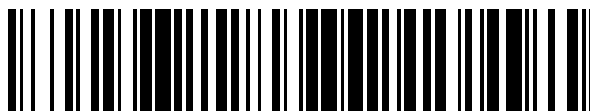


19



OFICINA ESPAÑOLA DE
PATENTES Y MARCAS

ESPAÑA



11 Número de publicación: **2 737 993**

51 Int. Cl.:

H04N 19/30 (2014.01)
H04N 19/46 (2014.01)
H04N 19/176 (2014.01)
H04N 19/102 (2014.01)
H04N 19/14 (2014.01)
H04N 19/17 (2014.01)

12

TRADUCCIÓN DE PATENTE EUROPEA

T3

- 86 Fecha de presentación y número de la solicitud internacional: **25.03.2013 PCT/IB2013/052349**
- 87 Fecha y número de publicación internacional: **03.10.2013 WO13144809**
- 96 Fecha de presentación y número de la solicitud europea: **25.03.2013 E 13721100 (9)**
- 97 Fecha y número de publicación de la concesión europea: **08.05.2019 EP 2880624**

54 Título: **Aparatos y métodos basados en región de brillo para codificación y decodificación de imágenes HDR**

30 Prioridad:

26.03.2012 US 201261615409 P

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

17.01.2020

73 Titular/es:

**KONINKLIJKE PHILIPS N.V. (100.0%)
High Tech Campus 5
5656 AE Eindhoven, NL**

72 Inventor/es:

MERTENS, MARK JOZEF WILLEM

74 Agente/Representante:

ISERN JARA, Jorge

ES 2 737 993 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

Aparatos y métodos basados en región de brillo para codificación y decodificación de imágenes HDR

5 Campo de la invención

La invención se refiere a aparatos y métodos y productos resultantes como productos de almacenamiento de datos o señales codificadas para codificación mejorada de al menos una imagen o video, en particular codificación de imagen o imágenes con un rango aumentado en comparación con imágenes de versiones anteriores (llamadas imágenes HDR de alto rango dinámico, y las imágenes de versiones anteriores se llaman bajo rango dinámico LDR), y codificación de información de imagen con una cantidad creciente de información de luminosidad (también conocido como alto rango dinámico) a o desde diversas representaciones de imágenes.

15 Antecedentes de la invención

Recientemente se han producido nuevos desarrollos relacionados con la codificación de imágenes/video (ya sea de escenas capturadas o gráficos de ordenador), en concreto, es deseable capturar mejor todo el rango de luminancias de objeto y colores que se producen en la naturaleza, hasta grandes valores de luma como, por ejemplo, 25000 nits (por ejemplo, nubes iluminadas por el sol) y a menudo también valores bajos como 0,01 nit, que se llama codificación HDR (alto rango dinámico). Hasta ahora, sistemas de captura de imágenes clásicos (es decir, la cadena comenzando en la cámara -e incluso iluminación de escena apropiada que era habitualmente relativamente uniforme- seguida por la codificación para, por ejemplo, almacenamiento o transmisión de imágenes, hasta la visualización de la imagen) han tratado escenas de alto rango dinámico (es decir, escenas en las que existen simultáneamente regiones oscuras importantes con bajas luminancias y objetos significativos en las mismas, y regiones brillantes con altas luminancias, en particular si existen también diversas regiones importantes de luminancias intermedias (diversos grises), en particular si varias de esas luminancias de escena no puede correlacionarse fácilmente con lo que es utilizable por un componente en la cadena, tal como, por ejemplo, una renderización basada en correlación lineal en un visualizador) de una forma muy distorsionada. Por ejemplo, si la acción estuviera sucediendo dentro de un volumen cerrado de un primer nivel de luz (iluminancia), tal como un coche o una habitación, regiones de iluminación más brillante, tal como el entorno visto a través de la ventana pueden haberse capturado, o al menos representarse en la señal con calidad muy baja (a saber, colores pastel, descoloridos o recortados). Esto es especialmente así para cámaras basadas en CMOS baratas, en comparación con el comportamiento más indulgente de, por ejemplo, la película de celuloide. En particular, únicamente unos pocos valores de código difícilmente representativos se han asociado con los objetos en estas regiones brillantes, que pueden resultar en mala representación de las texturas de objetos, en incluso recorte directo al valor máximo del espacio de colores usado para codificación. Tener tan pocos datos en estas regiones del eje de luminancia de la imagen capturada, también significa que las funciones de procesamiento, por ejemplo, que optimizan el contraste de imagen visualizada, pueden tener problemas para reproducir datos de píxeles finales buenos. Teniendo disponibles los visualizadores cada vez mejores hoy en día y en el futuro cercano (por ejemplo, con brillo máximo de varios miles de nits), o al menos tecnologías de procesamiento de imágenes más inteligentes, se puede desear mejorar esa situación, para ser capaz de crear imágenes renderizadas de calidad superior.

Por varias razones, al menos para un número de años en el futuro, se puede desear alguna forma de compatibilidad hacia atrás, que significa que datos de una así llamada codificación de bajo rango dinámico(LDR) debe estar disponible o al menos ser determinable fácilmente a partir de la codificación disponible, de modo que, por ejemplo, una caja de procesamiento de video mejorada nueva puede distribuir una señal LDR a un visualización de menor rango dinámico (por ejemplo, un visualizador móvil). También desde un punto de vista de almacenamiento, puede ser muy útil almacenar una señal de imagen de una manera tan versátil como sea posible, es decir, no solo con la máxima cantidad de datos útiles acerca de la escena, sino también de una manera que estos datos servirán a muchas potenciales aplicaciones futuras, especialmente si es de una forma simple. Habitualmente el rodaje de una película, por ejemplo, conlleva tanto esfuerzo, que la señal en bruto es altamente valiosa, y se codifica mejor esto de la mejor forma posible que permite la tecnología. No caer en una trampa de que incluso la codificación maestra de un programa es para una posterior generación de sistemas de visualización de mejor calidad por debajo de lo que podría haberse conseguido si los datos se codifican de forma diferente. Eso evita no únicamente tener que hacer una maniobra cara por todas partes, sino que el lector puede imaginar que algunos eventos a grabar, como el matrimonio de una pareja de la realeza o un video de una boda de una pareja normal, no se harán de nuevo. Y tratar de remasterizar un video de este tipo para una generación nueva de tecnología de visualización es, si no muy difícil, al menos complicado. Es preferible que la codificación permita la captura de la escena óptimamente en el primer lugar, e incluso permita fácilmente posteriores mejoras, por su propia estructura de codificación. Independiente de cómo se renderice en un visualizador particular más entorno de visualización, la información presente en codificaciones LDR actuales tal como JPEG (dependiendo entre otros de la escena capturada particular y sistema de cámara usado), se ve en la actualidad como (limitado a) aproximadamente 11 bits lineales o paradas. Por supuesto, si la codificación se tiene que usar directamente para renderización (por ejemplo, referido a visualización) algunos de los bits de información pueden no ser visibles. Por otra parte, un códec puede contener información de la escena original o composición de gráficos (referido a escena), que puede volverse relevante, por ejemplo, cuando un visualizador está cambiando su gama visible por seres humanos por medio de procesamiento de imágenes. Por tanto, es importante tener al menos los objetos de imágenes importantes bien representados en la imagen codificada.

Una cadena de captura HDR es más que simplemente apuntar una cámara a una escena con una gran relación de contraste de luminancia entre el objeto más oscuro y más brillante y grabar linealmente lo que hay en la escena. Tiene que ver con qué son exactamente los valores de gris intermedios para todos los objetos, ya que eso transmite, por ejemplo, el ambiente de una película (oscurecer ya algunos de los objetos en la escena puede transmitir un ambiente oscuro). Y este es un proceso psicológico complejo. Se puede imaginar, por ejemplo, que psicológicamente no es tan importante si una luz brillante se renderiza en un visualizador exactamente en tal proporción al resto de los valores grises renderizados como la luminancia de escena de esa luz fue para el resto de las luminancias de objetos de escena. En su lugar, se tendrá una impresión fiel de una lámpara real si los píxeles se renderizan con "alguna" luminancia de salida de visualizador alta, siempre que sea lo suficientemente más alta que el resto de la instantánea. Pero esa distribución entre objetos auto luminosos y reflectantes (en las diversas regiones de iluminación de la escena) es también una tarea que depende de la gama de visualizador y condiciones de visualización típicas. También, se puede imaginar que la codificación de las regiones más oscuras se hace preferentemente de modo que pueden usarse fácilmente en diferentes escenarios de renderización tal como diferentes niveles de iluminación envolvente promedio, teniendo diferentes niveles de visibilidad para el contenido de imagen más oscuro. En general, porque esta es una difícil tarea psicológica, artistas se implicarán en la creación de imágenes óptimas, que se llama gradación de color. En particular, es muy útil cuando los artistas hacen una gradación de LDR separada, incluso si eso se hace en una "estrategia de codificación HDR pura". En otras palabras en un escenario de este tipo cuando se codifica una sola señal RAW de cámara HDR, también se generará una imagen LDR, no necesariamente porque se usa para una gran fracción LDR del mercado de consumo de video, sino porque transmite una información importante acerca de la escena. A saber, siempre habrá regiones y objetos más importantes en la escena, y poniendo a estos en una subestructura LDR (que conceptualmente puede verse como una parte opuesta artística de un algoritmo de exposición automático, aún después de captura total, y en relación con luminancias capturadas fuera de ese rango), esto hace más fácil hacer toda clase de conversiones a representaciones de rango intermedio (MDR), adecuado para el accionamiento de visualizadores con características de renderización y visualización particulares. Usando un marco técnico de este tipo, se puede incluso con una única imagen de codificación, al mismo tiempo adaptar, por ejemplo, visualizadores LDR como un visualizador móvil con un brillo máximo de 50 nit (en interiores, o un mayor brillo pero compitiendo contra iluminación de exteriores alta), un visualizador MDR de brillo máximo de medio rango de digamos 1200 nits, y un visualizador de HDR de digamos brillo máximo de 8000 nits. En particular se puede ajustar esta parte LDR de acuerdo con varios criterios, por ejemplo, que renderiza con buena calidad en un visualizador LDR de referencia estándar (los colores parecen lo más similares posibles a los del visualizador HDR), o transmite un cierto porcentaje de la información capturada local (por ejemplo, una cierta cantidad de la imagen es visible), etc. Implementaremos en nuestro códec propuesto a continuación que tal visualizador de recepción puede, a partir de esa única codificación (o gradación) de escena que abarca todo, identificar fácilmente qué son, por ejemplo, las regiones oscuras, de modo que puede adaptar óptimamente la visibilidad incorporada de las mismas dadas sus características conocidas del sistema de visualización.

No existen tantas formas de codificar una señal HDR. Normalmente en la técnica anterior se codifica nativamente la señal HDR, es decir se mapea (linealmente) los píxeles con, por ejemplo, palabras flotantes de 16 bits, y a continuación el valor de luminancia capturado máximo en el blanco HDR en una filosofía similar a la codificación LDR (aunque psicovisualmente esto normalmente no es ningún blanco reflectivo en la escena, sino un color brillante de una lámpara). Esto es una codificación referida una escena nativa de las luminancias de objeto de escena original según se capturan por la cámara. También se podría correlacionar una señal HDR de rango total con el rango LDR de 8 bits a través de alguna función de transformación de luma "óptima", que habitualmente sería una función gamma o similar. Esto puede implicar la pérdida de precisión de color (en vista de la compensación entre rango y precisión para tales codificaciones) con correspondientes problemas de calidad de renderizado, especialmente si en el lado de recepción se espera procesamiento de imágenes tal como brillo local, sin embargo la gradación de valores grises dominante de los objetos de imagen (por ejemplo, el promedio sobre un objeto) se preserva difícilmente (es decir, sus relaciones de luma relativa/porcentual).

La técnica anterior también ha descrito algunas técnicas de codificación HDR usando dos conjuntos de datos de instantáneas para cada imagen HDR, habitualmente basándose en una clase de concepto de codificación escalable, en el que mediante alguna función de predicción, la precisión de una textura local codificada "LDR" se refina, o indica con más precisión, es decir, se proyecta a una versión HDR de esa textura, habitualmente escalando las luminancias LDR (el LDR en esas tecnologías normalmente no es un grado LDR de buen aspecto ya adecuado para renderización óptima en un visualizador LDR (de referencia) típico, sino habitualmente un procesamiento simple en la entrada HDR). A continuación la segunda instantánea es una instantánea de corrección para acercar la imagen HDR predicha a la imagen HDR original a codificar. Existe alguna similitud a las codificaciones de una sola imagen HDR, a través de las funciones de predicción que sirven como un criterio de definición de rango/precisión, únicamente en estas tecnologías la codificación se realiza con dos instantáneas.

Escalar las luminancias de una imagen de banda base implica aplicar una transformación, y esta transformación de predicción a menudo se define por bloque, para reducir la cantidad de datos a codificar. Esto puede ser ya un desperdicio de datos, ya que muchos bloques contendrán el mismo objeto y, por lo tanto, necesitan la misma transformación.

Como se ha dicho, la diferencia de la imagen HDR original con la predicción puede co-codificarse como una

instantánea de mejora al grado deseado, tanto como sea posible dado el rango y definición de la imagen de mejora. Por ejemplo, puede representarse un valor de gris HDR de 1168 con una división por 8 a un valor 146. Este valor HDR podría recrearse multiplicando de nuevo por 8, pero ya que un valor 1169 cuantificaría al mismo valor de capa base 146, se necesitaría un valor de mejora igual a 1 para ser capaz de recrear una señal HDR de alta calidad. Un ejemplo de una tecnología de este tipo se describe en la patente EP2009921 [Liu Shan et al. Mitsubishi Electric: Method for inverse tone mapping (by scaling and offset)]. Una cuestión interesante sobre tales métodos es siempre lo que el método de mejora realmente trae como mejora de información visual. Normalmente se aplica de forma ciega y puede, por ejemplo, para regiones texturizadas no contribuir en ocasiones información adicional relevante, especialmente para video con cambios rápidos.

Otra codificación de dos instantáneas se describe en la aún no publicada en la actualidad solicitud US61/557461.

Ahora existen problemas con todas las codificaciones HDR existentes. Solo aplicar transformaciones globales puede ser demasiada aproximada de acuerdo con lo que el creador de contenido desea después de haber invertido tanto en, por ejemplo, una película (efectos especiales). Otras aplicaciones pueden ser menos críticas, como la realización de un programa de televisión, pero aún es deseable un buen control sobre el aspecto final. Eso al menos tendría el coste de necesitar muchos bits de datos codificados. Por otra parte, especificar transformaciones intrincadas por píxel también implica una gran cantidad de datos a codificar. Esto se aplica, por ejemplo, a necesitar codificar una segunda imagen que es un mapa de potenciación de luminosidad, para que se codifiquen reflejos de textura de objetos en una primera imagen. También, junto con esto, se codifica ciegamente cualquier cosa que se produce posiblemente en la entrada, sin conocer mucho acerca de lo que hay realmente en la imagen (es decir, sin permitir un uso versátil), incluso sin darse cuenta que puede haber una gran cantidad de redundancia en la imagen potenciada. No digamos que estos datos ciegos son fáciles de usar por algoritmos inteligentes como, por ejemplo, mejora de instantánea o algoritmos de optimización en el lado de visualizador.

Trabajar en una base de bloques reduce la cantidad de datos, pero todavía no es óptimo. En particular siendo esta estructura de bloque también bastante ciega para el contenido de imagen real, y más molesta, imponiendo una nueva estructura geométrica que es la cuadrícula de bloques, que no tiene nada que ver con la imagen subyacente y, por lo tanto, puede coincidir más o menos convenientemente con las características de imagen (en particular la geometría de imagen), significa que pueden producirse varios artefactos relacionados con codificación de bloques. De hecho, un bloque no es mucho más que simplemente un gran píxel, y no realmente una estructura relacionada con contenido inteligente (ni en lo que se refiere a la estructura geométrica de color de ese objeto o región, ni su significado semántico, tal como, por ejemplo, siendo un objeto que debería estar mayormente escondido en la oscuridad).

El documento WO 2011/107905 A1 (KONINKL PHILIPS ELECTRONICS NV [NL]; MERTENS MARK JOZEF WILLEM [NL]) 9 de septiembre de 2011 (09-09-2011) divulga un método y un sistema para permitir una mejor coordinación entre la renderización deseada en el lado de producción de imagen y el renderizado real en el lado de visualizador. Información de definición de imagen, tal como regímenes de color, se codifica junto con la señal de imagen. Algoritmos de procesamiento de imagen en el lado de visualizador pueden aplicar a continuación operaciones de procesamiento de imágenes para modificar el aspecto de regiones locales de acuerdo con los descriptores de régimen.

Las realizaciones a continuación tienen como objetivo proporcionar medidas técnicas sencillas para mitigar al menos algunos de esos artefactos.

Sumario de la invención

Una codificación utilizable simple y fácil de imágenes HDR puede realizarse mediante los conceptos de realizaciones presentados en este documento que siguen los principios relacionados con un codificador de imagen (549) para codificar una imagen de una escena de alto rango dinámico, que comprende las características como se definen en la reivindicación 1.

Solo con únicamente uno o pocos tal valor o valores de gris de diferenciador de regiones, se puede transmitir ya la característica esencial de una escena HDR, tal como que existe una región "por encima de blanco" o "exceso de brillo" en la imagen. Por encima de blanco significa luminancias de escena por encima de blanco en la región iluminada, por ejemplo, el blanco que se grabaría desde un papel blanco iluminado normalmente (de acuerdo con diseñador de iluminación del set por ejemplo), en la parte principal de la escena. Los diferenciadores son una buena forma de codificar el contenido semántico de una escena. Por ejemplo, no existe solo un blanco en una escena real, como supone la codificación de imagen clásica. En la codificación LDR de imagen clásica, de hecho se ilumina la escena en la que se produce la acción aproximadamente uniformemente, y a continuación el objeto reflectante más blanco (en la iluminación más brillante en la región de imagen principal) habitualmente determinará el punto blanco de la codificación de imagen. En lugar de recortar, por ejemplo, objetos de exteriores, también se podrían incluir algunos objetos por encima de blanco, especificando el cámara un punto umbral de compresión para la curva de gamma de reproducción, pero que aún se vincula con el blanco principal (por ejemplo, 6x por encima de ese blanco).

En una escena real, puede existir, por ejemplo, un sol muy brillante en el exterior. Incluso cuando se agrupan estas dos regiones en una menor cantidad de codificación de bits (por ejemplo, representando la misma como una

instantánea clásica de 8 bits), se querría que estas dos regiones/rangos se separasen entre sí en el eje de luma. Esto significa que posteriormente (o, por ejemplo, intermedio en transcodificación y regradación automática, etc.) se podría tratar de forma más inteligente estas regiones. Ya hemos hablado anteriormente acerca de objetos de lámpara. El visualizador de renderización puede querer renderizar los mismos de acuerdo con un criterio que define uno o más de "tan brillante como puede" y "aunque no demasiado deslumbrante para el espectador". Para hacer esto, sin embargo, puede necesitar tratar esas dos regiones de imagen (lámpara frente al resto de la escena) de forma diferente e incluso de forma discontinua y, por lo tanto, puede necesitar saber qué es el objeto de lámpara en la imagen. Las codificaciones basadas en función de gamma clásicas moverán habitualmente la lámpara durante el posprocesamiento a alguna posición de luminancia renderizada que depende de esa gamma usada, pero no la semántica de escena junto con los detalles colorimétricos de sistema de renderización (tal como capacidades de visualizador, luz ambiental, etc.). Un razonamiento técnico similar puede hacerse para las regiones más oscuras, si se conoce su composición de región de luminancia, es decir, por ejemplo, un par de rangos de oscuro, por ejemplo, "oscuro brillante", "oscuro promedio", y "súper oscuro". Tales códigos (es decir, diferenciadores de valores de gris) podrían significar algo numéricamente, pero nuestro método permite que el graduador de color haga, por ejemplo, el HDR maestro final para almacenamiento en un disco Blu-ray, para colocar a los mismos con regiones semánticamente significativas. Por ejemplo, en un sótano oscuro en una película de terror, un negro promedio puede ser los colores con los que las paredes tienen que renderizarse (finalmente en el visualizador de renderización, de acuerdo con su mapeo de tono óptimo final para visualizar óptimamente), mientras que oscuro brillante (es decir, el rango de luminancia a renderizar entre negro promedio y oscuro brillante) podría ser las herramientas en esas paredes como cuchillos e instrumentos de tortura para hacer que se vean mejor (dados detalles colorimétricos de lado de renderización) y súper oscuro puede ser, por ejemplo, una esquina oscura, en la que el criminal puede estar escondiéndose. La región de esquina súper oscura es entonces nuestro primer objeto, y la región principal de oscuro promedio nuestro segundo, como en una escena de interior/exterior, el sol exterior puede ser el segundo objeto, y el interior el primer/principal objeto.

También, estas dos subregiones (por ejemplo, acción principal iluminada promedio y lámpara o sol exterior) pueden estar tan juntas en la representación de imagen codificada que se están tocando en realidad para no desperdiciar códigos de luma entre los mismos, que hace los mismos extremadamente difícil de separar ciegamente en el lado de recepción. Aún existiendo este valor de luma particular que marca el límite entre los mismos, que por lo tanto se co-codifica como un valor de gris de diferenciador de regiones (gTS) para un entendimiento de escena fácil (aunque simple) en el lado de recepción. Y esto permite entonces diversas aplicaciones, tal como codificación HDR y fácil reconstrucción a partir de una imagen de 8 bits en el lado de recepción, procesamiento de imagen tal como remapeo de color, etc.

Ventajosamente el codificador de imagen (549) comprende una unidad de determinación de mapeo de luma (553), dispuesta para determinar un mapeo de luma (TOM) para al menos uno del primer y el segundo objeto que define un mapeo entre lumas de píxeles como se codifica en la representación de imagen (Im_1) y luminancias de los píxeles en una segunda representación de imagen (IM_RC_HDR), y se dispone para suministrar el mapeo de luma (TOM) al formateador (554) que se dispone para co-codificar el mismo en la señal de imagen de salida (S(Im_1, MET(gTS), TOM). Tales mapeos de luma pueden determinarse de diversas formas, teniendo en cuenta tales principios como por una parte especificación óptima de la información en la instantánea (por ejemplo, la cantidad de códigos necesarios para codificar relativamente fiel una textura de una complejidad específica, como una granulosidad de madera), y por otra parte un aspecto, por ejemplo, definiendo una posición de luminancia en habitualmente un visualizador de referencia.

El creador de contenido podría dejar el mismo al lado de recepción para hacer este procesamiento deseado, por ejemplo, renderización de visualizador definitiva. Tener solo un gTS es suficiente ya para muchas situaciones, ya que el sistema de lado de recepción sabe entonces obviamente qué son los objetos brillantes, ya que tienen lumas por encima de gTS. Sin embargo, este sistema de co-codificar valores de gris de diferenciador de regiones permite una codificación mucho más versátil de escena HDR (conocimiento acerca de su composición o incluso significado semántico en metadatos) y en consecuencia diversos usos de esos datos. Por ejemplo el creador de contenido puede proporcionar uno o más escenarios en cómo mapear los colores de píxel/lumas codificadas en Im_1 a diversos otros espacios de color, tal como para renderización en visualizadores diferentes. Puede codificar, por ejemplo, varios valores para (aproximadamente) un tipo de visualizador (por ejemplo, que tiene un brillo máximo cercano a 4000 nits, es decir destinado para visualizadores de LCD u OLED con brillos máximo reales entre 3000 y 5000 nits), de modo que el visualizador puede finalmente elegir una estrategia de renderización final de todo el conocimiento de transformación codificado (codificando cómo desea el creador de contenido que sus imágenes se vean). Por ejemplo en visualizadores con un menor rango dinámico visualizable, un único diferenciador para las regiones más brillantes ya puede ser suficiente, ya que no tiene una capacidad alta de este tipo de renderización de regiones brillantes. Sin embargo, un visualizador de 8500 nits puede hacer un uso mucho más ventajoso del contenido si contiene más valores gTS que indican diferentes clases de regiones de brillo, ya que puede, conociendo su subgamma renderizable física de luminancias de brillo, asignar un subrango de luminancias diferente a, por ejemplo, objetos soleados de exteriores como a, por ejemplo, objetos de lámpara de una primera clase, y una región incluso más brillante cerca de brillo máximo para una clase más brillante de objetos de lámpara.

