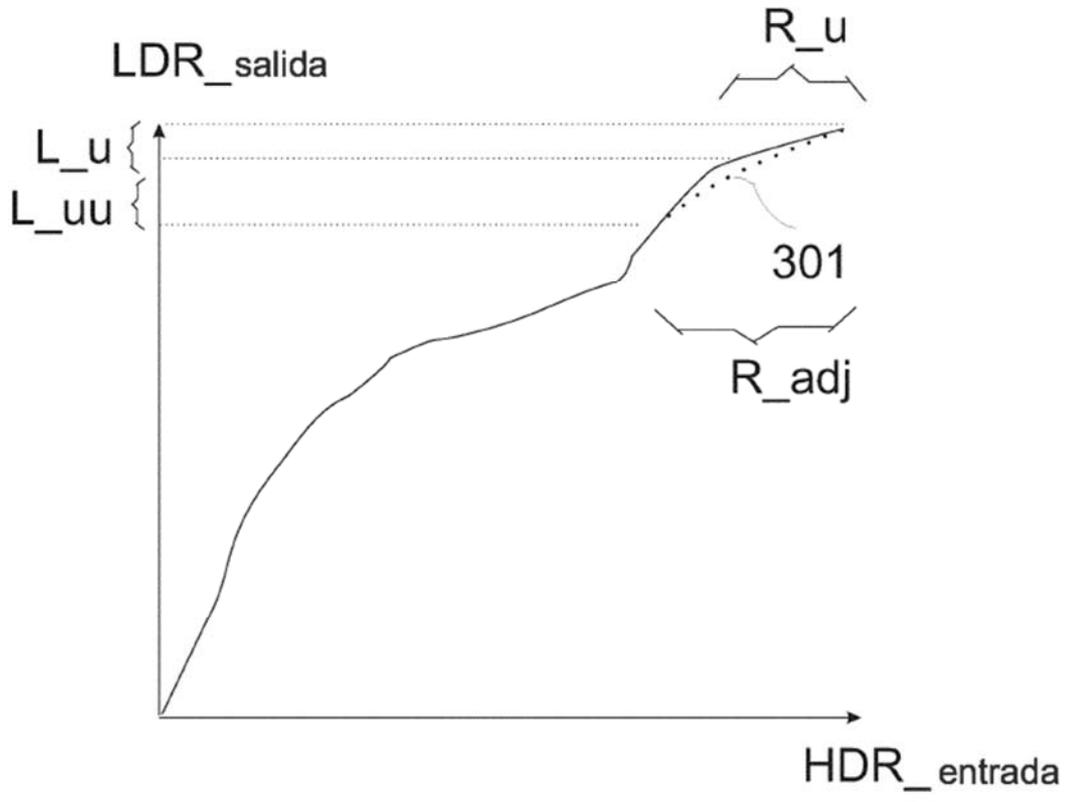


Fig. 3

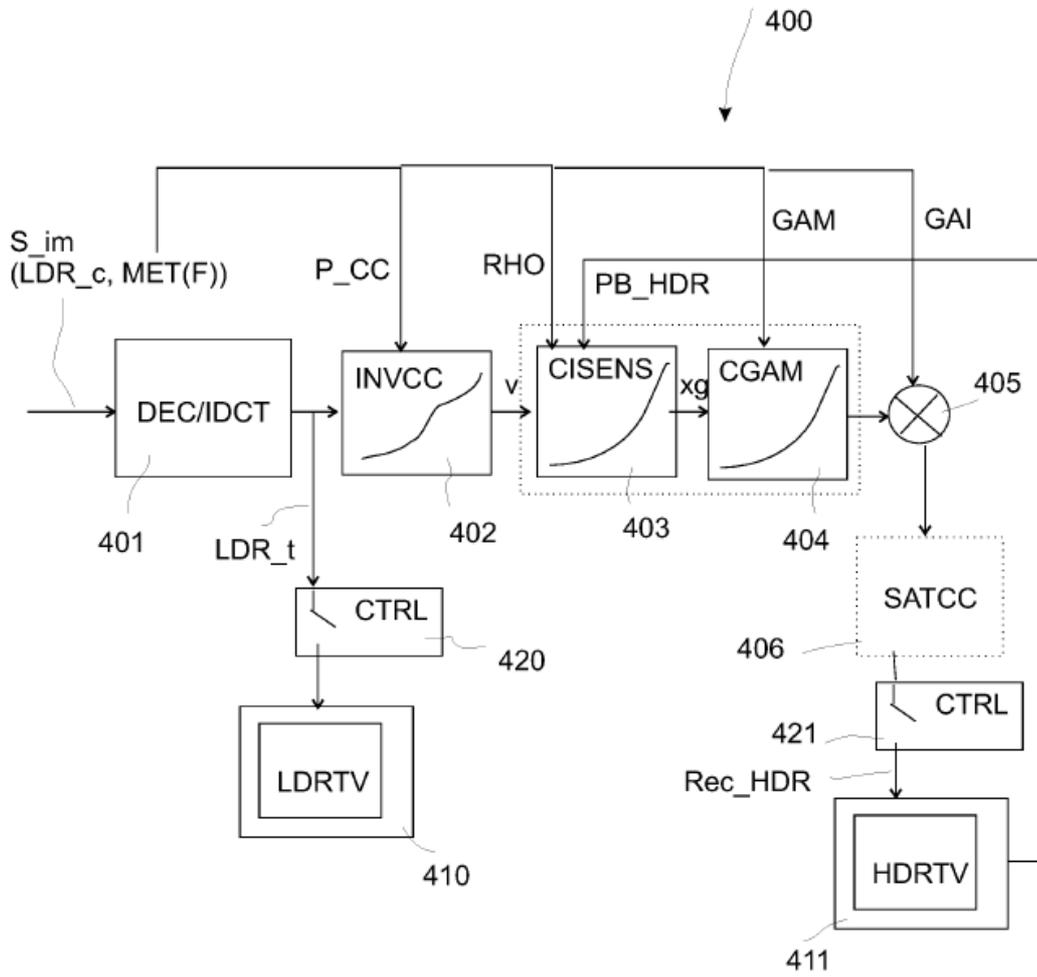


Fig. 4