Un creador de contenido con un menor interés en invertir mucho tiempo en la gradación puede, por ejemplo,

únicamente especificar dos gradaciones, por ejemplo, puede empezar desde Im_1 o alguna transformación automática de la misma, como siendo "suficientemente bueno" para renderización LDR, y a continuación tomar algún tiempo para retocar los mapeos para obtener una imagen HDR mejorada (por ejemplo, con regiones exteriores extra brillantes, luces o ventanas). Por tanto, puede especificar, por ejemplo, una imagen LDR de 8 bits (que llamaremos contenedor LDR), y a continuación algunas funciones, en primer lugar funciones de mapeo para recuperar aproximadamente la imagen HDR maestra original (por ejemplo, en una codificación de 16 bits flotante nativa) desde el LDR_container, así como en segundo lugar algunas funciones que permiten uno o más ajustes de esa imagen HDR. Por ejemplo, puede especificar una correlación de las regiones brillantes por encima del, por ejemplo, 90 % para visualizar en un segundo visualizador de referencia de 14000 nits (el primer visualizador de referencia puede ser el visualizador para el que se graduó el HDR maestro original antes de codificar el mismo con un contenedor LDR mapeando descendientemente, por ejemplo, un visualizador de 5000 nits). De manera similar, estas funciones pueden usarse para ajustar descendientemente a visualizadores MDR de aproximadamente 2000 nits, por ejemplo, invirtiendo su comportamiento de mapeo. En sus variantes más simples, el graduador que invierte menos tiempo puede especificar solo uno o más valores gTS para al menos algunas escenas en la película, y dejarlo para el visualizador o renderizador (por ejemplo, impresora) para averiguar qué sería una buena transformación para sus características de renderización.

Un aparato de lado de recepción de procesamiento de imagen puede a continuación, por ejemplo, determinar su grado final a partir de estos dos o más conjuntos de información (la instantánea de contenedor LDR codificada en Im_1 , y el al menos un valor de gris de diferenciación gTS, y si disponible cualquiera que sea la información de función de mapeo que el graduador especifique que está de acuerdo con sus deseos). Por ejemplo, mirando a la Figura 2b, el creador de contenido puede prescribir en la señal que para renderización HDR del objeto muy oscuro, tiene que usarse el mapeo parcial $PD_BLCK(i+2,j)$ (se explica en detalle a continuación), y que puede o debería usarse para renderización LDR del mapeo más brillante $PD_BLCK(i,j)$ (es decir, el objeto muy oscuro se trata entonces como la escalera). Ahora un visualizador de recepción de digamos brillo máximo de 1500 nits puede decidir usar cualquiera de estas dos estrategias (por ejemplo, lo más cercano a su brillo máximo, siendo la gradación/mapeo LDR para como máximo 750 nits (así que probablemente más de 400 nits) y el HDR para al menos, por ejemplo, 2000 nits), o puede interpolar entre las mismas de alguna manera, que significará para estas dos funciones lineales, por ejemplo, aplicar la función lineal a mitad de camino entre las mismas. El sistema permite que el creador de contenido vea el HDR como "efectos HDR", por ejemplo, potenciando una luz brillante, como una bola de plasma emitida desde la mano de un mago.

Nuestro método puede usarse cuando la codificación de Im_1 es una codificación de menores bits (que no es lo mismo que menor rango dinámico) que la imagen original (HDR maestro), por ejemplo, con una codificación de luma de 8 o 10 bits clásica. En este caso, esa imagen Im_1 puede definirse para un visualizador de referencia de un menor rango dinámico (y brillo máximo habitualmente de, por ejemplo, 500 nits) y diferenciadores gTS pueden ser útiles para determinar automáticamente gradaciones para rangos dinámicos de visualizador más altos (por ejemplo, para un visualizador con brillo máximo de 2000 nits). Pero de forma similar, por supuesto, la única imagen codificada Im_1 puede especificarse, por ejemplo, para ese visualizador de referencia de 2000 nits (es decir, utilizable directamente para accionar ese visualizador, o al menos requerir menores modificaciones colorimétricas antes del renderizado), y en un escenario de este tipo los valores gTS (y otros datos como especificaciones de funciones de transformación) pueden ser útiles entre otros para mapear descendientemente para obtener imágenes de accionamiento para visualizadores de menor rango dinámico, como, por ejemplo, un visualizador portátil.

Es decir ventajosamente el codificador de imagen (549) opera en un escenario de uso y configuración técnica de modo que uno de la primera y segunda representaciones de imagen es una representación de alto rango dinámico, codificándose la representación HDR, por ejemplo, para un visualizador de referencia con brillo máximo al menos por encima de 750 nits. Es decir, será utilizable sin gran modificación adicional para accionar un visualizador HDR para renderizar la imagen aproximadamente como pretendía el artista. Una representación HDR de este tipo puede, por ejemplo, ser un número entero de 3×32 bits o representación flotante de 3×16 bits (RGB, o YC1C2, etc.), etc. Aunque esta estrategia de codificación puede usarse en diversos escenarios entre diversas representaciones de espacio de color (por ejemplo, entre una primera representación HDR de 16 bits con primer blanco, función de gamma asignando los códigos de luma, etc. y una segunda, por ejemplo, representación HDR de 12 bits), es especialmente útil si al menos una de la imágenes (de entrada o salida) o al menos parte de la imágenes es alto rango dinámico (es decir, ya sea así codificada o así obtenida utilizable con alta calidad colorimétrica accionando un visualizador de renderización HDR, etc.), y en particular es útil cuando se codifica HDR con palabras de luma de pocos bits (es decir, por ejemplo, 8 o 10 en lugar de, por ejemplo, 14 o 20), en cuyo caso puede usarse en sistemas de capacidades de versiones anteriores o capacidades cercanas a la misma. Para una explicación completa, la reciente expresión comúnmente conocida alto rango dinámico habitualmente significa mayores brillos en escena original o renderización, mayor que en los escenarios de formación de imagen LDR actuales clásicos, o incluso más exacto como se describe a continuación: un rango mayor de aspectos de luminosidad (por ejemplo, de acuerdo con el sistema visual humano de un espectador, pero por supuesto incorporado en códigos técnicos como lumas). Aunque se puede definir bien un visualizador referido de señal de este tipo con referencia a un visualizador definitivo de capacidades que son máximas para tecnologías esperadas en un futuro razonablemente lejano, idealmente la imagen HDR se define al menos parcialmente referida a escena (ya que nunca se sabe qué hará un visualizador futuro para explotar el contenido codificado y los aficionados del formato bruto dirían que la imagen necesita almacenar al menos lo que potencialmente una cámara o programa gráfico de muy alta calidad puede capturar o producir, pero incluso entonces en lugar de usar un modelo de cámara de referencia, se pueden aún codificar esto como la escena que es aproximada mediante un

visualizador de muy alta calidad de la misma. De hecho, cualquier codificación entre 0 y L_{max} , con la función de asignación de código que sea, también puede verse como renderizable en un visualizador teórico de este tipo que tiene un brillo máximo de L_{max} , e incluso en el futuro lejano dadas las limitaciones fijas de la visión humana, no se necesitaría nunca realmente que renderizar la luminancia del sol fielmente, no en grandes visualizadores que abarcan una pared, y especialmente no para visualizadores pequeños en los que todo contenido de imagen se ve en un ángulo sólido pequeño. Por tanto, el graduador puede elegir codificar la imagen con cualquier visualizador de referencia que desee, ya sea uno teórico definitivo de 500.000 nits de brillo máximo, o uno más pragmático como de 10.000 nits, siempre que co-especifique esta colorimetría que define metadatos en su definición de códec.

Ventajosamente el codificador de imagen (549) se dispone de modo que codifica varios valores de gris de diferenciador de regiones (gTS_D_Loc_1, gTS_D_Loc_2) entre secciones que comprenden varias de las palabras de código de N bits que codifican los colores de píxel a partir de la representación de imagen (Im_1). Esto permite que el creador (o incluso software de procesamiento automático) pueda asignar, por ejemplo, diferentes valores de, por ejemplo, las "partes de sombra más oscuras" para diferentes partes de la imagen, y así tener mayor control sobre la capacidad de ajuste de la imagen. Por ejemplo, en una región geométrica central en la imagen se puede ocultar (mostrar) los objetos más oscuros que se definen como por debajo de, por ejemplo, valor de código 10, y en una esquina los objetos más oscuros están por debajo de valor de código 5. Esto puede manejar diversas situaciones físicas, como, por ejemplo, cambio geométrico de iluminación, en la que la relación entre píxeles de objetos oscuros y más oscuros puede cambiar varias veces para bloques que suceden a una redefinición de un gTS local. La codificación actual de los valores de gris de diferenciador de regiones en relación física (por ejemplo, en una memoria transportable) a los datos de textura de imagen (Im_1) puede hacerse de diversas formas, pero es ventajoso si los metadatos requeridos se codifican entremezclados con los datos de bloque de color de píxel, precisamente en estas ubicaciones en las que es aplicable, es decir habitualmente antes de que el primer bloque en una instantánea que tiene una relación dual de valores de gris por debajo y por encima de gTS (a usarse para segmentación o procesamiento de los bloques siguientes habitualmente).

Ventajosamente el codificador de imagen (549) se dispone por que codifica un valor de gris de diferenciador de regiones antes de una ejecución de varias imágenes sucesivas, siendo esto un valor de gris de diferenciador de regiones para todas esas imágenes sucesivas. Por supuesto, los valores de gris de diferenciador de regiones más importantes pueden codificarse en menos intervalos regulares, ya que pueden ser aplicables, por ejemplo, a una toma o escena entera. Por ejemplo se pueden codificar varias estrategias para codificar las áreas más oscuras para diferentes ambientes de renderización para una parte de terror aterrador oscura de una película. Más adelante, en la película, en una escena de día exterior, se puede codificar de forma separada una estrategia de abrillantado a usar para el cielo, antes de la primera imagen de esa escena. Esto permite especificar el procesamiento en una base de tomas o escenas, por ejemplo, definiendo las partes más oscuras de un sótano, y una definición de este tipo puede producirse de nuevo después de una toma intermitente de digamos fuera entre dos tomas de sótano oscuras.

Ventajosamente el codificador de imagen (549) se dispone por que codifica al menos un valor de gris de diferenciador de regiones en una memoria no físicamente adyacente a una memoria que almacena la representación de imagen (Im_1), junto con un código de asociación geométrica que permite la asociación de cada respectivo al menos un valor de gris de diferenciador de regiones con una región geométrica de la representación de imagen (Im_1) a la que es aplicable, comprendiendo el código de asociación geométrica habitualmente al menos las coordenadas de un bloque de la representación de imagen (Im_1). Esto permite, por ejemplo, servicios de visionado o remasterización. Una película puede tomar, por ejemplo, película de versión anterior (o incluso un programa o juego, etc. ya procesado de acuerdo con los principios presentes), y dejar a graduadores hacer un nuevo análisis de las imágenes. Pueden a continuación guardar los valores de gris de diferenciador de regiones, y nuevas funciones de mapeo, etc., por ejemplo, en un servidor en la internet. El espectador puede elegir a continuación, por ejemplo, ver la película en el "nuevo grado del artista X" descargando los metadatos de ese servidor (potencialmente anulando cualquier demarcación y/o metadatos de gradación existentes). Esta opción podría, por ejemplo, ofrecerse a través de la interfaz de usuario tras el inicio de la película. Diversos diferenciadores de grises gTS permiten co-especificación de diversas funciones de procesamiento pretendidas, y esta estructura puede manejarse paramétricamente para fácil especificación de nuevo de, por ejemplo, mapeos colorimétricos de dispositivo de renderización final, o gradación de nuevo de los datos (que no necesitan cambiar el código Im_1) etc., por ejemplo, pueden no necesitarse tres códigos de gTs en el subrango de luminancias más oscuras para una primera estrategia de procesamiento, que puede ser solo un estiramiento lineal o no lineal en todas las luminancias entre gTs1 y gTs3, pero una segunda especificación gTS2 de una región intermedia puede usarse en estrategias de mapeo más complicadas. Por ejemplo, la visualización de lado de renderización puede elegir procesar las luminancias entre gTS2 y gTS3 proporcionando buen contraste visual, pero casi recortar los valores por debajo de gTS2. Un transcodificador o aparato intermedio similar puede, por ejemplo, aplicar un recorte suave en las luminancias entre gTS1 y gTS2 que aún contienen alguna información de la captura original, aunque con poca precisión ya que esta serán regiones oscuras difícilmente visibles en la mayoría de visualizadores de todos modos, es decir, necesitando menos calidad de codificación. El creador ha usado de esta manera gTS2 para especificar información semántica adicional acerca de la escena representada, a saber qué partes más oscuras de la imagen son menos relevantes. Estructuras separadas pueden ser más complejas que metadatos intercalados con los bloques de datos de píxeles y manipularse más libremente.

Ventajosamente el codificador de imagen (549) se dispone por que codifica un primer valor reservado de un valor de

gris de diferenciador de regiones en la señal de imagen de salida (S(Im_1, MET(gTS)), indicando que, para al menos una región geométrica de la representación de imagen (Im_1), encontrarse de acuerdo con una dirección de exploración a través de la imagen más allá de una ubicación identificable con el primer valor reservado, se realiza una transformación desde los valores de píxeles como se codifica en la representación de imagen (Im_1) a valores de píxel en una segunda representación de imagen (IM_RC_HDR) de acuerdo con un algoritmo predefinido.

Valores especiales para el valor de gris de diferenciador de regiones, como, por ejemplo, "0" o "-1" (claramente no siendo una luma válida en el rango de [0,255]) puede indicar que la siguiente región de la escena tiene que tratarse de forma diferente. Por ejemplo en una codificación, el decodificador puede referirse a una parte muy diferente de la señal de imagen (por ejemplo, un sector diferente de una memoria extraíble conectada), que ahora tiene que consultarse para obtener la señal de salida final (por ejemplo, algunas imágenes pueden codificarse de acuerdo con alguna tecnología de dos capas para alguna razón, como, por ejemplo, características de señal diferentes, u origen, etc.). En ese caso el codificador puede, por ejemplo, copiar un bloque de este tipo de ese segundo sector de memoria a la posición local, por ejemplo, en Im_1 potencialmente antes de hacer una transformación adicional tras la misma, o como alternativa como valores de luma finales. Cuando se procesa la imagen, las lumas de salida podrían obtenerse parcialmente aplicando un algoritmo de renderización de gráficos por ordenador. O un código de este tipo puede indicar que una transformación de imagen adicionalmente tiene que aplicarse para cambiar las lumas locales de píxeles o aspecto de textura. La región podría ser cualquier cosa con la condición de que la trayectoria de exploración (trayendo los algoritmos a alguna ubicación de inicio (x,y) en la imagen, es decir que es la ubicación identificable) se complementa mediante algunos metadatos adicionales que especifican la región, por ejemplo, puede ser una elipse que comienza o tiene su centro en una posición desplazada de (x,y). Habitualmente, sin embargo las realizaciones se usarán ventajosamente en un sistema basado en bloques, en cuyo caso, por ejemplo, (algunos de) los sucesivos bloques de 16x16 píxeles son la región geométrica.

Ventajosamente el codificador de imagen (549) se dispone por que codifica un segundo valor reservado (gTR) de un valor de gris de diferenciador de regiones en la señal de imagen de salida (S(Im_1, MET(gTS)), indicando que para al menos una imagen sucesiva, un visualizador debería renderizar la misma con una luminancia de salida máxima por debajo de un valor predeterminado. Por ejemplo, un valor 255, o 260 puede indicar que una parte de una imagen, o varias imágenes sucesivas, tienen que renderizarse con brillo decreciente para ahorrar potencia.

Ventajosamente el codificador de imagen (549) tiene la unidad de determinación de mapeo de luma (553) dispuesta para determinar varios mapeos de luma (TOM) diferentes para al menos uno del primer y el segundo objeto a través de reglas de enlace de transformación, o se dispone para indicar con un indicador de procesamiento (PROC_IND) que varios mapeos de luma diferentes pueden usarse para transformar los colores de píxel de al menos uno del primer y el segundo objeto a una nueva representación de color de la segunda representación de imagen (IM_RC_HDR). Porque ahora los diversos objetos (de diferente brillo) relevantes son fácilmente identificables en la escena como se codifica en cualquier representación de imagen, es también fácil transformar los mismos de cualquier forma deseada. Por ejemplo, podrían aplicarse varias estrategias de transformación de color diferentes a digamos un objeto altamente brillante, para varios visualizadores de renderización diferentes pretendidos, o iluminaciones ambientales del entorno de visión, o configuraciones de preferencias de usuarios, etc., por ejemplo, algunos visualizadores con alto brillo máximo, es decir, capacidad de nivel alto en la renderización de subregiones más brillantes de la imagen pueden usar un mapeo final cercano a o inspirado por una primera estrategia que tiene un aspecto de contraste para las regiones más brillantes según se define mediante una primera estrategia de mapeo, mientras los visualizadores de menor brillo máximo de menor calidad pueden seguir exactamente o aproximadamente una segunda estrategia de mapeo que tiene un efecto decreciente en las distancias de interluminancia de al menos algunos de los píxeles de una región brillante de este tipo. Y estas transformaciones podrían co-codificarse (parcialmente) con o en la señal de imagen, o (parcialmente) dejarse a cualquier lado de recepción (ya sea final o intermedio), en este segundo caso podría ser útil si la señal de imagen contiene algunas indicaciones aproximadas de qué clases de transformaciones son o viceversa no son deseable, etc. Obsérvese que dependiendo del uso adicional, uno o más de los mapeos pueden especificar transformaciones a seguirse exactamente, frente a transformaciones que son una indicación aproximada de qué debería hacer el visualizador de renderización final. El primer caso se producirá habitualmente, por ejemplo, en el caso de que el mapeo codifique realmente alguna gradación precisa (como, por ejemplo, un grado maestro de una codificación de contenedor LDR de 8 bits del mismo), y el segundo caso puede aplicarse cuando la transformación es una transformación adicional que indica cómo pueden optimizarse adicionalmente los datos de luma de píxeles maestros para varias clases de visualizadores. Por ejemplo un visualizador de menor brillo máximo puede estudiar la curva funcional de esta estrategia de recorte suave (que puede especificarse entre varios valores semánticos importantes gTS) y a continuación usar un mapeo de tono definitivo que mantiene aproximadamente el aspecto visual prescrito.

En el lado de recepción se puede construir una tecnología en gran parte reflejada del lado de codificador, siendo un decodificador de imagen (605) para decodificar una representación de imagen codificada (Im_1, MET) de una escena de alto rango dinámico, que comprende las características como se definen en la reivindicación 8.

El al menos un valor de gris de diferenciador de regiones gTS puede usarse a continuación para procesamiento de imagen adicional, tal como, por ejemplo, la determinación del mapeo de color óptimo final para el visualizador de renderización dado y entorno. Por tanto, nuestro método es una buena forma de vincular la codificación de color

- independiente de visualizador original y la codificación de color dependiente de visualizador final, que puede tener como propósito, por ejemplo, que debería renderizar colores en el entorno de visualización de visualizador aproximadamente como le habrían parecido a un espectador humano en la escena original. La codificación de imagen real puede ser muy diferente de la misma (ya que habitualmente codificamos la mismas con referencia ya a algún visualizador de referencia realista, que sin embargo aún puede ser muy diferente de la situación de renderización real: por ejemplo, imagen HDR maestra se codificó para condiciones de visualización doméstica de ambiente relativamente oscuro, y esa televisión doméstica a continuación se ajusta para las condiciones finales en cierto modo más luminosas; sin embargo: mucha de la complejidad ya se hace en el grado maestro hacia uno o más visualizadores de visualización de referencia realistas, dejando una estrategia de transformación de color final más simple al visualizador), sin embargo, ya que normalmente no habrá inversión del orden de luminancias de píxel, una buena forma de caracterizar la escena como se representa adicionalmente, y permitir sostenibilidad de situación de visualizador fácil es dividiendo la misma semánticamente en subpartes de luminancia/luma, especialmente las que habitualmente serán importantes y altamente variables en cuanto a su aspecto en varios escenarios de visualizador, como, por ejemplo, las regiones más oscuras o más brillantes de la imagen. Obsérvese que podemos usar la palabra luma para especificar todas las operaciones matemáticas como, por ejemplo, segmentaciones, ya que una luma de este tipo se relacionará con una luminancia real (por ejemplo, cuando la imagen se emite en el visualizador de referencia) a través de alguna estrategia de mapeo de codificación, que es una generalización -potencialmente discontinua- de un mapeo de gamma como gamma 2,2.
- 5
10
15
- 20 Ventajosamente este decodificador de imagen (605) comprende una unidad de segmentación de imagen (606) dispuesta para usar el valor de gris de diferenciador de regiones (gTS) para obtener un segmento de menor luma y un segmento de mayor luma en la imagen decodificada (IM_INTRM), es decir hace la separación de entendimiento de imagen basándose en los valores de gris de diferenciador de regiones, de modo que procesamiento posterior como, por ejemplo, procesamiento de ruido optimizado puede hacerse de forma diferente para regiones que se renderizan de forma diferente finalmente (con, por ejemplo, menos visibilidad del ruido en las partes más oscuras).
- 25
- 30 Ventajosamente el decodificador de imagen (605) comprende una unidad de transformación de color de píxel (609), dispuesta para aplicar una primera transformación de color (PD_BLCK(i,j)) que transforma al menos valores de luma de los colores de píxel a píxeles de, por ejemplo, la imagen HDR maestra recuperada en el segmento de menor luma, y dispuesta para aplicar una primera transformación de color (PM_BLCK(i,j)) transformando al menos valores de luma de los colores de píxel a píxeles en el segmento de mayor luma. Por tanto, se puede determinar, por ejemplo, una instantánea de accionamiento óptimo para renderizarse en un visualizador de mayor rango dinámico (bajo y alto y menor y mayor se entenderá por un lector experto para referirse entre sí, por ejemplo, si una codificación de color de píxeles de imagen es para un visualizador de referencia de 350 nits, transformando el mismo a una representación pretendida para un visualizador de referencia de 2000 nits, significa que esta segunda imagen es para mayor brillo o, dicho de forma diferente, mayor rango dinámico que la imagen original). Una separación de este tipo significa mucha mayor calidad aunque codificación simple. Si se tuviera que codificar toda la imagen con una única estrategia, se puede llegar únicamente a un aspecto aproximado promediando todas las clases de errores (por ejemplo, una cara tiene que ser brillante, pero a continuación el brillo del sótano oscuro es demasiado alto, de forma que oscurecemos la cara en cierto modo por debajo de lo ideal, y el sótano es únicamente un poco más brillante). Ahora sin embargo podemos, por ejemplo, oscurecer el sótano según deseemos, y a continuación corregir localmente la cara definiendo la mismas con umbrales y una estrategia de actualización. También, esta definición parcial hace fácil cambiar solo algo de los mapeos. Por ejemplo, a través de varias imágenes de una toma de la escena de sótano, debido cambios a de luz y/o movimiento de cámara, el PM_BLCK(i,j) puede permanecer adecuado para toda la escena, aunque la captura (o aspecto necesario) de las partes más oscuras puede cambiar a medida que vamos a través de las sucesivas instantáneas de la toma. Podemos a continuación cargar una función de PD_BLCK(i,j) diferente después de, por ejemplo, la quinta imagen de la toma, contrarrestando que esa esquina oscura se ha vuelto de ahora en adelante en cierto modo más brillante, y necesita una estrategia de mapeo que contrarresta-oscorece la misma, también por supuesto usando la forma funcional apropiada de PD_BLCK(i,j) para tratar la visibilidad textual, etc.
- 35
40
45
50
- 55 Ventajosamente el decodificador de imagen (605) se dispone para aplicar una transformación de estrategia de coloración específica a los colores de píxel de al menos uno del primer y el segundo objeto si el desformateador (607) extrae un valor de gris de diferenciador de regiones (gTS) de un valor reservado, tal como, por ejemplo, un valor de 0 o 255. De nuevo, estos valores reservados cuando se detectan en cualquier sitio en la señal introducida pueden usarse para revertir inmediatamente a cualquier estrategia de procesamiento de repliegue. Habitualmente, detalles adicionales estarán disponibles en ese repliegue para aplicar (aunque no necesariamente, ya que el receptor puede solo realizar cualquier cosa por sí mismo basándose, por ejemplo, en análisis de imágenes). Por ejemplo, si la señal de imagen viene almacenada en una memoria, puede haber un sector de estrategias de repliegue sucesivas (por ejemplo, código algorítmico que define métodos de procesamiento de imagen, y sus datos requeridos), y a continuación cada vez que se detecta un código reservado de repliegue especial, el aparato de procesamiento de imagen de recepción salta al siguiente método de repliegue para aplicar el mismo. O los códigos pueden referirse a qué repliegue aplicar (potencialmente muchas veces), por ejemplo, 260 indica que debería usarse el primer algoritmo almacenado, 261 el segundo, etc.
- 60
- 65 Ventajosamente el decodificador de imagen (605) comprende una unidad de determinación de transformación (610) dispuesta para seleccionar una estrategia de transformación de color de píxel desde una fuente de memoria no

asociada con ninguno de los datos de la representación de imagen codificada (Im_1, MET). De esta manera el decodificador de lado de recepción tiene más versatilidad para decidir qué va a usar para transformar las lumas de píxeles. Por ejemplo puede tomar funciones de su propia memoria y decidir, por ejemplo, dependiendo de las propiedades de un objeto identificado, tal como su luma promedio. O podría tomar funciones a través de una conexión de red, determinada potencialmente en tiempo de ejecución por un servidor. La señal aún puede guiar parcialmente esto, especificando que es deseable aplicar (cualquier) mapeo de oscurecimiento, por ejemplo, (es decir, una transformación que tiene un resultado visual que el objeto parece más oscuro de alguna forma, por ejemplo, brillo promedio en conjunto con una modificación de contraste, y/o un área aumentada en el objeto de muy oscuro, por ejemplo, recortado a píxeles negros, etc.), en cuyo caso el lado de renderización preferentemente no debería aplicar un mapeo que abrillanta el objeto muy oscuro (teniendo en cuenta la visibilidad debido una iluminación ambiente, etc. por supuesto). Finalmente el lado de recepción, ya sea bajo el control específico del espectador o no, puede decidir por supuesto cumplir (parcialmente) con estas guías de co-codificación deseada o ignorar y atravesar las mismas. Habitualmente aunque la codificación de imagen (por ejemplo, en disco en el que se codifica) puede, por ejemplo, prescribir que la transformación no debería ignorarse ni incluso relajarse, sino que debería seguirse estrictamente, o viceversa no seguirse estrictamente.

Ventajosamente decodificador de imagen (605) se caracteriza por que unidad de determinación de transformación (610) se dispone para determinar la estrategia de transformación de color de píxel sobre la base de al menos un parámetro del entorno de renderización, tal como una característica del visualizador, o un nivel de iluminación ambiente, o el patrón de colores como se ven reflejados en la pantalla frontal del visualizador por una cámara, etc. Por tanto de nuevo, el aparato de lado de recepción puede al menos parcialmente optimizar los mapeos basándose en información importante únicamente disponible definitivamente en su lado. Un creador de contenido puede especificar que sus mapeos se usen bajo la suposición de que un cierto visualizador y entorno de visualización (por ejemplo, la mayoría de luces del salón apagadas, con únicamente alguna iluminación atmosférica, que de hecho puede realizarse aproximadamente en realidad con el espectador teniendo, por ejemplo, una lámpara de colores vivos en el suelo al lado del espectador), pero finalmente el lado de renderización puede cambiar los mismos, ya sea incluso un ajuste menor (que es el caso ideal). Aunque una cantidad de precisión de este tipo normalmente no se necesita, el creador de contenido podría especificar en la señal que, por ejemplo, PD_BLACK(i+2,j) se pretendía para el caso de que hubiera una luminancia de digamos 1 nit alrededor del visualizador, en cuyo caso si el visualizador de renderización mide 2 nits, puede decidir cambiar ligeramente la inclinación de PD_BLACK(i+2,j). En cualquier caso esto puede ser información útil para el procesamiento de algoritmos en el lado de recepción.

Las realizaciones descritas pueden realizarse de diversas formas, por ejemplo, mediante un método de codificación de imagen para codificar una imagen de una escena de alto rango dinámico, que comprende las etapas como se definen en la reivindicación 12.

O mediante un método de decodificación de imagen para decodificar una representación de imagen codificada (Im_1, MET) de una escena de alto rango dinámico, que comprende las etapas como se definen en la reivindicación 13.

O como un programa informático que comprende código de software que habilita un procesamiento para ejecutar cualquiera de los métodos que corresponden a las realizaciones descritas, ese software puede traerse en un disco u otro producto tangible, o descargarse a través de una red, etc.

Habitualmente el conocimiento codificado acerca de la escena representada viajará desde un locus/aparato a otro (ya sean unidades dentro de un mismo aparato o sistema de consumidor de aparatos conectados al mismo sitio como, por ejemplo, una caja de recepción o procesamiento de imagen y una televisión o visualizador conectado a través de, por ejemplo, HDMI, o servicios ejecutándose en aparatos en diferentes países) es decir por medio de una señal de imagen que codifica los colores de regiones de una escena de alto rango dinámico, que comprende palabras de código de N bits que codifican al menos las luminancias de los colores de regiones, y un valor de gris de diferenciador de regiones (gTS), indicando en el sistema de código usado para codificar las palabras de código de N bits que codifica al menos las luminancias de los colores de regiones, una demarcación entre al menos una región geométrica de píxeles de mayor luminancia en la escena de alto rango dinámico o mayores valores de las palabras de código de N bits que codifican los mismos, y al menos una región geométrica de píxeles de menor luminancia en la escena de alto rango dinámico o menores valores de las palabras de código de N bits que codifican los mismos. El sistema de código es la representación matemática técnica que define una derivada desde una luminancia de escena (a través de captura de cámara) y finalmente una luminancia a renderizar, habitualmente a través de una cantidad física llamada luma, que se define en un eje, y habitualmente con una palabra de código digital que cubre la extensión del eje (por ejemplo, entre 00000000 y 11111111), o un número flotante entre 0,0 y 1,0, y con una función de asignación (habitualmente una función gamma) que mapea tales luminancias no linealmente a luma. Habitualmente puede haber información adicional asociada con el sistema de código, tal como con qué luminancia máxima a renderizar corresponde el valor de código máximo. Cuando hablamos de esta señal, queremos decir que las propiedades especificadas se contienen de alguna forma en la señal, pero pueden contenerse de cualquier forma trasladada. Por ejemplo, algunos datos podrían fusionarse o dividirse, y estructurarse de cualquier forma. También puede haber transformaciones a otros códigos, tal como, por ejemplo, una modulación, o una codificación redundante para compensar potenciales daños de errores de bits, etc.

La imagen HDR puede codificarse (por ejemplo, como una imagen de textura de 8 bits LDR Im_1 llamada contenedor LDR, más metadatos para mapear que una reconstrucción del grado de HDR maestro por al menos un mapeo de tono global) en una memoria tal como una memoria extraíble, tal como, por ejemplo, un disco Blu-ray que almacena una señal de este tipo.

5 Realmente, las realizaciones de la invención pueden usarse en muchas realizaciones técnicas, escenarios o usos, tal como en un sistema de distribución de video a través de cualquier tecnología de red, que emplea cualquier codificador de imagen, decodificador de imagen, método, señal de imagen, u otro producto o implementación de cualquiera de las realizaciones descritas, o cualquier uso de ese sistema de distribución de video.

10 Por supuesto son posibles muchas variantes adicionales de las realizaciones descritas a continuación, y el experto entiende que pueden, por ejemplo, realizarse en diferentes aparatos en diferentes regiones geométricas del mundo, aplicando su funcionalidad parcial en diferentes momentos en tiempo, o varios tiempos después de cada uno, en diversos escenarios de uso comercial, etc.

15 Breve descripción de los dibujos

Estos y otros aspectos del método y aparato de acuerdo con la invención serán evidentes a partir de y se aclararán con referencia a las implementaciones y realizaciones descritas en lo sucesivo, y con referencia a los dibujos adjuntos, que sirven simplemente como ilustraciones específicas no limitantes y que ilustran el concepto más general, y en el que las líneas discontinuas se usan para indicar que un componente es opcional, los componentes no en línea discontinua no son necesariamente esenciales. También pueden usarse guiones para indicar que elementos, que se han explicado que son esenciales, se ocultan en el interior de un objeto, o para cosas intangibles tal como, por ejemplo, selecciones de objetos/regiones (y cómo pueden mostrarse en un visualizador).

25 En los dibujos:

La Figura 1 ilustra esquemáticamente diversas representaciones de una escena original de alto rango dinámico, ya que se renderizarán en diferentes escenarios, a saber: la Figura 1a muestra las luminancias de salida renderizadas absolutas comparadas entre sí para un visualizador de alto rango dinámico actual, un visualizador de cine, un visualizador de bajo rango dinámico y un visualizador portátil usado en exteriores, y la Figura 1b muestra las renderizaciones en un eje de aspecto universal, cuyo sistema de referencia absoluto se define mediante un espectador humano estándar;

30 La Figura 2 (es decir, la Figura 2a + la Figura 2b) ilustra esquemáticamente cómo diversas transformaciones de subcolores para transformar entre dos representaciones de color, definiendo ambas una misma vista de imagen en una escena, se aplicará a al menos las lumas de píxeles de diversos objetos de muy diferente luminancia (o luminosidad), cayendo en varios bloques de una descomposición de bloque de una representación de imagen;

35 La Figura 3 ilustra esquemáticamente una forma para codificar algunos metadatos adicionales de acuerdo con algunas realizaciones en una definición de señal de imagen particular, en particular cómo codificar valores de gris de diferenciador de regiones antes de bloques de color de píxel a los que son aplicables;

La Figura 4 ilustra esquemáticamente cómo un lado de recepción puede obtener segmentos de muy diferente luminancia o luminosidad en la imagen basándose en los valores de gris de diferenciador de regiones;

40 La Figura 5 ilustra esquemáticamente un sistema de lado de codificación, que puede operarse por un graduador de colores, una realización ilustrativa de un codificador que corresponde a nuestras descripciones de la invención;

45 La Figura 6 ilustra esquemáticamente un sistema de lado de decodificación, que puede ser, por ejemplo, un sistema de visualizador doméstico de consumidor que comprende tales aparatos como una televisión principal y un visualizador de imágenes portátil y un aparato de procesamiento de imagen tal como un ordenador central que gestiona la distribución o procesamiento óptimo de todos los videos para los diferentes visualizadores;

50 La Figura 7 ilustra esquemáticamente cómo el diseño de las regiones a las que se mapean los rangos de luminancia (o luma) pueden seleccionarse bien para mitigar problemas como errores de compresión; y

La Figura 8 ilustra esquemáticamente cómo puede usarse nuestro sistema en un escenario en el que colores de píxeles u objetos tienen que mapearse a colores óptimos para un número de visualizadores con características técnicas considerablemente variables.

55 Descripción detallada de los dibujos

La Figura 1 (es decir, La Figura 1a y la Figura 1b) muestra esquemáticamente cómo una escena HDR original (Orig_SCN) puede representarse óptimamente en 4 tipos de visualizador (3 típicos y un hipotético para ilustrar mejor el punto, a saber un libro electrónico de bajo contraste con iluminación del sol, teniendo únicamente un rango pequeño R_ERDR_OUTS de luminancias de salida reproducibles), y cómo una tecnología de codificación de imágenes debería acomodarse para tal. Enfatizamos que se necesita dividir conceptualmente las ideas con respecto a *renderización* final de una escena (es decir, las luminancias a emitir físicamente por un visualizador particular) a partir de la *codificación* de lumas de objeto de imagen. Esa es una filosofía tecnológica diferentes de tecnologías de imágenes de televisión clásicas como MPEG2, que siempre han equiparado estos dos correspondientes espacios de color, de modo que, por ejemplo, una señal codificada de gamma 2,2 puede aplicarse directamente a un visualizador (estándar), proporcionando (aproximadamente) la salida renderizada correcta (determinada por lado de estudio de una manera

calibrada). Esto es útil únicamente si se tiene una cadena cerrada, es decir, calibrada para un escenario particular, pero la historia se rompe si queremos tener otro contenido como en imágenes de alto rango dinámico (HDR) particulares, y/o diversos visualizadores y/o entornos de visualización con fundamentalmente diferentes características en las que renderizar estas señales. Incluso, aún gustaría de forma similar la simplicidad de tener únicamente una (o al menos unas pocas) señal o señales de codificación de imagen, y no imágenes codificadas diferentes para cada escenario (aunque pueden empaquetarse de nuevo (por ejemplo, transcodificarse, transformarse adicionalmente transformación de colore, etc.) y transmitirse a través de diferentes canales técnicos), que de otra manera significaría que un graduador de Hollywood u otro graduador tendría que hacer, por ejemplo, 20 gradaciones en lugar de 1 o 2 como anteriormente (por ejemplo, un grado de película maestro y un grado de DVD).

Definir imágenes HDR o una tecnología de formación de imágenes HDR puede conducir a debate. Por supuesto, no es simplemente el número de bits que es el criterio, ya que si, por ejemplo, la longitud de palabra máxima (por ejemplo, 2^8 frente a 2^{10}) se usa para un cierto blanco (por ejemplo, 500 nits), a continuación la diferencia es mayoritariamente o parcialmente solo una de precisión (realmente, visualizadores con relaciones de contraste atribuidas de 1.000.000:1 pueden incluso no renderizar discriminadamente el menor de estos códigos, y en una codificación de señales de gamma 2,2 tales negros profundos pueden no codificarse tampoco, a no ser que el visualizador haga alguna transformación de ennegrecimiento impresionante en los negros).

La definición normal de una escena de alto rango dinámico es la luminancia máxima dividida por la luminancia mínima. Esa puede ser una buena definición desde un punto de vista de hardware, por ejemplo, para un visualizador de renderización.

Por ejemplo, en la escena original, determina cuáles deberían ser las capacidades de un sensor de formación de imágenes de cámara, también si este opera con, por ejemplo, tecnología de múltiple exposición, porque cualquier cosa que no pueda grabarse fielmente o bien se recorta a blanco o bien a negro (y por supuesto también existe el redondeo y ruido). También es una buena forma indicar lo que puede renderizar físicamente un visualizador, por supuesto con la condición de que se haga de una forma justa, incluyendo, por ejemplo, dispersión en la placa frontal de cristal de la luz generada por el visualizador, así como reflejos del entorno (por ejemplo, la camiseta blanca del espectador enfrente de la TV). Toda clase de dispersión de luz y reflejos son las razones por las que el rango dinámico visto o capturado es a menudo mejor que los números de comercialización citados, ya sea porque la luz se fuga a través de toda clase de trayectorias desde puntos más luminosos de la escena hacia puntos más oscuros (durante la iluminación de la escena, a no ser que se construya con cuidado y gestionen las áreas de sombra), trayectorias espurias en cámaras (por ejemplo, neblina de lente, o reflejos corporales), o problemas de ambiente de visualización (por ejemplo, luz de visualizador o ambiental dispersándose en la placa frontal del visualizador, o reflejos dentro del visualizador que entra en el homogeneizador de luz, etc.) hasta el propio ojo del espectador (sin embargo, aunque el espectador puede empezar a perder precisión de discriminación de oscuros cuando tiene fuentes de luz intensas en su campo visual, especialmente cuando está cerca de las regiones oscuras, podríamos ignorar este factor ya que un visualizador puede necesitar idealmente ser mejor que el espectador, y al menos una codificación de imagen debería ser mejor, ya que no sabemos por adelantado cómo procesará la misma el lado de recepción e influenciará la visibilidad de las regiones de la imagen). Por lo tanto con una definición de relación de contraste de este tipo, se debería usar como el nivel mínimo lo que es de hecho finalmente discriminable en el ojo (ruido dado, etc.) y no, por ejemplo, el valor teórico que un LED apagado proporciona (casi) cero luminancia de salida (por lo tanto, normas imponen, por ejemplo, patrones de tablero de ajedrez para medir relaciones de contraste más justas), ya que nunca existe una situación de cero luz.

Una relación de luminancia, sin embargo, no es un buen criterio de rango dinámico para la CODIFICACIÓN de imágenes HDR. Lo que tiene que codificarse no es tanto un problema de lo que se puede renderizar, sino lo que está en una escena y lo que puede percibirse al menos teóricamente, es decir la señal de imagen necesita contener exacta o aproximadamente esos datos necesarios para poder recrear el aspecto deseado, y en todos los entornos de visualizador que se esperan que estén renderizando la imagen, incluso quizás en los mejores visualizadores en un futuro lejano (brillando directamente en el ojo, por ejemplo).

Por ejemplo, solo especificar una relación de contraste no tiene en cuenta el hecho de que en un entorno oscuro como un cine, el sistema necesita más contraste para ver la misma escena que aparece en la imagen (mientras que un escalado multiplicativo puro en el mínimo y máximo resultaría en la misma relación de contraste). Contrastes son de hecho también un fenómeno local, ya que un objeto relativamente iluminado puede hacerse que se perciba mucho más oscuro si se rodea por objetos más iluminados (contraste espacial). De hecho, psicológicamente el espectador comienza a analizar la instantánea, e identifica colores que piensa que son negros, blancos, por encima de blancos, etc. Y el espectador puede considerar que algo es negro o blanco, hasta que percibe un negro incluso más oscuro o blanco más brillante. Así pues, cómo de "dinámica" parece una instantánea no es únicamente una función del "negro" y el "blanco" en la escena, sino también una función de otras medidas de contraste locales que pueden definirse sobre la base de asignación de valores de gris (por ejemplo, se puede crear un aspecto diferente aumentando la distancia de luminancia de diferentes grises en texturas -haciendo texturas rocosas más ásperas, por ejemplo-, o haciendo sombras más oscuras, o incluso jugar con las interrelaciones entre nitidez y contraste). Se puede por lo tanto imaginar que si se quiere proporcionar a una cara con un aspecto diferente (más suave, más contraste, más arrugada, etc.), que los códigos que definen las texturas faciales tienen que permitir una operación de este tipo, es decir, por ejemplo, si se tuvieran únicamente dos valores de gris definiendo la textura facial, cambiar el contraste facial sería una operación

muy difícil. Tal irregularidad de colores faciales puede ser un problema de varias tecnologías de codificación de imágenes actuales.

Para poner el problema incluso más claramente, mostrar un ejemplo que desde un punto de vista de codificación el rango dinámico no es solo acerca del negro más oscuro y el blanco más brillante, sino acerca de qué exactamente hay en la escena representada, un dibujo en blanco y negro (es decir, teniendo únicamente dos diferentes valores de gris) puede RENDERIZARSE en un visualizador HDR con un blanco de 5000 nits y negro de 0,5 nits (es decir, con un alto rango dinámico de luminancia), pero ¿podríamos llamar realmente a esto una imagen HDR? Podemos incluso plantear preguntas adicionales, como si querríamos visualizar una señal simple de este tipo con las características de máximo blanco (blanco pico) respectivamente negro del visualizador de todos modos. ¿No sería antinatural, o al menos innecesario, y no digamos si se necesitase codificar directamente esos valores así (por ejemplo, con código 0 y 10000 y no solo como, por ejemplo, 0 y 2)? De hecho, por ejemplo, cuando se gradúan áreas más blancas en general un artefacto de aspecto que puede empezar a suceder es que la región blanca con textura comienza a parecer de tiza (como si se dibuja usando tiza) que es diferente de la textura física real que debería tener la región. Podríamos llegar a la pregunta: ¿qué es "negro" y "blanco" de nuevo? De hecho, bajo una iluminación soleada, suponiendo que nuestro ejemplo sería un dibujo en blanco y negro, por ejemplo, el blanco puede tener una luminancia de escena real de 5000 nits, pero bajo una iluminación diferente podría solo tener también 50 nits. Y a no ser que se protejan mucho las regiones negras de la iluminación de escena, estarían normalmente en algún sitio entre el 1 % del blanco, y no 1/10000ésimo. Por tanto, ignorando que imágenes renderizadas con más contraste pueden tener en algún grado un aspecto preferido, probablemente nos gustaría mostrar esa instantánea en blanco y negro con, por ejemplo, alto brillo aunque aproximadamente un rango de luminancia de 100:1 en el subrango de luminancia alta de un visualizador HDR para crear ese aspecto de dibujo iluminado por el sol. De otra manera de todos modos, arriesgamos sin la instantánea renderizada no parece extraña que el ojo descuenta algo de la diferencia en luminancia de renderización, por tanto con un graduador siempre querríamos usar óptimamente el rango dinámico de visualizador disponible dado que está presente como contenido de la imagen, y en vista de efectos temporales incluso imágenes anteriores y sucesivas. También obsérvese que aunque imágenes brumosas se consideran convencionalmente como bajo rango dinámico, una imagen de este tipo también con luces brillantes en la misma necesitaría al menos mapearse a una subregión alta de un eje de luminancia de colores renderizables de visualizador.

Nuestra filosofía de codificación es que una codificación necesita tener en cuenta ambos de estos factores, a saber por una parte cómo de dinámica finalmente se renderizará una imagen habitualmente, pero por otra parte qué clases de objetos más o menos brillantes contendrá la escena representada. Por tanto, para nuestros propósitos (es decir, representación o codificación de imágenes) sería más preciso decir que una imagen HDR es una imagen que contiene: una cantidad suficiente de valores de gris a lo largo de un número de rangos suficientemente lejos a lo largo de un eje de aspecto de luminosidad (siendo la luminosidad una cantidad psicofísica que no debe confundirse con luminancia o luma codificada). Por lo tanto podemos explicar mejor las físicas y realizaciones técnicas requeridas de imágenes HDR con el concepto de "*rangos de aspectos anidados*", como se explica con la Figura 1.