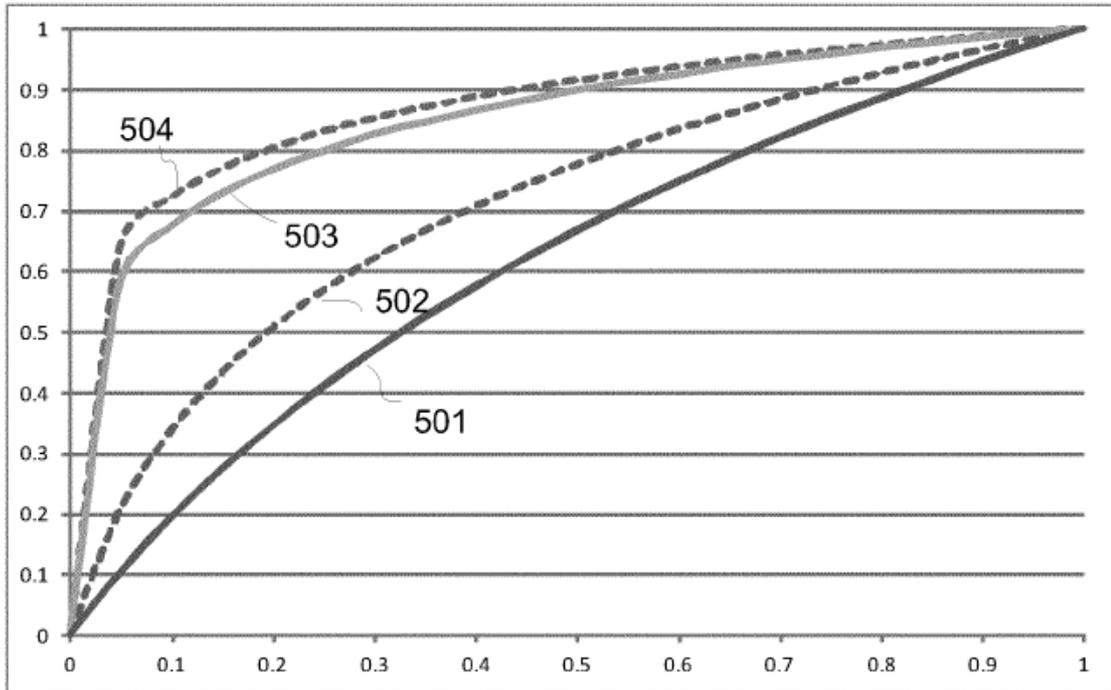


Fig. 5

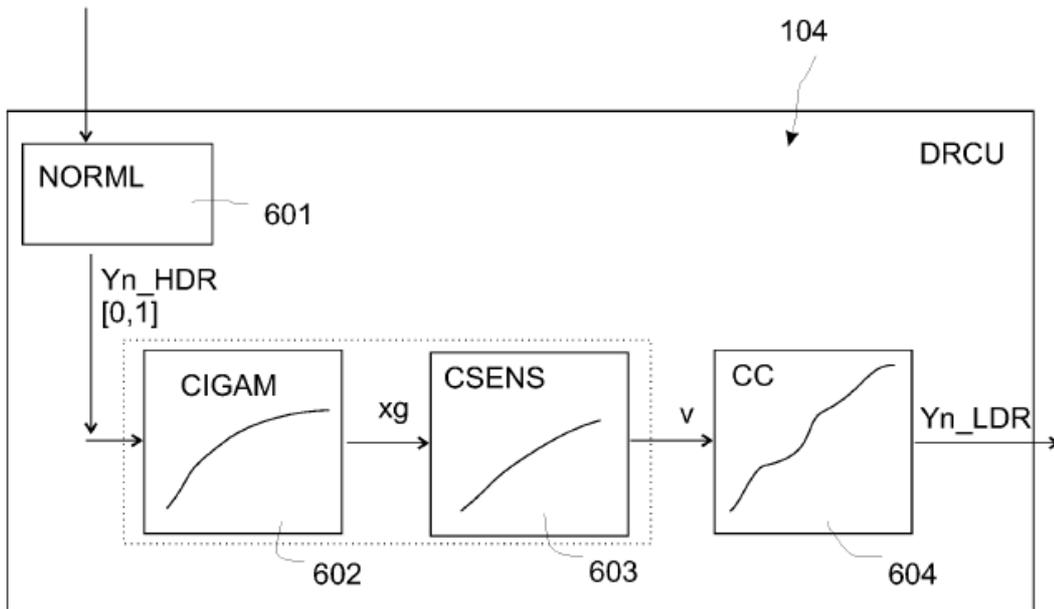


Fig. 6