En la Figura 1 vemos una escena (Orig_SCN) a capturar que tiene simultáneamente muchas luminancias oscuras y brillantes, es decir, detalles de textura significativos en un rango de luminancias tanto en áreas iluminadas oscura y brillantemente. Para la región/objeto brillante (BR_obj) hay un vitral, que tiene muchos colores bonitos que deberían codificarse y renderizarse de forma precisa. En el interior oscuro del edificio, hay una escalinata de madera oscura (DRK_obj) y un objeto incluso más oscuro (VDRK_obj). Es decir, un ser humano de pie en esa escena original vería muchas luminancias brillantes (y de hecho colores) en el vitral, y muchas diferentes luminancias oscuras en las diferentes regiones de sombras en la escalera. Cuando gira su cabeza, su procesamiento de retina y cerebral se adaptará al aspecto en el vitral, o viceversa, tratando de discriminar objetos de aspecto oscuro en las regiones más oscuras. Cómo de oscuro parece todo depende por supuesto de lo bien que los constructores de la escena hayan aislado las regiones más oscuras de las más brillantes, pero se puede imaginar, por ejemplo, el ejemplo de intentar mirar a través de un pequeño agujero de alcantarilla en el pavimento en un día muy soleado. Es decir las regiones "más oscuras" pueden variar de parecer gris oscuro a negro no discriminable y definitivo (o durante la noche negro más grisáceo no discriminable). En el visualizador de renderización, nos gustaría crear, dadas capacidades, al menos una en cierto modo experiencia similar (por ejemplo, colores algo oscuros no discriminables, con una luminancia lo suficientemente baja que al menos parecen razonablemente negruzcos), es decir equilibrando tanto el hecho de que un número significativo de luminancias de salida por subrango de luminancias aún renderizan las texturas de objeto de todos los objetos tantos brillantes como oscuros con buena calidad visible, frente a que el vitral debería aparecer significativamente más brillante que el promedio (dado el rango dinámico particular del visualizador esto puede ser más de una simulación usando efecto "ilusorio" psicovisual frente a una gran diferencia fotométrica real para visualizadores de alta brillo), y la escalinata más oscura que el promedio (siendo el promedio, por ejemplo, el nivel de grises del 18 % del ambiente de la sala iluminado). Independientemente de cómo el visualizador hará esto óptimamente, la codificación de imagen debería al menos contener toda la información, y preferentemente de una forma fácilmente gestionable.

Se puede capturar y codificar ahora esta escena, con una única señal de datos (por ejemplo, 0-1023, con una función gamma fija para mapear luminancias de entrada o salida a códigos; es decir, por ejemplo, si la función gamma define luminancias de salida, se puede convertir primero la imagen capturada a N nits de referencia, por ejemplo, visualizador de 16 bits (lineal o de otra manera, por ejemplo, con brillo máximo de 4000 nits), y a continuación codificar esos

"valores de nueva escena" a, por ejemplo, una representación de 10 bits, con la intención de que en el visualizador de referencia se produzca una reproducción exacta y, por ejemplo, un visualizador de 2000 nits aproximará el aspecto). O se pueden optimizar diversas señales codificadas para diferentes escenarios, por ejemplo, aplicar una diferente función gamma para una señal de cine, para compensar el comportamiento de entorno oscuro del sistema visual humano. Pero idealmente el procesamiento principal -que puede ser muy complejo en vista del comportamiento altamente local y no lineal de la visión humana- debería estar ya presente en la una o más imágenes graduadas codificadas (en codificaciones de una sola imagen HDR codificándose la gradación LDR simultáneamente dentro de la gradación de maestro HDR, usando el concepto de contenedor LDR, es decir, por el principio que se puede obtener de nuevo la imagen HDR maestra desde esa gradación codificada LDR (por ejemplo, una imagen MPEG-AVC clásicamente codificada), invirtiendo la estrategia de mapeado de color usada para hacer la misma desde la gradación maestra HDR, usando metadatos co-codificados que codifican ese mapeo; pero por supuesto la codificación de imagen puede contener varias gradaciones, ya sea con varias funciones de mapeo o al menos imágenes de píxeles adicionales parciales), es decir, su aspecto ya se determina aproximadamente correctamente para un número de escenarios de visualizador típicos. En ese caso, la optimización de visualizador real creará con una operación relativamente simple el aspecto final aproximadamente correcto, por ejemplo, una función gamma simple final para aumentar contraste para visualización de ambiente más oscuro, etc.

En cualquier caso, final aspecto parecerá como se muestra en la Figura 1b, y las luminancias de salida fotométricamente medibles serán como en la Figura 1a. El primer escenario es la señal mostrándose en un visualizador HDR (como se ha dicho, o bien con su propio grado HDR optimizado con como mínimo procesamiento mínimo (por ejemplo, algunas especificaciones de hardware realmente como imitando comportamiento de tipo CRT con una compensación adicional para la física de válvula LCD) utilizable directamente para accionar el visualizador HDR, o una señal de accionamiento derivada a partir de una única imagen HDR/señal de video codificada). Este visualizador es capaz de un brillo máximo (blanco) de 5000 nits, y una luminancia de salida mínima de 0,5 nits. Obsérvese que el menor valor es una aproximación promedio, ya que variará críticamente con varios parámetros de ambiente. Incluso en un entorno controlado, las luces de seguridad de cine pueden filtrar luz a la pantalla, y así hace el factor impredecible de la gente encendiendo sus teléfonos móviles (aunque en general el efecto se limitará, pero especialmente en las luminancias más oscuras, puede afectar la renderización). En una casa normal situaciones de iluminación pueden ser variables.

Pero aún la cuestión es cómo un ser humano percibirá tales luminancias, ya que esto dependerá del estado de su sistema visual, entre otros, determinado por la iluminación de la sala, si puede a veces ver a través de una ventana exterior, etc. U espectador puede controlar este aspecto cambiando los ajustes de instantánea en su control remoto. En cualquier caso, el visualizador HDR tiene un subrango relativamente grande de valores brillantes disponibles para la renderización del vitral (es decir, se muestra relativamente grande, cubriendo una parte superior del rango R_D_HDR). Al mismo tiempo, la escalinata puede mostrarse suficientemente oscura, es decir muy por debajo de 50 nits. Se supone para nuestro ejemplo que esto tiene un impacto psicovisual de que esta escalera parece tanto más oscuras en comparación con la estimación visual de la luminosidad promedio (por ejemplo, 18 % de gris), pero también que aún pueden verse significativamente fácilmente los detalles de textura sobre los reflejos de iluminación ambiente del cristal frontal del visualizador (por ejemplo, un escenario en el que el espectador ha atenuado su iluminación ambiental a un nivel de visualización de película, y el gris promedio se determina principalmente por la televisión y su propio contenido de imagen). Este visualizador HDR (+ entorno de visualización) es tan bueno que puede incluso mostrar el objeto muy oscuro con luminancia de salida de visualizador aún más oscura y correspondiente luminosidad psicovisual.

Si ahora mostramos la misma señal en un proyector de cine digital en un cine (de nuevo, ya sea gamma óptimamente corregida o no), sabemos que esta renderización de cine no mostrará ningún blanco por encima de aproximadamente 50 nits, sin embargo, siendo un entorno oscuro, al menos las tomas más oscuras pueden mostrar luminancias por debajo de digamos 0,05 nits, es decir mucho más oscuro que la renderización de visualizador HDR de sala doméstica. Es decir, el rango de luminancia de salida de cine R_CIN se encuentra entre 0,05 y 50 nits. No podemos decir que el vitral que se asignará un menor subrango de mayores luminancias en R_CIN, parecerá igualmente oscuro que la escalera en el visualizador HDR que tiene aproximadamente idénticas luminancias de salida, ya que el espectador se ha adaptado a la sala de cine oscura, por lo tanto, ve luminancias de salida más bajas como (casi) blanco. Es decir, también, en el cine podemos tener relativamente alto rango dinámico, al menos entre instantáneas (y al menos puede codificarse si no en la película positiva o señal digital, entonces en el negativo maestro). Especialmente con algo de la simulación psicovisual, como, por ejemplo, reproducir en colores culturalmente establecidos como diurnos o nocturnos, también el espectador de cine aún tiene la solución que después de una escena de sótano oscuro alguien sale al sol (ya sea menos impresionante en el visualizador HDR).

El hecho de que el sistema visual humano se adapta puede verse mejor en la representación de aspecto psicosocial de la Figura 1b, en la que ponemos las diversas imágenes de salida renderizadas en un eje de aspecto de luminosidad (Appear_SCAL). Esto es de hecho lo que ve el cerebro (con todo su complejo procesamiento), pero aproximadamente podemos mapear el mismo a cómo se comportan los conos retinales (o al menos junto con conexiones de células ganglionares). De cualquier modo, en nuestra filosofía técnica esa complejidad puede manejarse por el graduador humano, ya que siempre debería, como le gusta al creador de contenido, estar a cargo del aspecto de su contenido. Vemos de hecho que la renderización en visualizador HDR de sala doméstica (DISPL_HDR) y la renderización de cine

(MOV_THTR) son razonablemente similares (al menos pueden simularse entornos relativamente algo oscuros, así como exteriores brillantes). Sin embargo, la renderización de cine no es capaz de mostrar tales regiones brillantes, al menos con precisión sin ninguna deformación de color (que se muestra mediante el en cierto modo pictograma más oscuro, moviéndose desde el hiper brillo a la región brillante del eje de aspecto). Nos gustaría enfatizar que este efecto se debe también a la renderización separada en un cine frente a en un visualizador HDR en casa, ya que si se renderiza simultáneamente en un visualizador HDR en un cine, la comparación se vuelve diferente de nuevo (ya que ahora las regiones brillantes en la pantalla de proyección relativamente tenue pueden compararse directamente con las del visualizador HDR). Sin embargo, estar en una oscuridad relativamente profunda, la renderización de cine puede simular escenas muy oscuras, como, por ejemplo, una escena nocturna, en la que el sol empieza a elevarse lentamente hacia el horizonte. Sentado en una sala de estar iluminada por un sol brillante nunca se puede tener ese aspecto. Cualquier entorno que tiene también regiones brillantes (por ejemplo, un visualizador HDR brillando brillantemente co-situado) destruirán en mayor o menor medida esa "ilusión" visual de una escena nocturna totalmente oscura. Incluso ignorando que los colores oscuros renderizados en el visualizador se encontrarán por debajo de la luminancia de reflexión del cristal frontal, todos los colores claros que entran en el ojo en grandes ángulos desde el entorno romperían la ilusión (que se ilustra mejor incluso con el ejemplo del libro electrónico). Por supuesto en principio se podría hacer una sala de estar incluso mucho más oscura que en el cine ya que en casa la seguridad no es problema, que significaría que a continuación el visualizador HDR también tiene mayores capacidades para un negro más profundo, pero normalmente a la gente en casa le gusta algún tipo de iluminación ambiente acogedora (en cualquier caso, una imagen codificada que atiende para todas las situaciones de renderización podría también optimizarse fácilmente por esa gente que no le gusta ver sus películas de miedo de la forma más aterradora en una sala de estar negro intenso, que significaría que las regiones más oscuras de la imagen necesitan codificarse tanto con suficiente precisión como estar accesibles fácilmente para el procesamiento de optimización colorimétrica). Obsérvese también que en entornos muy oscuros, el contraste de escena como se ve por el sistema visual humano puede degradarse severamente (es decir, como se vería la escena original), por tanto se puede necesitar simular las mismas (en, por ejemplo, un cine en el que ese efecto no es aún muy intenso) renderizando los objetos más oscuros con un gris oscuro un par de paradas por encima de negro intenso, y los objetos blancos con un gris claro un par de paradas por debajo de la luminosidad de zona de referencia de blanco.

Por tanto, existen regiones que pueden no poderse renderizar con precisión en cada posible visualizador, incluso aún querríamos codificar las mismas, ya que puede o no haber visualizadores que son capaces de renderizar las mismas (ya sea, por ejemplo, después de un abrillantamiento), proporcionando este ejemplo una región ultra oscura para el cine, y una región hiper brillante para algunos visualizadores HDR. Obsérvese que la región ultra oscura del sistema visual humano puede finalizar en algún sitio en el lado inferior de la adaptación visual humana, tal como, por ejemplo, para codificar una cueva muy tenuemente iluminada en la que algunas luces se filtran a través de una grieta en la distancia. Sin embargo, un nivel de este tipo es irrelevante para visualizar incluso en los cines (teóricamente) más oscuros, porque las partes más brillantes de la imagen/contenido de video no permitirán que el sistema visual se adapte óptimamente (nadie ve películas de cuevas en una cueva). Puede igualarse sin embargo con el nivel en el que el ojo empieza a ver ruidosa y borrosamente, tal como, por ejemplo, cuando uno entra en una sala oscura después de haber estado al sol. Una experiencia visual de este tipo es algo que se puede querer renderizar, porque transmite un nuevo nivel de calidad visual, como luces deslumbrantes en el lado brillante. Es decir, es el régimen en el que se equilibra lo que (solo) puede verse con lo que no puede verse. Pero el punto es que una renderización de entorno oscuro puede mostrar mejor el objeto muy oscuro, ya que puede renderizar el mismo por debajo de la región oscura del eje de aspecto, existían inicios de región ultra oscura.

Un tercer visualizador es un visualizador doméstico LDR (renderización DISPL_LDR), por ejemplo, una televisión "clásica" con digamos un brillo máximo presente de 300 nits (que para nuestra descripción supondremos que se comporta relativamente similar a visualizadores más antiguos de, por ejemplo, 100 nits de brillo máximo). Suponerlo puede mostrar en cierto modo menores negros profundos (por supuesto en los negros podrían ser similares al visualizador HDR, pero por explicación, digamos que tiene, por ejemplo, atenuación global en lugar de atenuación LED de 2D). También puede renderizar menos colores oscuros porque quizás en vista del menor brillo máximo necesita reservar un mayor subrango de su rango LDR R_D_LDR para luminancias más brillantes o intermedias, de forma que renderizará tanto la escalinata como el objeto muy oscuro con al menos visualmente aproximadamente los mismos grises oscuros. Realmente, reservará únicamente unos pocos niveles de luminancia para la escalinata, haciendo la misma con textura menos detallada, y el objeto muy oscuro habitualmente se recortará a negro (y quizás incluso inoportunamente invisible contra el recorte a partes negras de la escalinata). Otra propiedad típica del visualizador LDR es que no puede renderizar fielmente los objetos hiper brillantes, y habitualmente recortará (suave) los mismos a un rango muy pequeño de (casi) blancos, todo dependiendo entre otras de qué contraste se quiere para los rangos medios cerca de grises medios. Estrategias de recorte y aproximación pueden tener un fuerte impacto psicovisual, ya que el cerebro reconoce que está sucediendo algo especial.

Así que vemos que la renderización es realmente una asignación de luminancias (ajustadas por adaptación visual humana) (es decir, de hecho correspondientes luminosidades y brillos para el sistema visual humano) de la escena a diferentes subrangos del respectivo rango de luminancia renderizable por visualizador. Algunos visualizadores pueden renderizar únicamente una subparte del rango total que está anidado (desde al menos un lado) en el rango total, y algunos visualizadores pueden renderizar casi todos los aspectos relativamente fielmente. Es decir cuando se mapean a luminancias de salida o de hecho visualizador accionando valores de imágenes (es decir, para accionar, por ejemplo,

las válvulas LCD y algún accionamiento de retroiluminación), debe tomarse alguna aproximación, ligeramente cambiando el aspecto exacto de un objeto de escena o región, en un aspecto que es aún razonablemente similar, y si no es convincente, al menos aceptable. El ejemplo del libro electrónico en luz solar exterior se eligió para enfatizar el punto de distorsión. En este punto se debe forzar el gran rango de luminancias de escena casi en un único valor de luminancia renderizable (siendo su rango de luminancia E_{ERDR_OUTS} muy pequeño), y se deben desplazar regiones de la imagen en distancias considerables del eje de aspecto (en cualquier caso ya que la mayoría de los negros serán excesivamente brillantes por los reflejos solares, al menos el rango de aspecto será pequeño, el visualizador puede también compensar eso usando simplemente valores de accionamiento físico en un correspondiente rango de luminancia de salida pequeño). Esto tiene, por ejemplo, como consecuencia que regiones oscuras pueden no renderizarse totalmente, y deben tomar elecciones severamente distorsionadas. En lugar de mostrar, por ejemplo, 10 % de negro si el 50 % es el menor valor visible, se podría simplemente también renderizar todos esos valores cerca del 50 %, o incluso mejor, con un mapeo de tono a valores por encima de eso. Por ejemplo, se puede recortar toda la región más oscura a lo que este visualizador tiene como su "negro" (es decir, su menor valor renderizable), que con un rango de luminancia tan pequeño incluso puede no parecer negro, porque la alternativa de extender las luminancias de objeto oscuras sobre luminancias más brillantes no es una opción, ya que pueden a continuación volverse más claras que algunos de los píxeles de vitral. De forma similar, se debe abandonar el deseo de que algunas escenas puedan renderizarse fielmente en una impresión. Se puede únicamente hacer lo mejor que se puede para usar principios de mapeo y psicovisuales para tener al menos un buen equivalente (pero no ventanas brillantes a menos que se incorporen tintas fluorescentes o similares y se ilumine intensamente con una fuente de UV). Obsérvese que por simplicidad solo analizamos los principios en un eje de luminosidad de una dimensión simplificado. La naturaleza tridimensional de las gamas reales (es decir, principalmente las de los dispositivos de renderización) también tiene un impacto interesante en el procesamiento cromático de colores, por ejemplo, su saturación, que visualmente incluso puede confundirse/equilibrarse parcialmente con brillo en algunas situaciones.

Obsérvese que para integridad también hemos mostrado aspectos de saturación, ya que se producen en escenas naturales, por ejemplo, cuando se mira a lámparas, o, por ejemplo, cerca del sol. Esto es cuando los niveles de conospina obtenidos durante un corto tiempo se distorsionan severamente (blaqueamiento) y se ven puntos. Por ejemplo en una escena invernal se puede mirar a un sol bajo, y el aire alrededor del mismo puede ser muy brillante, y la luz del sol reflejándose en partículas de las nubes alrededor del sol pueden ser incluso más brillantes. Por supuesto, no es deseable renderizar estas regiones en ningún visualizador HDR con colores brillantes con saturación de pigmento visual, pero se pueden asignar dos diferentes subrangos de luminancia en la región híper brillante, es decir, por ejemplo, mostrar esas regiones al menos un poco irritantemente brillantes. Por otra parte, se puede también considerar que estos colores ya no son importantes de cualquier modo (a quién le importa el brillo real o color de un filamento de lámpara incandescente de cualquier modo, aunque la codificación de casas de colores iluminadas brillantemente, o incluso algunos reflejos especulares, o signos comerciales de tubo TL de color, etc. puede ser importante aún de cualquier modo) y codificar los mismos con un valor similar a recorte, que se puede llamar simplemente híper brillo, o una región cercana a los códigos máximos (por ejemplo, solo el valor 1023). El visualizador puede a continuación elegir si quiere renderizar ese brillo irritante, o con una alguna menor luminancia de salida, en cuyo caso el cerebro puede estimar el brillo a partir del recorte. También permite que el creador de contenido se centre en lo que necesita codificar de forma precisa, y que cuando se usa casi directamente para accionar, por ejemplo, un visualizador HDR producirá una buena calidad (por ejemplo, contraste) para todas esas regiones (por ejemplo, tanto interiores oscuros y una sala incluso más oscura, como exteriores soleados), y que muchas regiones brillantes que considera menos relevantes pueden siempre recortarse (potencialmente con una luminancia de salida por debajo del brillo máximo, por ejemplo, en un modo de ahorro de potencia), incluso en visualizadores HDR. Tales modos de ahorro de potencia pueden realizarse mejor por el visualizador si el graduador define un número de tales regiones "irrelevantemente brillantes", habitualmente con varios valores gTS, que el ahorrador de potencia puede usar para distorsionar la imagen por encima de tales valores para un número de modos de ahorro de potencia aumentados. De hecho el creador puede incluso usar artísticamente uno o más códigos "de saturación" para descartar contenido importante de la escena como se representa, por ejemplo, para realizar un aspecto de alta saturación.

Ahora se puede querer transformar representaciones de una escena en una primera colorimetría -en particular un espacio de color que define los objetos de escena con primeras coordenadas a lo largo de una luma o luminancia o eje relacionado con valor de gris similar (suponiendo por simplicidad que de las dos coordenadas de color cromáticas son fijas en ambas representaciones, por ejemplo, una tonalidad y una saturación) de acuerdo con una primera regla de asignación (definiendo una luminancia de parche de objeto de escena local como una luma de píxel; y aunque en lugar de lumas podríamos codificar también los píxeles, por ejemplo, con luminancias en un sistema XYZ, por simplicidad llamaremos los valores de gris codificados lumas)- a una segunda colorimetría. Como solo un ejemplo para describir fácilmente a continuación algunos conceptos y realizaciones de la invención, supondremos que tenemos una escena original con una relación de luminancia de, por ejemplo, 2097152:1, o 21 bits si se codifica linealmente. Por supuesto, esto puede aún suplementarse con un valor de luminancia exacto del punto de brillo con el que corresponde el valor 2^{21} (que puede ser diferente para una escena de exteriores soleada que para una escena de interiores de tarde oscura). En la práctica, ya que ningún visualizador puede codificar el sol de cualquier modo, por simplicidad supondremos adicionalmente que podemos codificar relativamente fielmente (es decir, con distorsiones psicovisualmente menos importantes, como disminuir la luminancia del sol en su versión renderizada de visualizador) estas escenas HDR originales con una codificación HDR maestra de 16 bit (al menos para luma Y, y no nos importa por ahora si eso es flotante o int). Esto es porque se puede definir que la codificación sea no lineal a lo largo de su eje

de luma, es decir usando una gamma maestra para mapear luminancias de objeto de escena a códigos de espacio de color HDR.

Otro ejemplo es codificar, es decir, mapear esa codificación de 16 bits en una nueva colorimetría/espacio de color, a saber un código de 8 bits, por ejemplo, con la gamma 2,2 estándar. Existen varios mapeos de gama para eso, por ejemplo, se podría simplemente comprimir linealmente el rango de luma, pero ya que eso resulta en malos resultados, normalmente se usa una curva sigmoidea más gradual, y se pueden usar modelos más complejos, como, por ejemplo, aplicar la compresión a una versión de filtro paso bajo de la imagen, y a continuación añadir de más intensamente el detalle de paso alto a eso. O el mapeo puede modelar cómo el sistema visual humano aproximadamente (existiendo por supuesto las imposibilidades anteriormente descritas para hacer alguna clase de renderizaciones en hardware limitado) vería la escena original, si se vieran en el nuevo marco de, por ejemplo, un visualizador con mucho menor rango dinámico, es decir un visualizador LDR. El sistema visual humano se comporta no linealmente, disminuyendo los aspectos visuales menos importantes, y, por ejemplo, una sombra áspera en la escena original (al menos cómo algunas cámaras ven la misma) puede verse como grisáceo relativamente claro. No debería cometerse el error de mapear la misma a una gama LDR de modo que mucha de la sombra se vuelve casi el negro mínimo de ese visualizador, porque a continuación por supuesto el sistema visual lo interpretará como demasiado oscuro. Debería suavizarse en cierto modo disminuyendo el contraste (local), de modo que parecerá menos profundo, que en la escena original. En general, el mapeo de gama a la gama LDR puede usar toda clase de matemáticas que aplican optimización local, etc.

Por tanto, en conclusión, se aplica una función de transformación a los píxeles de la representación de 16 bits, para obtener la representación de 8 bits. Por ejemplo primero una transformación global, y a continuación algunas transformaciones adicionales espacialmente locales. Y viceversa, se puede transformar (por ejemplo, predecir) una representación HDR (por tanto, por ejemplo, nuestra representación de 16 bits) a partir de una codificación de 8 bits, mediante otro mapeo de color/luma. Un ejemplo de un sistema de este tipo se publicó en el documento WO2012004709 (generación de imágenes de alto rango dinámico a partir de imágenes de bajo rango dinámico).

Simplifiquemos de nuevo la explicación centrándonos en el mapeo para una codificación LDR de 8 bits, a una representación HDR de 16 bits, utilizable para accionar un visualizador HDR de digamos 5000 nits de blanco máximo, y proporcionando de este modo una renderización artísticamente agradable (es decir, buena calidad) de la escena original (por ejemplo, parece razonablemente similar, en que las sombras aún parecen amenazantemente oscuras, etc.; nota importante, si la codificación maestra de 16 bits original fuera una gradación óptimamente ajustada por un artista informático de acuerdo con las direcciones del director y/o DOP, por ejemplo, haciendo la región de sombras incluso más tenebrosa o amenazantemente oscura, a continuación la intención de calidad puede ser que el visualizador HDR definitivamente de renderización transmita ese aspecto amenazante tan bien como sea posible, es decir, según se pretende).

Pueden existir diferentes formas de mapear los píxeles desde su valor de código de 8 bits, a píxeles (para la misma posición espacial) con un nuevo y diferente valor de código de 16 bits. Por ejemplo, este mapeo puede potenciar el brillo renderizado del vital, ya que el visualizador HDR es capaz de renderizar tales regiones brillantes, que corresponderán con una correspondiente transformación para obtener la luma de píxel de la imagen HDR (supóngase por simplicidad que este acciona directamente el visualizador HDR), basándose en cómo se comporta el visualizador HDR y se define el código HDR. Obsérvese que cuando describimos el comportamiento de brillo de los objetos representados y hablamos acerca de, por ejemplo, potenciar por simplicidad compararemos luminancias de salida (por ejemplo, luminancia renderizada en el visualizador LDR = 400 de entre 500, frente a 3000 en el visualizador HDR), en las que puede realizarse lo mismo en un espacio de luma codificado real, por ejemplo, incluso atenuado las regiones más oscuras (proporcionando relativamente el mismo resultado), y mantenido los vitrales altos tanto para la codificación HDR como LDR.

Una transformación puede ser global, es decir, siempre que el píxel se sitúe en la imagen, la forma funcional de la transformación depende únicamente del valor de píxel de imagen de 8 bits/LDR, es decir: $Y_{16b}=f(Y_{8b})$, en la que Y_{16b} es la luma de 16 bits, que puede representarse, por ejemplo, como una palabra de código binaria, o un valor flotante entre 0 y 1, etc., y de forma similar para la luma de 8 bits Y_{8b} . Un ejemplo de una función de este tipo es una gamma global: $Y_{16b}=g \cdot Y_{8b}^{\gamma}$, en la que g es un factor de ganancia, y γ el coeficiente de una función de potencia.

La ventaja de una función global de este tipo es que se necesita codificar únicamente una pequeña cantidad de datos, por ejemplo, se puede transmitir la gamma y ganancia antes de cada instantánea, o incluso una toma de instantáneas de la misma escena, teniendo las mismas características de imagen.

La desventaja es que si se usa para ir desde HDR/16 a LDR/8 bits (es decir, una señal que se supone que se ve bien en un visualizador LDR de digamos 200 nits de blanco máximo), aunque aproximadamente hace la apariencia correcta (un comportamiento de compresión linealmente de una instantánea de HDR con regiones de alto brillo, es que se comprimen las partes más oscuras demasiado, haciendo que la instantánea se vea en promedio demasiada oscura en LDR visualizadores, pero una función gamma ya puede manejar de forma equilibrada aproximadamente dos regiones, una más oscura frente a una más brillante), porque se corrige la compresión en las partes más oscuras de

la instantánea haciendo eso menos con la forma de gamma apropiada. Pero esta es únicamente una única y simple forma funcional. Cuando se ajustan críticamente algunos colores en un segundo plano, colores similares en un objeto de primer plano pueden cambiar en una forma que es menos deseable para ese objeto. También cuando se mueve desde 8 a 16 bits, se pueden poner, por ejemplo, las luces brillantes en la posición de luminancia de salida de visualizador HDR correcta (es decir, la Y_{16b} correcta), pero haciendo eso ajustando/estirando la función gamma, se puede hacer, por ejemplo, las regiones más oscuras más brillantes de lo deseado. O se puede usar una función de mapeo más compleja como una función polinómica a trozos con puntos de control, óptimamente seleccionados, pero aún se arriesga a desplazar algunos de los grises intermedios a luminancias no deseadas, sin hablar acerca de esto quizás no sea la mejor forma de controlar el aspecto de color de la imagen.

El problema se agrava porque se puede haber hecho, por ejemplo, un mapeo con pérdidas de la escena original HDR a la imagen LDR de 8 bits, que puede suceder, por ejemplo, para la escalinata oscura y el objeto muy oscuro. Aunque originalmente en la escena a capturar ese objeto muy oscuro estaba mucho más oscuro que la escalinata, en la imagen de 8 bits puede tener valores de luma que corresponden a valores de al menos algunos de los píxeles de escalinata. Es decir píxeles que deberían tener valores de luma (muy) diferentes ahora tienen los mismos valores (o al menos los histogramas de conjuntos de píxeles pueden solaparse donde no deberían), aunque las buenas noticias es que pueden residir en diferentes regiones espaciales de la imagen. Una única función que opera en los valores de gris codificados ya no puede discriminar estos dos casos. Es decir, si se quiere transformar el objeto muy oscuro a luminancias Y_{16b} que son muy bajas, sucederá lo mismo erróneamente con parte de la escalera de escalinata (resultando, por ejemplo, en un oscurecimiento de contraste excesivo de algunas partes de la escalera), o viceversa. Es decir el artista o aparato de transformación de coloración querrá poder aplicar una transformación diferente a esos dos objetos.

La otra clase de transformaciones son transformaciones de luma locales (o en general color), que aplican una función diferente a cada píxel. Se podría mirar, por ejemplo, a un área de máscara alrededor del píxel, e potenciar su luma un poco dependiendo de cuáles son los valores circundantes, por ejemplo, si son casi los mismos aunque en cierto modo diferentes. Un ejemplo de esto es el pico alrededor de bordes de objetos, en los que se quiere potenciar las lumas de píxeles locales en cierto modo por encima frente a por debajo del valor de perfil de paso en la vecindad del borde. Un ejemplo en transformación/codificación de imágenes HDR es el principio JPEG-HDR, en el que se usa una imagen JPEG normal para la textura, y a continuación co-codifica una imagen de potenciación que tiene un factor de potenciación para cada píxel. La ventaja es que se podría co-codificar deseo de transformación local algorítmica según se realiza como un resultado final en una imagen de este tipo (por ejemplo, aumentando el contraste de textura de una primera forma, y un gradiente de iluminación semi global en otra, que el artista de gradiente podría optimizar según su deseo), aunque eso implique un precio severo de una cantidad aumentada de datos a codificar, ya que ahora para cada imagen HDR dos imágenes LDR tienen que codificarse (frente nuestra única, por ejemplo, imagen de contenedor LDR mencionada anteriormente). Se podría preguntar si se codifica $8\text{bit_texture} * 8\text{bit_boost}$, si no es mejor simplemente codificar en bruto 16bit_HDR .

Ahora existe una forma intermedia, porque si es deseable cierta potenciación, normalmente será deseable para todo un objeto, por ejemplo, el vitral. Es decir, los valores de píxeles luma/de potenciación en la imagen de incremento no estarán totalmente decorrelacionados, incluso más, si se procesa/codifica de forma inteligente, pueden correlacionarse de tal forma que se pueden representar los mismos de una forma más simplificada. Es decir se pueden especificar los mismos paramétricamente de una forma funcional, quizás incluso tan simple como con un único número de potenciación que puede almacenarse en metadatos co-codificados.

Eso requeriría codificar objetos, o más genéricamente regiones de imagen geométricas.

Un ejemplo fácil de esta segmentación en piezas es definir una cuadrícula de bloques, y a continuación definir una transformación óptima para cada subárea rectangular. Por ejemplo se podría definir un multiplicador de ganancia y compensación para cada uno de esos bloques en el documento EP2009921 [Liu Shan et al. Mitsubishi Electric: Method for inverse tone mapping] o co-codificar una gamma a local para cada bloque diferente. Tales métodos normalmente sufren rápidamente de artefactos de bloques. Por ejemplo, se puede llegar a una ganancia óptima diferente o gamma a aplicar al bloque $BLCK(i+1, j-1)$ (véase la Figura 2a) y quizás para bloques más allá de ese como hasta $BLCK(i+1, j)$ que para el bloque $BLCK(i+2, j)$. Esto es porque el primer bloque puede optimizar la transformación valorando altamente un aspecto óptimo de la escalinata, mientras que para el segundo bloque se puede optimizar, por ejemplo, centrándose en un criterio de visibilidad del objeto muy oscuro. Incluso desviaciones pequeñas en una parte de la curva (es decir, para algunas lumas de píxeles disponibles Y_{8b}) puede resultar en una visibilidad de una diferencia de las estadísticas de las luminancias Y_{16b} de las partes/objetos de segundo plano en estos bloques, es decir, resultando en la percepción de un límite visual, ya que el cerebro está entrenado para detectar tales diferencias estadísticas, para eso puede significar detectar un tigre escondiéndose en la hierba amarillenta. Cuando se aplican algunos algoritmos, se puede ver una cuadrícula gruesa, esa visibilidad se aumenta mediante modulaciones temporales de las estadísticas de color de las regiones subyacentes después de la transformación a Y_{16b} .

Ahora también existe una posible solución a ese problema, a saber podrían codificarse de forma precisa todos los objetos y, por lo tanto, garantizar que todos los píxeles del objeto oscuro de primer plano consiguen su transformación local óptima, y los píxeles de la región de segundo plano de todos los bloques en esa región consiguen la misma transformación proporcionando a los mismos la renderización óptima y, por lo tanto, sin bordes de bloque visuales.

5 Pero todas las esperanzas de hacer eso se evaporarían en vista de la eficacia de codificación es decir la cantidad de bits necesarios de nuevo, conduciéndonos hacia la obligación de aceptar la codificación de dos imágenes, o probablemente la codificación en bruto de Y_16b (ya sea quizás que a continuación para compatibilidad hacia atrás sería necesaria adicionalmente otra codificación de Y_8b). Adicionalmente, no únicamente la codificación precisa el límite actual de, por ejemplo, nuestra escalera implica muchos datos a codificarse para, por ejemplo, una función polinómica a trozos, sino también graduadores a menudo les gusta hacer sus selecciones de objetos menos precisa, especialmente cuando necesitan hacer cientos o miles de tomas en una película.

10 En el pasado, tales codificaciones orientadas a objetos se han intentado en el marco de MPEG4-2, pero nunca han sido satisfactorias por diversas razones. No se puede simplemente extraer estos objetos, ya que no se sabe cuáles son sus definiciones de patrón de textura espacialmente variable, por tanto se conduce a la codificación de sus límites. Pero eso por un lado conduce a estructuras de codificación complejas (en comparación con la simplicidad universal de tecnología de codificación basada en bloques), tal como, por ejemplo, funciones polinómicas a trozos o serpientes, y en segundo lugar probablemente la necesidad de intervención humana para colocar óptimamente esas serpientes (ya que desalineamiento de límites es una lacra de muchos algoritmos, por ejemplo, perdiendo una pieza de esquina de un objeto), y en tercer lugar muchos valores de código adicionales para codificar tales curvas de límite. Todos estos factores de complicación no parecen favorecer la fácil adopción en normas de codificación de video o instantáneas fijas prácticas.

20 Pero el inventor se dio cuenta que en la situación de codificación HDR particular (es decir, cuando se transforma entre una primera representación de, por ejemplo, menor rango de luma dinámico y segunda representación de mayor rango de luma dinámico de una escena) existe casi siempre una propiedad particular de la imagen que permite una codificación todas las ventajas de segmentación precisa, aún con la ventaja de necesitar únicamente también unos pocos bits de datos adicionales. En todas las renderizaciones (o codificaciones de imágenes subyacentes) de la Figura 1, existe siempre una jerarquía de brillos de región (abarcando diferentes rangos de luminancia o luma), por ejemplo, la ventana siempre va a ser el objeto más brillante. Y aunque espacialmente pueden existir objetos más oscuros a la derecha, objetos más brillantes en el medio y de nuevo objetos más oscuros en la derecha, habitualmente en cada región local siempre existe parte de la instantánea que es más oscura, (realmente, pueden existir varias clases, como objetos de oscuridad media, pero al menos algunos píxeles son los más brillantes y algunos son los más oscuros, y normalmente incluso tienen unas estructuras geométricas relativamente simples como la estructura de relleno sólida convexa de la ventana de cristal). Pero obsérvese que incluso si tenemos un patrón de barrotes de prisión contra un sol brillante en un bloque eso no es un problema, ya que los barrotes de prisión se discriminan fácilmente dentro del bloque como que tienen los píxeles más oscuros. También, la distribución sobre varios bloques normalmente es fácilmente gestionables, incluso si implica reiniciar algunos valores gTS en algunas ocasiones entre bloques a lo largo de una trayectoria de exploración. Para un caso extraño que accidentalmente resultaría más difícil, se puede tomar por supuesto recurso a métodos o estrategias auxiliares.

40 El principio se explica con la Figura 2a y la Figura 2b.

45 En la Figura 2a ahora hemos mostrado nuestra imagen con iluminación de vitral de la escalinata de madera oscura con su subdivisión de bloques superpuesta. Es en estos bloques que, por ejemplo, un algoritmo análisis de imágenes automático analizaría las estadísticas de imagen locales, tal como, por ejemplo, el histograma de luma local (o derivado del mismo histograma de luminancia, por ejemplo, de una representación de escena en una colorimetría de referencia, de renderización de visualizador), y llegar a una posición para crear una imagen HDR Y_16b imagen HDR transformando una imagen HDR Y_8b. Por ejemplo, puede usar principios estadísticos y conocimiento acerca de cómo se vería una imagen típica (si la escalera está ya relativamente oscura, un primer mapeo particular puede hacer la misma habitualmente demasiado oscura en un visualizador LDR, o el graduador puede solo probar un escenario de este tipo comprobándolo), y a continuación seleccionar una gamma de mapeo de, por ejemplo, 4,3. Un ejemplo de una transformación deseada de este tipo se muestra en la Figura 2b. Como se ha dicho anteriormente, no es necesario que haya una función o algoritmo de transformación completa (en lugar de una función gamma, se podría tener un conjunto de reglas programadas, como, por ejemplo, calcular una medida de textura local, local en una variación local de luma, etc., para llegar a un valor de luminancia final para uno o más local píxeles) por píxel, pero deseamos una transformación optimizada semi global, es decir, una transformación por objeto o clase habitualmente. En la región de imagen cubierta por el bloque BLCK(i-1,j-4) vemos una subescena local en el área seleccionada con ese bloque comprendido de dos objetos, a saber a parte del vitral, y la pared alrededor del mismo (que podría, por ejemplo, tener ladrillos o papel pintado cuya textura también tiene que renderizarse con buena calidad, pero por simplicidad eso no se muestra), que ocupa esos píxeles del bloque que no son vitral. Porque estos objetos son muy diferentes (un contra luz que representa los píxeles más oscuros contra el exterior brillante, que no empieza a explicar incluso que los colores saturados luminosos del vitral puede demandan un tratamiento especial), podemos desear aplicar transformaciones muy diferentes para obtener una nueva codificación de la imagen tal como, por ejemplo, Y_16b, al menos para algunas categorías de visualizadores para las que se destina habitualmente o al menos es útil esa señal. La ventana y la pared son objetos muy diferentes y, en particular, se iluminan de forma diferente. No únicamente se ilumina la pared (con cualesquiera propiedades físicas que tenga, tal como un BDRF, albedo, etc.) con cualquier luz que está en el interior de la sala, sino que habitualmente crea su color/luminancia reflejando luz desde sus alrededores, y en particular su principal fuente o fuentes de iluminación. La ventana por otra parte es un color translúcido, ya que

modula directamente por absorción de la luz del exterior. Al menos nos gustaría ver que la ventana es más brillante en cualquier renderización de visualizador, pero podrían existir criterios de calidad de renderización adicionales, en vista de este diferente mecanismo de generación de color. Podría ser que se quiere mostrar en el visualizador HDR la pared con luminancia de salida de visualizador relativamente atenuada, no muy diferente de lo que mostraría un LDR, o una pared real reflejaría estando en el entorno de visualización del visualizador y espectador oscurecido. Por otra parte, se puede querer potenciar la ventana de cristal, que digamos se codifica en la imagen LDR con valores de luma no mucho mayores que los de la pared, ya que de otra manera un visualizador LDR no puede mostrar los mismos (relativamente fiel) de cualquier modo, en luminancias que están en algún sitio cerca de la parte superior de la gama realizable del visualizador HDR, es decir, teniendo una coordenada de luma alta Y_{16b} . Es decir, una imagen HDR apropiada tiene que construirse con paredes más oscuras y ventanas muy brillantes.

En la Figura 2b mostramos otro ejemplo que muestra qué hacer con la escalera, y se muestra una función de mapeo de luma de comportamiento total $TM_BLCK(i,j)$ para lo que es deseable: en caso de que cualquier posible luma de entrada $Luma_{in}$ estuviera presente en para un píxel en ese bloque, que entonces es la luma de salida $Luma_{out}$ transformada de la imagen HDR Y_{16b} . Por supuesto, algunos colores no están presentes en realidad (en ese bloque no hay ningún vitral), por tanto los hemos mostrado con línea discontinua. Los que es relevante son las funciones de transformación para esos rangos de Luma en los que están presentes. Por tanto, el experto entenderá que esto permite varias realizaciones de realizaciones, basándose en, es decir, versatilidad deseada o complejidad de codificación. Se podría almacenar la función entera $TM_BLCK(i,j)$ con las partes con línea discontinua proporcionando algunos valores (ya que esto se consigue fácilmente si se codifica la transformación con una forma funcional tal como función gamma, pero también si la transformación se codifica como una tabla de consulta, y los valores intermedios pueden ser útiles en partes de la imagen o imágenes en las que están presentes), o se podría almacenar en ubicaciones separadas únicamente las subtransformaciones, tal como la transformación parcial $PD_BLCK(i,j)$ necesitada para la escalinata definida sobre la luma en el rango RNG_COD_DRK . Juntar tales transformaciones parciales tienen muchas ventajas. Pueden almacenarse en cualquier sitio y por cualquier razón. Se puede entender que la transformación parcial $PD_BLCK(i,j)$ puede almacenarse en cualquier sitio (por ejemplo, en el comienzo de esta toma de imágenes, o incluso en el comienzo de la película) como una transformación más grande que codifica el comportamiento de mapeo del papel pintado, también el lugares en los que es mucho más claro porque, por ejemplo, se ilumina por una lámpara local en su vecindad. Pero a continuación únicamente la parte dentro del rango RNG_COD_DRK se toma del mismo (y, por ejemplo, almacena en una memoria temporal cuando se aplica el algoritmo de mapeo de luminancia a todos los píxeles de ese bloque $TM_BLCK(i, j)$). Tales transformaciones parciales podrían incluso distribuirse como, por ejemplo, una internet u otro servicio de red, por ejemplo, en un servicio de protección de derechos de autor o solo un servicio separado que ofrece una renderización más bonita de algunos objetos, o en escenarios sobre la marcha como juegos, etc.

Por tanto, $TM_BLCK(i,j)$ de este ejemplo muestra dos mapeos de luma relevantes parciales, a saber en primer lugar la parte $PD_BLCK(i,j)$ para la escalera que es un estiramiento lineal con compensación, que abriga la escalera en cierto modo en comparación con su codificación de imagen LDR oscura (es decir, $Luma_{in}$), y a continuación potencia el contraste en cierto modo, haciendo los granos en la madera más visibles. En segundo lugar existe la transformación parcial $PM_BLCK(i, j)$ para el segundo plano de la sala (que en este caso puede ser algún suelo en lugar de papel pintado), que en este ejemplo es un estiramiento (curvado) variable. El mismo mapeo habitualmente se aplicaría en ambas partes del bloque $BLCK(i+1, j-1)$.

Si ahora sin embargo llegamos en el bloque $BLCK(i+2,j)$, esta estrategia de mapeo aún podría funcionar muy bien en la parte de segundo plano, pero no para los píxeles con luma in en el rango RNG_COD_DRK , ya que ahora codifican el objeto muy oscuro. Esto debería mostrarse mucho más oscuro en el visualizador HDR, es decir tener menores luma $out_{Y_{16b}}$ en la imagen HDR mapeada desde la imagen LDR. Eso se muestra mediante la línea más gruesa que muestra la nueva estrategia de transformación parcial $PD_BLCK(i+2, j)$ para ese bloque, es decir ese diferente y nuevo objeto en ese bloque. Tiene un factor de estiramiento mucho más suave, que quiere mantener todos los colores de objeto muy oscuros muy oscuros, y casi indiscriminables.

Por lo tanto necesitamos un nuevo mecanismo técnico que permita cambiar rápidamente tales estrategias de mapeo parcial sobre partes de diversos bloques, que realmente corresponden a objetos reales que requieren diferentes renderizaciones o gradaciones óptimas.

Ahora el inventor se ha dado cuenta que en el mundo de las imágenes HDR (es decir, implicando habitualmente mapeo entre diferentes representaciones de colores de la misma imagen o imágenes, por ejemplo, espacios de color basados en Y_{8b} a Y_{16b}) casi siempre existe una relación especial entre tales regiones parciales u objetos dentro de bloques, en concreto, sus respectivas lumas (o luminancias o representaciones similar) son diferentes. Una luma representativa podría ser una luma promedio, pero habitualmente una propiedad más estricta es que la luma más oscura del primer objeto (en el ejemplo del bloque $BLCK(i+2,j)$ el segundo plano (suelo)) es más claro/alto que la luma más clara de un píxel en la región parcial más oscura (en este ejemplo del objeto muy oscuro). Es decir, se puede demarcar esos dos codificando meramente un "valor de gris de diferenciador de regiones" para al menos ese bloque, y habitualmente un número de bloques más allá (suponiendo una cierta dirección de exploración, por ejemplo, zigzag de izquierda a derecha). Ese valor de gris de diferenciador de regiones es por lo tanto un límite de luma (o coordenada de luminosidad similar de la representación de color, de hecho se puede siempre recodificar la misma para diferentes

definiciones de rangos de luminancia de imágenes, igual que se puede redefinir las estrategias de mapeo desde, por ejemplo, una codificación de [0,255] a una codificación de [0,0, 1,0] de los mismos datos de textura de imagen) más allá del cual se codifica el primer objeto y por encima del cual el segundo objeto. Y aunque la escalinata en bloque BLCK (i+1, j-1) puede necesitar otro valor de gris de diferenciador de regiones, porque esas escaleras contienen algunos valores más brillantes en la imagen LDR que el objeto muy oscuro, el principio permanece el mismo. Para el vitral que contiene bloques se invierte el orden y ahora el segundo plano es la región parcial más oscura en esos bloques, pero el principio permanece el mismo. Teniendo un valor de gris de diferenciador de regiones tan simple, un aparato de lado de recepción puede perfectamente, reconstruir con precisión de píxel los objetos necesarios. Esto no sería posible en codificación orientada a objetos genérica, ya que, por ejemplo, un pez podría contener, por ejemplo, colores azules en su interior que también se producen en el océano alrededor del mismo, pero el problema de representación de imagen HDR es siempre el de tener regiones mucho más claras co-situadas en algún sitio en la imagen con regiones más oscuras. Esto puede suceder habitualmente porque, por ejemplo, esas regiones se iluminan de forma diferente, o incluso se auto iluminan como lámparas. Y otra propiedad es que tales regiones de luminancia (muy) diferentes se separan en cierto modo geométricamente en la imagen, es decir a menudo en diferentes bloques, que permite optimización adicional. Este es el ejemplo del objeto muy oscuro, que es de hecho algo oscuro como la escalera, y puede incluso usar (algo de) los mismos códigos de luma en la imagen LDR. Pero ya que se produce en un bloque diferente, la única cosa que se necesita optimizar son los metadatos semánticos de representación, que pueden ser tan simples como un único valor de gris de diferenciador de regiones, que puede ser, por ejemplo, en este ejemplo el valor superior de RNG_COD_DRK. Es decir, un módulo de segmentación de objeto en el extremo de recepción (que puede realmente ser la misma clase de aparato que en el extremo de envío o existir en el extremo de envío de hecho, pero es un módulo que habitualmente consigue la imagen LDR + metadatos que contienen los diversos uno o más valores de gris de diferenciador de regiones necesarios) puede segmentar con precisión todos los objetos relevantes basándose en el valor del valor de gris de diferenciador de regiones que recibió antes de que comenzó el primer bloque con escalera, y de forma similar para todos los bloques consecutivos. Al menos esta codificación se usará para todos los bloques que contienen escaleras, es decir, hasta que la situación muy nueva se produzca la primera vez, en BLCK (i+2, j), en el que reside el objeto muy oscuro. Antes de que comience ese bloque, la nueva situación se comunica transmitiendo un nuevo valor del valor de gris de diferenciador de regiones. Ahora también en el extremo de recepción el decodificador se reinicia y ordena con los nuevos valores apropiados, para hacer la segmentación de nuevo correctamente, como se habría verificado antes de la finalización del almacenamiento en el extremo de transmisión. habitualmente, el codificador puede conectarse con, por ejemplo, software que permite fácilmente que el graduador defina valores gTS relevantes. Por ejemplo, puede tener un deslizador para establecer un valor, y a continuación ver en pseudocolores (por ejemplo, rojo frente a verde) qué partes de la escena (quizás local para bloques seleccionados) se determinan como por debajo o encima de gTS. O puede seleccionar regiones aproximadamente, y el aparato puede ya ayudar semi automáticamente al graduador, analizar las estadísticas y, por ejemplo, proponer un primer valor de gTs basándose en una estimada de regiones coherentes que varían considerablemente en brillo. O el graduador puede garabatear rápidamente en una región, por ejemplo, dentro del vitral, y ya seleccionar al menos valores de inicio para gTS con la misma, que a continuación podría ajustar a través de cualquiera de diversos controladores de interfaz de usuario.

Y una vez que se tienen estos segmentos, es una tarea fácil asociar los mismos con las transformaciones requeridas. El decodificador puede, por ejemplo, etiquetar todos los píxeles de segundo plano como "objeto=0", y, por ejemplo, aplicar un color global o estrategia de mapeo de luma como se codifica antes del inicio de la instantánea (o incluso por defecto para un tipo de visualizador HDR de referencia, tal como un gamma 4,0). O el decodificador (y codificador para emular primero la capacidad de decodificación) puede antes de un bloque particular actualizar el mapeo a aplicar para objetos de segundo plano/objeto=0. La escalinata puede etiquetarse como "objeto=1", y alguna regla de vinculación asocia un mapeo con esos píxeles (segmentados). Por ejemplo la regla por defecto puede ser que si se codifica un nuevo mapeo antes del bloque, esa función de mapeo (o algoritmo) tiene que aplicarse para las lumas de píxeles que están por debajo del "valor de gris de diferenciador de regiones" actual. O la función de mapeo puede codificarse de tal forma, por ejemplo, aplicable en (principalmente o únicamente) un rango de lumas de este tipo, que tiene que usarse claramente para el objeto más brillante de las dos (o más) regiones.

Por tanto, necesitamos únicamente unos pocos bits de datos adicionales para codificar los objetos, a saber una o más veces, dependiendo de la complejidad de la imagen, un valor de gris de diferenciador de regiones. Para las imágenes más simples con, por ejemplo, únicamente una ventana al exterior, puede ser suficiente un único gTs. Podríamos incluso usar esa estrategia para ajustar mapeos en caso de que no haya discontinuidad de luma clara entre las dos regiones parciales, por ejemplo, para gradientes de iluminación a lo largo del papel pintado de segundo plano se podría usar este mecanismo con valores de gris de diferenciador de regiones variables para aplicar en cierto modo diferentes mapeos a, por ejemplo, las partes iluminadas más oscuras, por ejemplo, para modificar la visibilidad.

Son posibles varios escenarios. Para la mayoría de escenas HDR, habrán únicamente dos diferentes regiones de luminosidad por bloque, y puede incluso haber únicamente un par de diferentes regiones de luminosidad, por ejemplo, 2 (en caso de que únicamente el vitral necesitase un mapeo diferente en comparación con un mapeo global que se evalúa satisfactoriamente para el resto de la instantánea). En ese caso se necesitaría únicamente un par de veces para transmitir un valor de gris de diferenciador de regiones entre los códigos de color de píxel de los bloques (o codificaciones similares, como, por ejemplo, fuera de los datos de píxeles, en una estructura de datos que se puede co-rastrear con la exploración de los bloques). De hecho, en el escenario simple del vitral un único valor de gris de

diferenciador de regiones puede ser suficiente, es decir, podría co-codificarse antes de la toma de imágenes en la película que contiene la escena. En ese caso el módulo de segmentación entenderá que cada luma por encima del valor de gris de diferenciador de regiones se supone que se tratará/mapeará como una ventana brillante. En algunas ocasiones puede haber más de dos objetos que solapan con una única ubicación de bloque, en cuyo caso tendremos un objeto más oscuro, uno con brillo medio y uno más claro. En ese caso pueden segmentarse mediante el mismo principio transmitiendo dos valores de gris de diferenciador de regiones, por ejemplo, antes de ese bloque. Se podría lo mismo también en caso de que únicamente el objeto más oscuro esté en el bloque actual (por ejemplo, con el de brillo medio), y el objeto más claro se produce únicamente un par de bloques después, es decir, se co-codifican estos dos valores de gris de diferenciador de regiones a continuación para una serie de digamos 10 bloques sucesivos.

5 Existe únicamente un escenario que se produce infrecuentemente en el que dos objetos de similar brillo/luma se producen en los mismos bloques, es decir tienen un número de píxeles de la misma luma, significando que no puede asignarse definitivamente a ninguno de ambos objetos, o dicho de otra manera, sus rangos integrales se solapan (considerable, de otra manera tampoco es tan problemático). Ese sería el caso si el objeto oscuro: 1) se codifica realmente con códigos asignados doblemente (es decir, no se reservan, por ejemplo, únicamente tres códigos, luma 0, 1 y 2 para solo nuestro objeto muy oscuro, esos valores no están a continuación presentes en la escalera; y 2) esos dos objetos no se separan como en nuestro ejemplo, sino que se colocan en el mismo bloque, por ejemplo, habitualmente solapándose. En ese escenario y en caso de que el creador de contenido aún se preocupe acerca de tener una codificación de alta calidad de este tipo de esas regiones oscuras de cualquier modo, nuestro codificador necesitaría usar un escenario de repliegue, por ejemplo, en una estrategia de codificación de imagen HDR, en lugar de predecir toda la instantánea mediante mapeos de segmentos locales basándose en nuestra segmentación guiada por metadatos presentemente descrita, necesitaríamos una codificación diferente, por ejemplo, se podría además co-codificar una imagen pequeña del tamaño de ese bloque que contiene directamente los valores requeridos Y_16b, y a continuación superponer los mismos en la imagen HDR en las ubicaciones de píxeles de esos bloques. Y se podría aún usar el mecanismo de valor de gris de diferenciador de regiones para eso usando umbrales reservados particulares. Por ejemplo, un valor de gris de diferenciador de regiones de cero o -1 parecería que "no tiene sentido", porque no hay luminancias por debajo de eso, es decir, puede significar la estrategia de codificación de repliegue (por ejemplo, superponer). Además de señalización de una estrategia de codificación HDR (u otra imagen) de sustitución, como, por ejemplo, codificar una pequeña parte de una imagen a partir de un video no como Y_8b LDR, sino un RAW parcial, por ejemplo, imagen Y_16b, a substituir (también después de transformación apropiada habitualmente) para esa región de la imagen cuando se genera una imagen de mayor rango dinámico, se puede también usar valores reservados por otras razones. Por ejemplo un valor de gris de diferenciador de regiones de 260 puede indicar que el siguiente bloque es tan difícil de segmentar, no será posible basándose en uno o más valores de gris de diferenciador de regiones en el rango de luma codificado (por ejemplo, 16, 200 y 240), sino que se necesita otra estrategia de segmentación. Por ejemplo, tras detectar este valor reservado de 260, el lado de recepción puede usar un mapa ya segmentado para al menos el actual o más bloques segmentados. Es decir, mirará a continuación en una imagen de segmento co-codificado pequeño, en la que para al menos los bloques sucesivos los píxeles se etiquetan, por ejemplo, como o bien "0", "1", o bien "5", en caso de que estos sean tres tipos de objetos presentes. Después de que esta segmentación de repliegue ya no se necesite, el algoritmo regular basado en "valor de gris de diferenciador de regiones" puede reiniciarse mediante, por ejemplo, re-codificación antes de que el primer bloque al que la segmentación regular aplique los antiguos valores no reservados de nuevo (por ejemplo, 16, 200 y 240), u, otro código gTS reservado como 270 podría usarse para indicar que reanuda la codificación de metadatos de segmentación regular, y que los valores gTS anteriores (habitualmente almacenados en memoria de trabajo en el lado de recepción) son de nuevo aplicables.

10
15
20
25
30
35
40

45 Pero de cualquier modo, incluso cuando se necesita ocasionalmente otra estrategia de repliegue para las situaciones raras muy complejas, tenemos una codificación que es muy eficiente en datos (porque principalmente necesitamos únicamente los mapeos y los valores de gris de diferenciador de regiones que delimitan en qué píxeles cuyo mapeo necesita aplicarse, y normalmente algunos metadatos adicionales para especificar precisamente ese enlace (por ejemplo, regla de vinculación de transformación: objeto=3 usa mapeo=5)) porque únicamente muy raramente necesitamos codificación de respaldo alternativa con más consumo de bits. Pero adicionalmente, es también muy versátil en aplicaciones de procesamiento, como, por ejemplo, ajuste para renderización en visualizadores diferentes. Porque mediante nuestro método hemos definido de forma fácil la semántica HDR de la escena, que se necesita para ajustar hacia diferentes visualizadores. Y la regla de vinculación puede ser dinámica, por ejemplo, puede haber un número de reglas almacenadas. Por ejemplo mapeo=5 puede rellenarse adicionalmente mediante diferentes mapeos dependiendo de, por ejemplo, qué representación de color de imagen HDR de salida será el resultado al que se mapea (por ejemplo, Y_16b frente a Y_10b), o para qué visualizador estará (por ejemplo, un visualizador HDR de 5000 nits, o un visualizador HDR de 50000 nits, o un visualizador HDR de 1500 nits), etc.

50
55

60 Y el experto entenderá que esta codificación puede incorporarse de diversas formas, por ejemplo, mediante diferentes reglas, por ejemplo, cambiando en inicialización la regla de vinculación desde "objeto=3 usa mapeo=5", a "objeto=3 usa mapeo=7", o habiendo producido el módulo de segmentación diferentes códigos de segmento dependiendo de los particulares de la imagen HDR de salida, o el mapeo puede cambiarse dinámicamente, por ejemplo, haciendo referencia a particulares adicionales (como un apuntador variable al inicio de diferente algoritmos o LUT o entradas en listas de diferentes ecuaciones, etc.). Esto también permite, por ejemplo, manejo fácil de órdenes de interfaz de usuario, tal como, por ejemplo, una modificación controlada por usuario del "aspecto de brillo general" de la escena, que puede implementarse reasignando diversas nuevas funciones de mapeo a diferentes objetos de la escena,

65

modificándose la luminosidad de todos los objetos/regiones de valor gris en una relación de mapeo particular (por ejemplo, el vitral puede no modificarse en gran medida, ya que ya es brillante, quizás abrigantado un poco adicionalmente en comparación con el resto de la imagen para no perder mucho del aspecto HDR debido a la relación de luma con el resto de la imagen, pero el interior de la sala circundante puede abrigantarse, ya que es esa parte de luminosidad promedio que influencia principalmente el aspecto de brillo de imagen; y viceversa cuando se atenúa el visualizador, porque el usuario percibe el mismo como incómodamente brillante, por ejemplo). Únicamente esa parte sobrescrita necesita a continuación tratarse de forma diferente, pero si es tan difícil y crítica, probablemente necesitara procesamiento de renderización complejo de cualquier modo. Por tanto, el creador de contenido tiene una mucho mejor opinión en qué se, por ejemplo, abriganta y cómo se abriganta, ya que el abrigantamiento no necesita ser una compensación o multiplicación simple, sino que puede ser una estrategia compleja que equilibra brillos de subregiones (por ejemplo, recortando porcentaje de los píxeles a blanco), que habitualmente es inteligente en escenarios de visualizador limitado por gama, en los que los algoritmos actuales pueden conducir a artefactos.

Otro uso muy útil de umbrales de demarcación de segmentos reservados se ilustra en el sistema doméstico de lado de recepción de la Figura 6. En este punto, la televisión 602 recibe una señal SB, por ejemplo, desde una estación de televisión a través de vías aéreas. Comprende en metadatos METB, y código o códigos de especificación de configuración de HDR SET_HDR, que pueden co-transmitirse de diversas formas (pero habitualmente como metadatos en el inicio de una serie de imágenes) y especifica cómo el visualizador debería comportarse en lo sucesivo. Un código SET HDR interesante puede usarse para conmutar entre renderización HDR y LDR, repliegue, por ejemplo, para ahorrar potencia, porque en la actualidad difundimos en continuo, por ejemplo, un programa de noticias de estudio, que no necesite la cantidad máxima de efectos cinematográficos de HDR. Por tanto, entre, por ejemplo, el anuncio o película antes del mismo y las noticias, puede transmitirse un código SET HDR de "render_LDR" (o co-codificarse en un programa de video almacenado en, por ejemplo, un grabador de disco duro doméstico, o servidor de almacenamiento de video conectado a internet), que significa que desde ahí en adelante el visualizador HDR renderizará con un brillo de blanco máxima de, por ejemplo, únicamente 500 nits (aunque tiene una capacidad de brillo máximo de 5000 nits). Ahora como realizaciones de nuestra invención divulgada presente, se puede hacer tan fácilmente estableciendo el valor de gris de diferenciador de regiones gTR igual a 255, que significa que todas las luminancias por debajo (es decir, todas las luminancias posibles en la imagen de 8 bits) necesitan tratarse con el mismo mapeo, que puede, por ejemplo, co-almacenarse con la imagen, o prealmacenarse en el mapeo de gamma de visualizador que a continuación renderiza todo a máximamente 500 nits. Por ejemplo se puede usar un valor gTS que demarca que grises se renderizarán, quizás atenuarán, y por encima de esos todos los grises pueden recortarse a un blanco atenuado relativamente oscuro.

Ahora es importante entender que hay dos clases de mapeos/transformaciones a aplicar a las lumas (o codificaciones relacionados con luminosidad).

El primero son simples mapeos "relacionados con hardware" matemáticos, que solo corrigen para el visualizador de visualización particular y entorno. Por ejemplo, si una imagen se codifica para una CRT de gamma 2,2 (visualizador de referencia), pero se muestra en un LCD con una función de transferencia electro-óptica sigmoidea, el propio visualizador puede usar matemáticas colorimétricas elementales para corregir eso, haciendo que el LCD renderice la imagen como si fuera una CRT de referencia. De manera similar, se puede optimizar en gran medida para características de entorno de visualización con matemáticas simples. En primer lugar por supuesto, cuando se escala a una renderización más oscura, se necesita reducir el brillo de referencia asociada con la codificación de imagen. Esto se realiza en gran medida mapeando el valor de código máximo (por ejemplo, 255) al brillo máximo del visualizador de renderización. Sin embargo, puede también hacerse más complejo, por ejemplo, un subrango particular de luminancias de la imagen podría asignarse a un rango particular de luminancias renderizadas del visualizador. Pero habitualmente, también tiene que aplicarse una corrección de gamma, teniendo en cuenta tales cosas como un cambio por el contrario dependiendo de la luminancia de la imagen renderizada y su ambiente. Esto proporciona unos resultados bastante aceptables si el contenido de información de rango de luminosidad (es decir, la codificación de información sobre los diversos subrangos de aspecto) en los dos sistemas es relativamente similar, incluso es difícil cuanto los rangos de aspectos son muy diferentes. Para ir a un rango dinámico mucho más estrecho se tiene que decidir qué subrangos aún tienen que mostrarse con qué calidad, es decir habitualmente con qué contraste dentro de objeto y qué contraste entre objetos, y a menudo se generan rangos de solapamiento por el algoritmo de mapeo. A la inversa es incluso más difícil. Si tenemos únicamente un par de rangos de objetos comprimidos, es difícil evaluar dónde poner los mismos en un rango de aspecto HDR de salida, y mucho menos inventar nuevos valores de luma/aspecto. Se vuelve incluso más difícil generar una imagen HDR buena y natural a partir de una imagen mapeada LDR en la que los rangos de luma de 8 bits de objetos distintos de luminancia de escena se han codificado inapropiadamente solapándose entre sí (como cuando se simula una iluminación muy uniforme, destruyendo toda o suficiente información de la iluminación de escena original).

Todas las transformaciones de los valores de gris (o en general colores) de los píxeles de objeto, pueden verse como transformaciones genéricas, que pueden ser locales el lugar de globales, que se hacen habitualmente por un artista (puede incluso corregir si la calibración matemática simple anterior para un escenario de visualización diferente no es lo suficientemente precisa). Los artistas pueden hacer gradaciones artísticas complejas, en las que, por ejemplo, cambian las diferentes luminancias presentes en nubes en la instantánea, para hacer que una tormenta parezca más amenazante. O pueden incluso usar un efecto de renderización de gráficos por ordenador, para hacer que rayos de

luz salgan de los ojos de un robot, como se representa por las lumas de píxeles/colores deseadas. Para nuestra explicación podemos ilustrar dos escenarios típicos. O bien todo el posicionamiento importante en el rango de luminancia de los valores de gris de píxeles de objeto (la gradación de color) se ha hecho en una imagen HDR maestra (IM_MSTR_HDR, véase la Figura 4, que puede ser, por ejemplo, una imagen de 16 bits con una gamma de definición particular), y la imagen LDR (Im_1) se deriva puramente mediante transformaciones matemáticas simples en esa maestra HDR, como, por ejemplo, una función de gamma de la que el factor de gamma se ajusta basándose en tales características como el histograma de la imagen HDR, es decir la conversión HDR a LDR es meramente un desplazamiento adecuado simple de los valores de gris (habitualmente no a desviaciones grandes), de modo que toda la información se contiene relativamente con precisión, ya sea en una estrategia de codificación diferente. En este caso una imagen HDR final a renderizar en un visualizador HDR puede derivarse a partir de la imagen LDR aplicando esta estrategia de mapeo matemática inversa. O, un graduador humano 520 puede como alternativa derivar la imagen LDR basándose en una gradación ajustada óptimamente adicional empezando desde la gradación maestra como se codifica en la imagen HDR maestra IM_MSTR_HDR (es decir, él, por ejemplo, comienza con la imagen [0, 1,0] como si fuera LDR y libremente empieza a transformar la misma coloriméricamente de acuerdo con cualesquiera que sean sus gustos). Es decir en este escenario, existe tanto una codificación del aspecto óptimo para sistemas de renderización HDR en la imagen HDR IM_MSTR_HDR, como otro aspecto óptimo para sistemas LDR, codificados en la imagen graduada LDR (por ejemplo, ventajosamente de 8 bits) Im_1. Aunque nuestros métodos son aplicables a cualquier transformación de color o luma para objetos entre una primera y segunda definición de espacio de color para los colores de píxel, habitualmente de diversa profundidad de bits y/o brillo máximo de visualizador previsto (la única cosa necesaria que existe una buena relación de orden entre regiones más luminosas y más oscuras regiones en al menos una de las codificaciones de representaciones de color/imagen), nos centraremos en nuestra ilustración en un ejemplo del segundo tipo. Es decir, el graduador puede haber especificado muchos mapeos de color ajustados, es decir, de luma_in general a forma funcional de luma-out (por ejemplo, como una LUT) para diversos subobjetos o regiones de la imagen o imágenes, que en nuestras estrategias se convertirán en una serie (uno o más) de valores de gris de diferenciador de regiones (gTS), un número de funciones o algoritmos de transformación, y habitualmente también una o más reglas de vinculación, segmentos obtenibles de vinculación con transformaciones a aplicar (por ejemplo, si existen tres valores de gris de diferenciador de regiones sucesivos, los objetos por debajo del primer gTS1 pueden segmentarse como "0", por encima de gTS1 como "1", y a continuación si el segundo gTS2 es un valor mayor aplicable al mismo conjunto de objetos (es decir, el mismo rango de luma), por encima de ese gTS2 serán segmentos "2", pero las lumas por debajo ya pertenecen a "1". Si el gTS2 es solo una redefinición de objetos más oscuros y más luminosos, como en nuestro ejemplo de objetos muy oscuro, las lumas por encima del umbral serán en ambos casos segundos planos de segmento "1", pero las lumas inferiores serán segmento "0" respectivamente segmento "2". Si las relaciones están claras, no necesitan codificarse datos adicionales, pero habitualmente pueden existir metadatos adicionales que explican el significado de los valores de gris de diferenciador de regiones. Por ejemplo puede ser suficiente simplemente definir el tipo de valores de gris de diferenciador de regiones como "further_demarcation_in_same_luma_range", o "modified_demarcation", etc. pero para los casos más complejos, y de hecho porque no se necesitan muchos más datos adicionales un codificador puede elegir hacerlo siempre así, se puede solo co-codificar lo que tiene que hacerse para cada situación, con, por ejemplo, reglas de asignación de valores de segmento. Por ejemplo "if luma < gTS1 → object/segment = 0", si "luma > gTS2 → segment = 2", etc. De esta manera se garantiza una posible mala interpretación y transformación incorrecta resultante.

La Figura 3 aclara un posible ejemplo ventajoso de cómo codificar las realizaciones anteriores, adaptando el marco de las tecnologías de codificación de imágenes actuales, tal como, por ejemplo, una norma de codificación de video MPEG como AVC o similar.

Se puede comenzar poniendo algunos de los metadatos en un encabezamiento global Gen_Im_HDR (por instantánea, o para la primera instantánea de una toma de instantáneas, por ejemplo), que es habitualmente útil para transformaciones predominantes. Por ejemplo, para tomas simples, puede ser suficiente codificar en este documento un primer mapeo Glob_TM1 a aplicar a la mayoría de la instantánea, y un segundo mapeo global Glob_TM2 a aplicar a alguna región, por ejemplo, mucho más luminosa. El primer mapeo podría aplicarse a nuestra sala de la Figura 1 (es decir, siendo todo el segundo plano, escalinata, y objeto muy oscuro), y a continuación el segundo mapeo podría aplicarse para abrillantar/potenciar el vitral. Y la diferencia entre estos dos objetos se encuentra rápidamente en el lado de recepción por medio del valor de gris de diferenciador de regiones gTS_glob codificado (habitualmente esta ventana tendrá lumas (mucho) mayores en la imagen LDR Y-8b que el resto del objeto, pero sin estos metadatos, esto puede ser muy difícil de determinar automáticamente). Si se rota la cámara en la sala, puede, por ejemplo, ser que el vitral empieza a volverse más luminoso porque más del sol brilla a través. Esto puede codificarse variando gradualmente Glob_TM2 y posiblemente gTS_glob para imágenes sucesivas en la toma. Esto permite, por ejemplo, mantener la codificación del vitral en la imagen Y_8b igual a través de las imágenes sucesivas (por ejemplo, usando la mejor posible asignación de código para retener la cantidad máxima de detalle en las pinturas del vitral, porque se puede potenciar el brillo de ventana a través del mapeo variable Glob_TM2 (es decir, los cambios de iluminación son el la transformación funcional el lugar de la codificación de color de textura pixelizada). A continuación se codifica un número de bloques de datos de píxeles, por ejemplo, a través de una DCT. Si la codificación global es suficiente para toda la imagen, a continuación todos los datos de píxeles siguen ese encabezamiento global, hasta el fin de la toma, o incluso clip de película. Sin embargo, suponemos en este ejemplo que tenemos el escenario más complejo en el que en algún sitio en la imagen, empezando antes de un bloque particular (i-1, j-2), tenemos que empezar haciendo transformaciones locales. Es decir, habitualmente podemos aún usar el conocimiento de transformación global como

ya codificado Glob_TM1 etc., por ejemplo, para transformar los píxeles de segundo plano del papel pintado, pero tenemos que hacer para al menos un nuevo objeto local una nueva transformación. Es decir, algunas de las estrategias de transformación se redefinirán localmente, por ejemplo, se sobrescribirán. En este ejemplo los metadatos locales Loc_MFT_1 contienen una nueva estrategia para demarcar las partes más luminosas por encima de gTS_L_loc_1, por ejemplo, porque hay un objeto incluso más luminoso ahí como una luz. También, se co-codifica un valor de gris de diferenciador de regiones para determinar uno o más objetos oscuros gTS_D_loc_1. En el ejemplo, el objeto claro puede aún transformarse suficientemente con la transformación aplicable y disponibles en la actualidad para regiones luminosas, pero se codifica un nuevo mapeo Loc_TM_DK para transformar los objetos oscuros (por ejemplo, en este punto para la primera vez la escalinata se produjo, y ya sabemos cómo transformar la ventana y el papel pintado, pero no aún la escalera oscura). Un ejemplo de una regla de vinculación de transformación LnkRL_1 también se co-codifica, esa regla indica que lumas por debajo del valor de gris de diferenciador de regiones de objeto oscuro gTS_D_loc_1 se tienen que mapear con la transformación para la escalinata Loc_TM_DK.

Esta información es suficiente de nuevo para un número de bloques sucesivos (o en general una forma definida general), que contiene segundo plano, o escaleras, hasta que terminamos antes del bloque (i+2,j), en el que tenemos que codificar el valor de gris de diferenciador de regiones gTS_D_loc_2 permitiendo segmentación del objeto muy oscuro, y su estrategia de mapeo Loc_TM_DK_2. DAT_TM proporciona un orden de los datos, tal como, por ejemplo, orden espacial (o temporal en transmisión) de bloques a lo largo de una trayectoria de exploración, como se conoce bien de codificación de imágenes.

Aunque hemos descrito únicamente un ejemplo intercalado, también pretendemos cubrir sistemas en los que los metadatos están físicamente separados de los datos de bloque de píxeles, aunque asociados con posiciones de bloque particulares. Aunque algunas estructuras de codificación de video pueden incluir perfectamente el ejemplo anterior (porque ya se han dedicado, o genérico a usarse a voluntad en los marcadores de posición de memoria de metadatos en su lugar, otras estructuras de codificación de video podrían no tener suficiente memoria de metadatos para almacenar todo, o perder compatibilidad hacia atrás confundiendo sistemas antiguos si algunos datos se escriben intercalados. Por lo tanto, otras realizaciones equivalentes podrían codificar todos los metadatos (gTS, etc.) en una parte separada de la señal (por ejemplo, en el comienzo de una película en un disco, o en intervalos regulares durante la difusión, etc.), y a continuación hacer esos datos asociables por medio de un código de asociación geométrica con bloques particulares u otras regiones. La forma más simple de hacer esto es escribir el número del bloque (y potencialmente también el número de imagen, película/número de contenido, etc.) después de los datos, por ejemplo, como: "gTS1= 230 / ImageNr/TimeCode = 2541, Block_loc= (X=20, Y=12)". Esa parte de metadatos separada puede incluso, por ejemplo, residir en una señal diferente, y suministrarse a través de medios diferentes, por ejemplo, la película se pone en un blu-ray en el reproductor, pero los metadatos "que explican" los mismos como los valores de gris de diferenciador de regiones se recuperan desde un almacenamiento de red (por ejemplo, un servicio especializado que permite renderización HDR mejorada) a través de, por ejemplo, internet.

La Figura 4 muestra cómo se verá una segmentación habitualmente, en cuyo ejemplo también hemos explicado cómo, por ejemplo, el vitral puede subdividirse en subregiones, que sería útil, por ejemplo, si la parte inferior se ilumina menos luminosamente, por ejemplo, debido a partes del exterior que bloquean algo de la luz. En ese caso, resultará un nuevo tipo de segmento SEGM_TYP_2, por ejemplo, segmento = "5". Entendemos ahora cómo las reglas de segmentación mediante comparación simple con los valores de gris de diferenciador de regiones óptimamente determinados pueden generar fácilmente objetos de diferente luminosidad (habitualmente diferente iluminación en la escena HDR) que puede segmentarse con precisión, independientemente de su relación con los bloques. Por tanto, se pueden codificar todos los otros datos útiles, tal como mapeos a usarse en la actualidad, en una base de bloque, mientras los resultados se aplican a precisión de objeto, es decir, sin ningún artefacto de halo o bloque, etc.

Queremos decir un poco más acerca de los valores gTS. Ya hemos mencionado que pueden definirse independientemente de la codificación de luma técnica que se use, por ejemplo, se puede usar valores gTS de luma en un espacio de color YCrCb de gamma 2,2, o valores de diferenciador Y de luminancia en una codificación XYZ de los colores de imagen etc. permanece la pregunta interesante de si los valores gTS se definen en el espacio de color de referencia de la primera o la segunda imagen, por ejemplo, la imagen inicial o final. Si se usa una representación LDR mapeada para codificar un grado maestro de HDR, se recuperaría esa imagen HDR mediante mapeo de escalado ascendente de luminancia desde esa imagen LDR. Por tanto, tendría sentido definir los valores gTS a lo largo del eje de luma de codificación de imagen LDR, aunque en principio en situaciones usuales también se podría especificar los mismos a lo largo del eje de luma HDR, ya que a través de la inversa de las funciones de mapeo de recuperación de HDR esos valores basados en HDR gTS podrían convertirse en valores basados en LDR equivalentes gTS. Habitualmente, metadatos especificarán en el comienzo de la codificación de video qué definición es aplicable. Ahora, profundicemos un poco en qué puede suceder en algunos escenarios para valores basados en LDR gTS. En principio uno podría tener un mapeo a partir de HDR maestro a la segunda imagen LDR que (ligeramente) solapa los histogramas de luma de regiones que en la imagen HDR original estaban separados (por ejemplo, algunas de las partes más oscuras de la escalera pueden obtener luminancias en LDR que también se producen en el objeto muy oscuro). Podríamos a continuación especificar el diferenciador de gTS a la mitad de las colas de histogramas superpuestos, o en una mejor posición de luma. Aunque en principio, podría haber problemas cuando se escala ascendientemente, esto no necesita ser un problema para varias estrategias de mapeo, en particular sin tienen un comportamiento relativamente suave alrededor del solapamiento (es decir, no potencia contraste entre píxeles). Sin

embargo, a continuación nos limitaremos a sistemas que normalmente deberían tener histogramas separados tanto en la imagen LDR como HDR. Las diversas realizaciones de aparato pueden limitarse para tener esta limitación en cuenta, por ejemplo, limitando la elección de mapeo de HDR a LDR que puede elegir un graduador, etc. Esto será fácil para codificaciones no comprimidas (por las que queremos decir presión espacial psicovisual como técnicas de frecuencia como DCT, y no compresión de un subrango de luma). Para codificaciones comprimidas debemos ser un poco más cuidadosos con problemas como, por ejemplo, las estructuras de tablero de ajedrez de codificación de DCT incompleta. Aunque tal necesidad no siempre es un problema en la práctica, en ocasiones el artefacto puede volverse más severo visualmente, por ejemplo, pareciendo un área más ruidosa. En particular esto puede suceder si en el LDR sin comprimir original, se separasen los histogramas de escalera y segundo plano (quizás tocando, con algunos códigos no usado entre medias), pero después de descomposición de base DCT, señales recuperas tendrían algunos puntos de tablero de ajedrez más oscuros a partir del ambiente más luminoso que se encuadran dentro del subrango asignado para la escalera más oscura. Si también se tiene una curva de mapeo de tono que se estira mucho a lo largo de ese valor gTS entre escaleras y segundos planos (por ejemplo, una función discontinua con una gran compensación entre las dos partes de mapeo de tono) a continuación podría ser que esos puntos se vuelvan significativamente más oscuros en el segundo plano al menos cerca de la escalera. Diversas realizaciones de aparato (o método) pueden ocuparse de eso de varias formas, y en particular una interfaz de usuario puede ofrecer al graduador diferentes formas de interactuar y especificar este comportamiento de codificación. En primer lugar, puede hacer la curva de mapeo de tono menos inclinada, y el aparato puede o bien inicialmente ofrecerle a él una elección de mapeos únicamente menos inclinados, o bien puede corregir iterativamente (al menos una iteración) los mismos, ofreciendo al graduador que especifique de nuevo los mapeos únicamente para regiones en las que él evalúa que el artefacto es demasiado severo. También, el mapeo puede ser de tal forma que hay algunos códigos de repuesto. En particular, se puede definir fácilmente tal comportamiento con dos valores gTS. La Figura 7 ilustra esquemáticamente un escenario de este tipo. En este gráfico Luma_in será las luminancias de la imagen HDR, y luma_out de la correspondiente codificación LDR de las mismas, que enviaremos a través de un codificador MPEG heredado, por ejemplo, en la imagen HDR, las regiones brillantes son de luma/luminancias muy separada de las regiones oscuras, que muestra desde su separación a lo largo de la luma en axis. En teoría podríamos diseñar un mapeo que hace que se toquen a lo largo del eje de Luma out (LDR), pero ahora diseñamos un mapeo que deja algún rango ProtRng de códigos vacíos entre ellos. Estos códigos no deberían existir en la codificación de LDR, pero después de descompresión de DCT algunas de las partes más oscuras de estos tableros de ajedrez se encuadran dentro de este ProtRng. El decodificador sin embargo puede reconocer esto, y eliminar las mismas de la señal, por ejemplo, recortando las mismas al menor valor de Luma_out del rango de brillo, antes de hacer el escalado ascendente de luminancia para recuperar la imagen HDR. Con este principio podríamos incluso reducir ese rango protector ProtRng, incluso hasta tal extremo que algunos códigos después de descompresión de DCT pueden encuadrarse dentro del rango oscuro de la imagen LDR, y mapearse (potencialmente lejos del rango de brillo en la imagen HDR) mediante la inversa del mapeo oscuro MapDrk en lugar del mapeo correcto para esos píxeles, a saber el mapeo luminoso MpBrght. Pero tales artefactos de DCT normalmente tienen una estructura que va desde un par de valores intermedios a los puntos más oscuros en el tablero de ajedrez. Por tanto, el decodificador puede, por ejemplo, detectar a partir de esos algunos valores incorrectos en el bloque que puede haber un problema potencial, y cambiar después de descompresión de DCT pero antes de escalar ascendentemente de luminancia tales valores a valores en el rango de brillo de la imagen LDR, incluso si esos píxeles realmente píxeles del objeto oscuro, solo para estar en el lado seguro (un efecto de HDR ligeramente incorrecto, pero tampoco artefacto potencial fuerte). El codificador puede usar un código reservado para este "recorte a rango" (0 o 1), para indicar si esto debería suceder a un bloque, o si debería dejarse solo y solo escalar ascendentemente, y el graduador puede indicar bloques problemáticos clicando, por ejemplo, en los mismos con su ratón o garabateando a través de un conjunto conectado de bloques problemáticos. Aunque el decodificador puede no conocer la diferencia, el codificador puede, teniendo la señal original y toda la información, determinar si un problema de este tipo puede suceder, por tanto puede hacer un modo de pseudocolor con el que el graduador puede alternar entre los píxeles incorrectos mostrados como, por ejemplo, rojo saturado brillante, frente a su color real después de reconstrucción (incorrecta) de la imagen HDR. Varias otras opciones (interactivas) están disponibles también, por ejemplo, el codificador puede usar más palabras de código de DCT para bloques que se seleccionaron como problemáticos por el graduador, o a la inversa menos bloques de DCT, de modo que existe aún un menor error de frecuencia, pero los patrones de ajedrez rápidos se eliminan en caso de que eso proporcione un mejor aspecto final. O, por ejemplo, puede hacerse un pequeño cambio a los datos originales o coeficientes de DCT, por ejemplo, un patrón de contador puede aplicarse al bloque LDR antes de que se codifique por DCT, de modo que los menores valores de tablero de ajedrez ya no se encuadran dentro del rango LDR oscuro, etc.

La Figura 5 muestra un ejemplo de un posible aparato de gradación 510 en un lado de creación de contenido, controlado por un graduador 520 (el experto entenderá cómo las mismas realizaciones de realización de nuestra invención se aplicarían en, por ejemplo, un aparato de transcodificación automático basado en colorimetría matemática, o cualquier otra realización).

Dentro del aparato de gradación 510 hay un codificador de imagen 549 para codificar una imagen de una escena de alto rango dinámico, cuya imagen puede haberse capturado anteriormente, por ejemplo, mediante una película de celuloide o un sistema digital de cámara electrónica, a la que se han añadido efectos especiales, y que en caso de video puede ser una imagen en una secuencia de composición temporal final. El codificador de imagen (que ahora por simplicidad asumimos que es una unidad como un IC, pero puede un paquete de software, en el que algunos componentes pueden incluso ejecutarse en un servidor remoto, etc.) puede habitualmente comprender diversos

subcomponentes (habitualmente bajo el control de software, que permite al graduador elegir parámetros), y habitualmente comprenderá alguna variante de una unidad de codificación de textura de píxel 552, que se dispone para codificar los colores de píxeles de la imagen de acuerdo con una representación de imágenes definida particular (Im_1), por ejemplo, con luminancias de palabras de código de N bits, como, por ejemplo, palabras de código de 8 bit o 10 bit o 12 bit, y codificaciones de crominancia como Y_Cr y Y_Cb. Ya que ya existen varias variantes de codificación, que abarca VC1, VP8, y codificaciones de tipo MPEG similares, hasta incluso codificadores fractales menos populares, no necesitaremos aclarar adicionalmente ese aspecto.

Se comprende también sin embargo una unidad de análisis de imagen 550 que puede aplicar análisis de imágenes más simple o más complejo. En tales aparatos de gradación profesionales como se muestran en el ejemplo, habitualmente muchos de los algoritmos embebidos en software están disponibles, dando al graduador casi control total de la imagen, tanto cuando espera estudiar sus propiedades y composición, como cuando quiere cambiar la misma arbitrariamente. Puede, por ejemplo, usar una pipeta para muestrear un color particular (y puede a continuación definir a partir de ese color de píxel muestreado un "color de objeto" típico, por ejemplo, eligiendo límites colorimétricos apropiados alrededor del color muestreado), o mirar a formas de onda de señal, o histogramas, u otras representaciones de regiones (por ejemplo, el sistema puede mapear subrangos de luma encima de una región, por ejemplo, mediante pseudocolores). Puede, por ejemplo, (temporalmente) iluminar una región particular, para inspeccionar visualmente más claramente su textura en uno o más visualizadores de referencia 530. Puede habitualmente aplicar un número de procesamiento de imagen, como ajustar la nitidez de una región, o aplicar un efecto de iluminación u otro efecto. Puede demarcar un objeto dibujando un límite alrededor del mismo con un lazo, etc.

Ahora habitualmente la unidad de análisis de imagen convertirá al menos un objeto en un valor de gris de diferenciador de regiones (gTS), o en otras palabras asociará con un objeto al menos un determinado gTS relacionado. Puede, por ejemplo, determinar el histograma de la región de objeto seleccionada, y determinar que el valor de luma mínimo que contiene es mayor que el de la región circundante, por ejemplo, toda la imagen. Puede implicarse tratamiento interactivo, por ejemplo, el graduador puede primero abrillantar la región, de forma que su nuevo valor mínimo es mayor que el mayor valor en el resto de la imagen, o una parte relevante de la misma relacionada geoméricamente con el objeto (por ejemplo, bordeando el objeto).

Este valor de gris de diferenciador de regiones gTS se emitirá a un formateador 554 que puede co-codificar (de nuevo siguiendo las reglas de alguna norma de codificación de imágenes, heredada o nueva) en una señal de imagen de salida (S(Im_1, MET(gTS))) la representación de imagen (Im_1) y el valor de gris de diferenciador de regiones (gTS), este último habitualmente en un formato de metadatos textual preacordado. Por ejemplo que señal de imagen puede grabarse en un disco Blu-ray 511, o grabarse en alguna otra memoria como el disco o memoria de estado sólido de un servidor de red, o transmitir la señal de imagen en tiempo real de una conexión de transmisión de señal, etc. Debería estar claro para el experto que aunque describimos esta funcionalidad en la presente construcción física, son posibles otras realizaciones.

Habitualmente cuando se gradúa en el aparato de gradación el graduador usará simultáneamente su unidad de determinación mapeo de luma 553 para determinar un mapeo de luma (TOM) para al menos algunos de los objetos (teniendo por supuesto también los otros objetos a continuación una transformación, quizás una transformación de identidad, pero esa transformación puede predefinirse, por ejemplo, como por defecto, o elegirse por el visualizador de renderización etc.). Definirá un mapeo entre lumas de píxeles como se codifica en una primera representación de imagen (por ejemplo, una imagen HDR Y_16b de entrada) y lumas de los píxeles en una segunda representación de imagen (por ejemplo, una imagen LDR Im_1), o a la inversa. La unidad de determinación de mapeo de luma 553 puede determinar matemáticamente una función de mapeo por sí misma, por ejemplo, proponiendo la misma como una sugerencia inicial al graduador, mirando las propiedades visuales de las diversas regiones de, por ejemplo, la imagen HDR, y cómo pueden representarse aún razonablemente en una codificación LDR. Esto puede resultar en aplicar, por ejemplo, un mapeo sigmoidal o multisegmento, con las curvaturas de las rodilla y hombros, por ejemplo, determinado aislando sublóbulos particulares del histograma global, o entendimiento de imagen tal como detección facial o cualquier variantes de los mismos. El graduador puede a continuación ajustar esta función, por ejemplo, desplazando o doblando el hombro del sigmoide. En nuestro método él puede hacer esto en relación con los valores gTS. Por ejemplo, el aparato de gradación puede ya definir un valor gris importante (por ejemplo, 999) que puede ser un punto de control para, por ejemplo, una parte de una curva de mapeo multisegmento, pero el graduador puede a continuación mejorar este punto, por ejemplo, desplazar el mismo de modo que una porción más relevante de un objeto como, por ejemplo, la escalera se transforma ahora mediante un mapeo de luma/(tono) parcial. Ilustramos algunos aspectos adicionalmente con el ejemplo de la Figura 8. Como ya se ha mencionado, podemos usar nuestro método en un mero método de codificación, por ejemplo, en el que una imagen HDR tiene que codificarse a través de una imagen LDR utilizable heredada y codificada (contenedor LDR). En esa situación únicamente habrá algunas funciones fijas para mapear entre las dos imágenes. Con la Figura 8 describimos sin embargo cómo nuestro sistema puede usarse con los valores gTS en un escenario adicional de capacidad de ajuste de visualizador, en el que se determinan grados adicionales para diferente visualizadores, si esta información ya se graduó (es decir, el graduador al menos comprobó cómo se verían tales transformaciones en un par de muy diferentes visualizadores de referencia como HDR, sub LDR con un rango dinámico pequeño) y codificó en la señal de imagen habitualmente como diversas funciones a aplicarse en uno o más imágenes de textura (Im_1), o si únicamente había datos codificados para un grado HDR de buen

aspecto y quizás un buen grado LDR, y en el lado de renderización un sistema de visualizador (por ejemplo, visualizador u ordenador) está determinando basándose en esos datos y nuestros valores gTS al menos un grado adicionalmente para, por ejemplo, un visualizador MDR de rango dinámico medio. En este gráfico, usamos una representación de luminancia final que es absoluta. Luminance_in puede ser la señal de entrada como se define como se vería en, por ejemplo, algún visualizador de referencia, y luminance_out puede ser las luminancias renderizadas de salida en diversos visualizadores reales con diferentes capacidades de brillo. Suponemos que los objetos de menor luminancia se codifican en gran parte correctamente y, por tanto, renderizados, por tanto ambos visualizadores (Dis1, Dis2) usarán el mismo mapeo de tono TM_FDrk, que puede ser una transformación de identidad, o algún estiramiento de contraste. Ahora por encima de gTSh1 comienzan las regiones brillantes en la imagen, y existen dos regiones brillantes (por ejemplo, iluminada estableciendo la salida del sol a gTSh2, e iluminada por la fuerte iluminación de un estadio de fútbol por encima de gTSh2). El visualizador 1 puede tener un brillo máximo muy alto, por tanto tenemos mucho margen para asignar subrangos de brillo de la misma a diversas clases de iluminación visual. Primer mapeo de tono de procesamiento de brillos TM_TBri1_Dis1 puede para ese visualizador de brillo estirar considerablemente los datos originales, de modo que esa región se ve bien luminosa y con contraste. Un segundo mapeo de tono de procesamiento de brillos TM_TBri2_Dis1 puede incluso compensar esa región con luminancias renderizadas muy altas, de modo que visualmente esa región es muy diferente de las partes iluminadas por el sol, por ejemplo, la iluminación del estadio se ve muy áspera. Esta discriminación puede hacerse fácilmente con los valores gTS (por ejemplo, en el ejemplo de este mapeo lineal pueden incluso parametrizar la función de mapeo). Para un visualizador con menor brillo máximo, el ordenador, por ejemplo, que determina el mapeo final puede hacer otra cosa para las diversas regiones determinadas por los valores gTS. Por ejemplo puede procesar los menores brillos con una función de mapeo de menos contraste TM_Bri1_Dis2, de modo que aún hay margen para las regiones iluminadas por las luces del estadio, que sin embargo necesitan recortarse suavemente con la función TM-Bri2_Dis2.

Este mapeo de luma (TOM) se co-codifica finalmente en la señal de imagen de salida (S (Im_1, MET(gTS), TOM) mediante el formateador 554, de acuerdo con señal de imagen acordada que define especificaciones. De nuevo, un mapeo de este tipo puede mapear en principio desde usar cualquier primera especificación de codificación de color que determina cualquiera primer imagen para cualquier primer visualizador de referencia (en particular con cualquier rango dinámico de luminancia de salida) a de forma similar cualquier segunda codificación de imagen (en particular con mayor o menor brillo máximo, menor o mayor rango dinámico, etc.), siempre que se especifique y acuerde claramente con el lado de recepción.

Habitualmente un codificador de imagen 549, que de acuerdo con los conceptos de realización se dispone para co-codificar valores de gris de diferenciador de regiones (gTS) inteligentemente elegidos, es útil para demarcar regiones de brillo promedio de regiones de alta brillo, es decir, por ejemplo, la parte del histograma de luma por debajo de un cierto percentil, y un porcentaje de los valores más altos, especialmente cuando se separan mediante códigos no usados (o una definición similar basándose en luminancias renderizadas). Esto es, por lo tanto, muy útil para codificación de escena HDR, en cualquier formato/espacio de color la imagen se codificará finalmente en al menos una versión (por ejemplo, Y_8b o Y_10b, o Y_16b, y definiciones adicionales como luminancia de blanco prevista, negro, curva de gamma, etc.), ya que estas escenas HDR habitualmente no tienen luminancias de objeto de escena similares y, por lo tanto, después de que la cámara captura las lumas de imagen en vista de la iluminación uniforme usada mediante los diseñadores de iluminación, como en producción LDR, sino que tienen regiones de iluminación muy diferentes. Y los valores gTS pueden caracterizar adecuadamente las mismas.

Por tanto, básicamente el graduador solo aplica sus operaciones clásicas en la imagen o imágenes como selección de objeto, definiendo las curvas de mapeo óptimas para diferentes partes (subrangos de luma típicos) de ese objeto, etc., y el codificador 549 traduce eso en parámetros de las realizaciones de esta invención, tal como valores de gris de diferenciador de regiones gTS.

Hemos aclarado la invención en la Figura 5 con un sistema de producción de contenido de entretenimiento doméstico, por ejemplo, teniendo acceso a un servidor de video 580 a través de una conexión 581, que contiene archivos de video, tal como una gradación maestra HDR IM_MSTR_HDR de digamos alguna película o programa de televisión, que se produjo en el momento de producir la película o programa, como siendo la gradación de referencia definitiva. Esto se convertirá a continuación a una gradación de cine doméstico para un lanzamiento de versión doméstica, por ejemplo, como en una imagen MPEG-AVC de 8 bits Im_1, y los metadatos de acuerdo con cualquiera de las realizaciones presentadas. Por supuesto, el codificador puede también incorporarse en otro sistema, aparato o escenario de uso, por ejemplo, para determinar uno o más grados maestros directamente desde una señal de cámara en bruto desde la cámara 501 a través de conexión (por ejemplo, inalámbrica) de imagen/video 505, o para remasterización, etc.

La Figura 6 muestra una posible realización de un sistema de lado de recepción, a saber un sistema de renderización de imagen o video de consumo doméstico. Una televisión 602 puede recibir directamente una primera señal de video (o imagen) SB (Im_1, MET) a través de, por ejemplo, las vías aéreas. Este suministro de video de ejemplo ya se explicó anteriormente, y usa los valores de gris de diferenciador de regiones en entre series de imágenes (habitualmente yendo de un contenido a otro o partes de un programa como reportajes en un programa de noticias) que o bien deberían renderizarse cinematográficamente, con alto brillo y precisión HDR (es decir, también en cuanto a asignación precisa en el eje de luminancia de salida de varios objetos que determina el aspecto de la imagen), o

bien debería ejecutarse (cerca de LDR) con brillo (y potencia) reducido. También puede haber un aparato de procesamiento de video 601 (como, por ejemplo, un decodificador de salón o PC), que puede conseguir su video a través de una o más conexión a la internet (I_net). Por ejemplo, un servidor de YouTube o similar puede suministrar una señal HDR, que preferentemente es tanto simplemente codificada como utilizable de una manera versátil para diversos diferentes posibles visualizadores de renderización (el así llamado criterio de "capacidad de ajuste de visualizador"). Además de, por ejemplo, una codificación Y_8b Im_1 de la señal HDR, contendrá uno o más de los metadatos de realización anteriores, y, por ejemplo, también un indicador de procesamiento PROC_IND, que especifica cómo puede procesarse esta imagen Im_1, por ejemplo, para obtener una versión de imagen HDR. Por ejemplo, puede especificar que el lado de recepción puede usar varias estrategias de transformación de color/luma, por ejemplo, con un indicador como "receiver_determines_optimal_mapping". En ese caso el aparato de recepción como el decodificador de salón o tv puede determinarse a sí mismo para aplicar un primer mapeo, por ejemplo, si el espectador tiene las luces encendidas en su sala de visualización, y un segundo mapeo si las luces están apagadas. De hecho, procesamiento permitido puede especificarse en términos de tolerancias o cambios porcentuales, por ejemplo, un aparato de lado de renderización puede permitirse que aplique una función gamma de hasta 1,2, pero no más fuerte a un grado, por ejemplo, si el visualizador tiene un brillo máximo dentro de un rango de el del visualizador de referencia (por ejemplo, el grado puede determinarse para un visualizador de referencia de 700 nits, y permitirse que se pueda modificar ligeramente si el visualizador real está dentro del rango del 50% del mismo, es decir, con un brillo máximo de entre 350 y 1050 nits). El indicador de procesamiento puede también especificar que únicamente puede usarse uno o uno de un par de transformaciones específicamente determinadas, etc. El indicador puede tener una definición variable, que puede volverse compleja, por ejemplo, puede comprender guías detalladas para control de interfaz de usuario, guiando al espectador a través de selecciones para tener una visión óptima de la película (proporcionándole algunas opciones de optimización aprobadas por el creador, como un par de formas de mejorar los oscuros, haciendo los mismos más coloridos, pero en cierto modo disminuyendo el ambiente de la imagen), según desee el creador de contenido (calibración manual con subconjuntos seleccionados de imágenes por ejemplo), etc. Normalmente habrá escenarios de repliegue, ya que el espectador tiene el control definitivo, por tanto, estas guías pueden ignorarse o anularse, pero las presentes realizaciones permiten un alto grado de versatilidad, como, por ejemplo, un comentario más cercano por el creador de contenido sobre cómo tiene que renderizarse su contenido en el entorno de renderización final (ya sea visualizador doméstico, de cine, en exteriores, profesional en, por ejemplo, un estadio de fútbol, etc.).

El decodificador de imagen 605 habitualmente puede comprender algunas de las siguientes unidades. Una unidad de decodificación de textura de píxel 608 necesita disponerse de modo que pueda realizar cualquier cálculo matemático necesario para decodificar las señales de imagen entrantes, que pueden codificarse de acuerdo con muchas normas, de forma que, por ejemplo, software puede ejecutarse, que puede actualizarse si se lanza un nuevo codificador de onduladas. Por supuesto, habrá desempaquetamiento de señal y quizás demodulación, etc. (que habitualmente se harán mediante un desformateador 607, junto con extracción, y potencialmente también decodificación de los metadatos como los valores de gris de diferenciador de regiones), pero la unidad de decodificación de textura de píxel 608 será capaz de hacer tales cosas como, por ejemplo, decodificación aritmética, decodificación de DCT inversa, etc., como todos los componentes en normas de MPEG visual y similares. Una unidad de segmentación de imagen 606 hará la segmentación, y como se ha dicho, eso puede hacerse fácilmente mediante umbrales desde los valores gTS, pero también pueden soportarse estrategias de segmentación más complicadas. A continuación una unidad de transformación de color de píxel 609 realizará el mapeo de al menos las lumas de píxeles, que puede ser tan simple como grabar, por ejemplo, el valor de salida de la función PD_BLACK(i+2,j) que pertenece al valor de luma de píxel de ese píxel particular Im_1 como valor de entrada (Luma in). Este valor de salida se escribirá en la imagen de salida HDR IM_RC_HDR en esa ubicación de píxel. Esa imagen puede ser una enviada a través de una conexión 688 a la TV (por ejemplo, para accionamiento directo o procesamiento adicional por una unidad de procesamiento de imagen 620 en la TV o visualizador en general, que también es capaz de hacer transformaciones de color).

Puede haber una imagen intermedia IM_INTRM implicada, y aunque esto puede ser cualquier representación de referencia, asumimos en la actualidad para simple explicación que es una imagen de luma de 8 bits (también con palabras de 8 bit para dos representaciones de canal de color). Si la representación de imagen de entrada Im_1 no se comprime (por ejemplo, DCT), entonces esto puede ser una simple copia de Im_1, y de otra manera habitualmente es la imagen resultante de la descompresión.

El sistema también muestra distribución de red doméstica a través de un medio de enlace de comunicación de red como la antena 699 a un visualizador portátil 630 (por ejemplo, un IPAD que usa el espectador para continuar viendo la tv en su cama en su habitación). Este ilustra bien la versatilidad de las realizaciones, ya que el aparato 601 puede a continuación precondicionar otra señal de imagen IM_RC_MDR óptimamente para este dispositivo, que puede, por ejemplo, únicamente tener un rango dinámico medio, entre LDR (que podemos definir extremos aproximadamente por encima de 750 nits de brillo máximo) y HDR de alta calidad, que comienza por encima de digamos 2500 nits. La imagen MDR puede a continuación codificarse en IM_RC_MDR usando incluso la misma Im_1 para las texturas de píxel, y los mismos valores de gris de diferenciador de regiones pero funciones de mapeo cambiadas para mapear al rango diferente de luminancias de salida renderizadas de visualizador.

Las presentes realizaciones también permiten mejorar la interactividad de interfaz de usuario en el lado de renderización, ya que el espectador puede, por ejemplo, ajustar sus funciones de mapeo paramétricamente. Por

ejemplo, aumentar el brillo del objeto muy oscuro puede ser tan simple como controlar la inclinación de la función PD_BLACK(i+2, j). Algoritmos inteligentes pueden aplicar modificaciones de luma coordinadas a todos los objetos en la imagen o imágenes en sincronía estética en el toque de un único botón (habilitando, por ejemplo, una función de brillo inteligente), pero también es posible habilitar control de los diversos objetos definiendo una interfaz de usuario más compleja. Por ejemplo, cuando ve la t.v., el usuario puede usar su visualizador portátil 630 como un control remoto, y tener una copia de la imagen de la TV en el visualizador de ese control remoto, con los diversos objetos significativos ya preseleccionados con los métodos de valor de gris de diferenciador de regiones. El espectador puede a continuación indicar rápidamente con un par de cambios (por ejemplo, algunos deslizadores apareciendo encima de los objetos) su renderización preferencial en una o un par de imágenes (por ejemplo, en el inicio de la película, algunas escenas de características importantes o en una orden de pausa, la imagen actual para que se reproduzca la escena). Un botón de deshacer puede restaurar la situación, etc. Puede usarse inteligencia artificial para deducir las preferencias de los espectadores a partir de sus acciones, incluso almacenando especificaciones para programas no relacionados en momentos de reproducción muy diferentes como en días diferentes. El sistema puede, por lo tanto, inferir que al espectador le gusta sus negros muy intensos, o a la inversa con más brillo, y a continuación aplicar este conocimiento a otras imágenes.

Los componentes algorítmicos desvelados en este texto pueden realizarse (completamente o en parte) en la práctica como hardware (por ejemplo, partes de un CI específico de aplicación) o como software que se ejecuta en un procesador de señales digitales especial, o un procesador genérico, etc. Pueden ser semiautomáticos en el sentido que al menos alguna entrada de usuario puede estar/haber estado (por ejemplo, en fábrica, o introducida por consumidor, u otra entrada humana) presente.

Debería ser entendible para el experto en la materia a partir de nuestra presentación qué componentes pueden ser mejoras opcionales y pueden realizarse en combinación con otros componentes, y cómo las etapas (opcionales) de los métodos se corresponden con medios respectivos de aparatos, y viceversa. El hecho de que algunos componentes se desvelan en la invención en una cierta relación (por ejemplo, en una única figura en una cierta configuración) no significa que no sean posibles otras configuraciones como realizaciones bajo el mismo pensamiento inventivo según se desvelan para patentar en el presente documento. También, el hecho de que por razones pragmáticas únicamente se haya descrito un espectro limitado de ejemplos, no significa que otras variantes no puedan caer bajo el alcance de las reivindicaciones. De hecho, los componentes de la invención pueden realizarse en diferentes variantes junto con cualquier cadena de uso, por ejemplo, todas las variantes de un lado de creación como un codificador pueden ser similares como, o corresponder a aparatos correspondientes en un lado de consumo de un sistema descompuesto, por ejemplo, un decodificador y viceversa. Varios componentes de las realizaciones pueden codificarse como datos de señal específicos en una señal para transmisión, o uso adicional tal como coordinación, en cualquier tecnología de transmisión entre codificador y decodificador, etc. La palabra "aparato" en esta solicitud se usa en su sentido más amplio, en concreto un grupo de medios que permiten la realización de un objetivo particular, y puede ser, por ejemplo, (una pequeña parte de) un CI, o un dispositivo especializado (tal como un dispositivo con una pantalla), o parte de un sistema en red, etc. "Disposición" o "sistema" también se pretende que se use en el sentido más amplio, por lo que puede comprender entre otros, un único aparato comprable físico único, una parte de un aparato, una colección de (partes de) aparatos cooperativos, etc.

La indicación de producto de programa informático debería entenderse que abarca cualquier realización física de una colección de comandos que posibilitan que un procesador genérico o de fin especial, después de una serie de etapas de carga (que pueden incluir etapas de conversión intermedias, tales como traducción a un lenguaje intermedio, y un lenguaje de procesador final) introduzca los comandos en el procesador, para ejecutar cualquiera de las funciones características de una invención. En particular, el producto de programa informático puede realizarse como datos en un soporte tal como, por ejemplo, un disco o cinta, datos presentes en una memoria, datos que viajan mediante una conexión de red -alámbrica o inalámbrica-, o código de programa en papel. Aparte del código de programa, los datos característicos requeridos por el programa pueden materializarse también como un producto de programa informático. Tales datos pueden suministrarse (parcialmente) de cualquier manera.

La invención o cualquier dato utilizable de acuerdo con cualquier filosofía de las presentes realizaciones como datos de vídeo, puede usarse como señales en soportes de datos, que pueden ser memorias extraíbles como discos ópticos, memorias flash, discos duros extraíbles, dispositivos portátiles escribibles mediante medios inalámbricos, etc.

Algunas de las etapas requeridas para la operación de cualquier método presentado pueden estar ya presentes en la funcionalidad del procesador o cualesquiera realizaciones del aparato de la invención en lugar estar descritas en el producto de programa informático o cualesquiera unidades, aparatos o métodos descritos en el presente documento (con detalles específicos de las realizaciones de la invención), tales como etapas de entrada y salida de datos, habitualmente etapas de procesamiento bien conocidas incorporadas tales como control de visualización convencional, etc. También deseamos protección de los productos resultantes y similares resultantes, como, por ejemplo, las señales novedosas específicas implicadas en cualquier etapa de los métodos o en cualquier subparte de los aparatos, así como cualesquiera nuevos usos de tales señales, o cualesquiera métodos relacionados.

Debería observarse que las realizaciones anteriormente mencionadas ilustran en lugar de limitar la invención. Donde el experto en la materia pueda realizar fácilmente un mapeo de los ejemplos presentados a otras regiones de las

reivindicaciones, no hemos mencionado por concisión todas estas opciones en profundidad. Aparte de las combinaciones de elementos de la invención según se combinan en las reivindicaciones, son posibles otras combinaciones de los elementos. Cualquier combinación de elementos puede realizarse en un único elemento especializado.

5
Cualquier signo de referencia entre paréntesis en la reivindicación no se pretende para limitar la reivindicación, ni es algún símbolo particular en los dibujos. La expresión "comprendiendo/que comprende" no excluye la presencia de elementos o aspectos no enumerados en una reivindicación. La palabra "un" o "una" precediendo a un elemento no
10
excluye la presencia de una pluralidad de tales elementos.

REIVINDICACIONES

1. Un codificador de imagen (549) para codificar una imagen (IM_MSTR_HDR) de una escena de alto rango dinámico, que comprende:
- una unidad de codificación de textura de píxel (552), dispuesta para codificar colores de píxel de la imagen como una representación de imagen (Im_1) que comprende palabras de código de N bits;
 - una unidad de análisis de imagen (550) dispuesta para determinar y emitir un valor de gris de diferenciador de regiones (gTS), que es un valor de luminancia que delimita por debajo del mismo luminancias de todos los píxeles de un primer objeto en al menos un bloque de la imagen, y por encima del mismo luminancias de todos los píxeles de un segundo objeto en el al menos un bloque de la imagen; y
 - un formateador (554) dispuesto para co-codificar en una señal de imagen de salida (S) la representación de imagen (Im_1) y el valor de gris de diferenciador de regiones (gTS),
- el codificador de imagen comprendiendo además una unidad de determinación de mapeo de luma (553), dispuesta para determinar un mapeo de luma (TOM) para al menos uno del primer y el segundo objeto que define un mapeo entre lumas de píxeles como se codifica en la representación de imagen (Im_1) y luminancias de los píxeles en una segunda representación de imagen (IM_RC_HDR) en la que una de la representación de imagen y la segunda representación de imagen es una imagen HDR con un brillo máximo establecida fija, y la otra es una imagen LDR, que son dos gradaciones de la escena HDR creadas por un graduador de colores humano o automático, y se dispone para suministrar el mapeo de luma (TOM) al formateador (554) que se dispone para co-codificar el mismo en la señal de imagen de salida (S).
2. Un codificador de imagen (549) de acuerdo con la reivindicación 1, dispuesto para codificar varios valores de gris de diferenciador de regiones (gTSD_Loc_1, gTS_D_Loc_2) en entre secciones de imágenes espaciales que comprenden varias de las palabras de código de N bits que codifican los colores de píxel a partir de la representación de imagen (Im_1).
3. Un codificador de imagen (549) de acuerdo con la reivindicación 1, dispuesto por que codifica un valor de gris de diferenciador de regiones antes de una serie de un número de imágenes sucesivas, siendo un valor de gris de diferenciador de regiones para todas esas imágenes sucesivas.
4. Un codificador de imagen (549) de acuerdo con la reivindicación 1 dispuesto por que codifica al menos un valor de gris de diferenciador de regiones en una memoria no físicamente adyacente a una memoria que almacena la representación de imagen (Im_1), junto con un código de asociación geométrica que permite la asociación de cada respectivo al menos un valor de gris de diferenciador de regiones con una región geométrica de la representación de imagen (Im_1) a la que es aplicable, comprendiendo habitualmente el código de asociación geométrica al menos las coordenadas de un bloque de la representación de imagen (Im_1).
5. Un codificador de imagen (549) de acuerdo con cualquiera de las reivindicaciones anteriores, dispuesto por que codifica un primer valor reservado de un valor de gris de diferenciador de regiones en la señal de imagen de salida (S), indicando que, para al menos una región geométrica de la representación de imagen (Im_1), que se encuentra de acuerdo con una dirección de exploración a través de la imagen más allá de una ubicación identificable con el primer valor reservado, se realiza una transformación desde los valores de píxeles como se codifica en la representación de imagen (Im_1) a valores de píxel en una segunda representación de imagen (IM_RC_HDR), de acuerdo con un algoritmo predefinido.
6. Un codificador de imagen (549) de acuerdo con cualquiera de las reivindicaciones anteriores, dispuesto por que codifica un segundo valor reservado (gTR) de un valor de gris de diferenciador de regiones en la señal de imagen de salida (S), indicando que para al menos una imagen sucesiva, un visualizador debería renderizar la misma con una luminancia de salida máxima por debajo de un valor predeterminado.
7. Un codificador de imagen (549) de acuerdo con la reivindicación 1, en el que la unidad de determinación de mapeo de luma (553) se dispone para determinar varios diferentes mapeos de luma (TOM) para al menos uno del primer y el segundo objeto a través de reglas de enlace de transformación, o se dispone para indicar con un indicador de procesamiento (PROC_IND) que pueden usarse varios diferentes mapeos de luma para transformar los colores de píxel de al menos uno del primer y el segundo objeto a una nueva representación de color de la segunda representación de imagen (IM_RC_HDR).
8. Un decodificador de imagen (605) para decodificar una representación de imagen codificada (Im_1, MET) de una escena de alto rango dinámico, que comprende:
- una unidad de decodificación de textura de píxel (608), dispuesta para obtener de la representación de imagen codificada (Im_1) colores de píxel de píxeles de una imagen decodificada (IM_INTRM); y
 - un desformateador (607) dispuesto para extraer de la representación de imagen codificada (S) un valor de gris de diferenciador de regiones (gTS), que es un valor de luminancia que delimita por debajo del mismo luminancias

de todos los píxeles de un primer objeto en al menos un bloque de la imagen (IM_INTRM), y por encima del mismo luminancias de todos los píxeles de un segundo objeto en el al menos un bloque de la imagen; y para extraer un mapeo de luma (TOM) para al menos uno del primer y el segundo objeto en el al menos un bloque de la imagen; de la representación de imagen codificada (S);

5 - una unidad de segmentación de imagen (606) dispuesta para usar el valor de gris de diferenciador de regiones (gTS) para obtener un segmento de menor luma y un segmento de mayor luma en el al menos un bloque de la imagen decodificada (IM_INTRM); y

10 - una unidad de transformación de color de píxel (609), dispuesta para aplicar una transformación de color (PD_BLACK) como se define en el mapeo de luma (TOM), transformando al menos valores de luma de los colores de píxel, a píxeles en el segmento de menor luma, y dispuesta para aplicar una transformación de color (PM_BLACK) como se define en el mapeo de luma (TOM), transformando al menos valores de luma de los colores de píxel, a píxeles en el segmento de mayor luma, para obtener una segunda representación de imagen (IM_RC_HDR), con lo que una de la representación de imagen y la segunda representación de imagen corresponde a una imagen HDR maestra graduada (IM_MSTR_HDR) y la otra es una imagen LDR.

15 9. Un decodificador de imagen (605) de acuerdo con cualquiera de las reivindicaciones de decodificador anteriores, que comprende una unidad de determinación de transformación (610) dispuesta para seleccionar una estrategia de transformación de color de píxel de una fuente de memoria no asociada con ninguno de los datos de la representación de imagen codificada (S).

20 10. Un decodificador de imagen (605) de acuerdo con la reivindicación 9, caracterizado por que la unidad de determinación de transformación (610) se dispone para determinar la estrategia de transformación de color de píxel sobre la base de al menos un parámetro del entorno de renderización.

25 11. Un decodificador de imagen (605) de acuerdo con cualquiera de las reivindicaciones de decodificador anteriores, dispuesto para obtener la representación de imagen codificada (Im_1) y el valor de gris de diferenciador de regiones (gTS) a partir de memorias separadas físicamente, y se dispone para asociar el valor de gris de diferenciador de regiones (gTS) con una parte geométrica de la representación de imagen codificada (Im_1).

30 12. Un método de codificación de imagen para codificar una (IM_MSTR_HDR) de una escena de alto rango dinámico, que comprende:

- codificar colores de píxeles de la imagen con una representación de imagen (Im_1) que comprende palabras de código de N bits;

35 - determinar y emitir un valor de gris de diferenciador de regiones (gTS), que es un valor de luminancia que delimita por debajo del mismo luminancias de todos los píxeles de un primer objeto en al menos un bloque de la imagen, y por encima del mismo luminancias de todos los píxeles de un segundo objeto en el al menos un bloque de la imagen; y

40 - co-codificar en una señal de imagen de salida (S) la representación de imagen (Im_1) y el valor de gris de diferenciador de regiones (gTS), y un mapeo de luma (TOM) para al menos uno del primer y el segundo objeto que define un mapeo entre lumas de píxeles como se codifica en la representación de imagen (Im_1) y luminancias de los píxeles en una segunda representación de imagen (IM_RC_HDR) en la que una de la representación de imagen y la segunda representación de imagen es una imagen HDR con un brillo máximo establecida fija, y la otra es una imagen LDR, en el que la representación de imagen (Im_1) y la segunda representación de imagen (IM RC HDR) obtenible aplicando una transformación de color que comprende el mapeo de luma (TOM) son dos gradaciones de la escena HDR creadas por un graduador de colores humano o automático.

50 13. Un método de decodificación de imagen para decodificar una representación de imagen codificada (Im_1) de una escena de alto rango dinámico, que comprende:

- obtener de la representación de imagen codificada (Im_1) colores de píxel de píxeles de una imagen decodificada (IM_INTRM);

55 - extraer de la representación de imagen codificada un valor de gris de diferenciador de regiones (gTS), que es un valor de luminancia que delimita por debajo del mismo luminancias de todos los píxeles de un primer objeto en al menos un bloque de la imagen (IM_INTRM), y por encima del mismo luminancias de todos los píxeles de un segundo objeto en el al menos un bloque de la imagen;

- extraer un mapeo de luma (TOM) de la representación de imagen codificada; para al menos uno del primer y el segundo objeto en el al menos un bloque de la imagen;

60 - segmentar el al menos un bloque de la imagen usando el valor de gris de diferenciador de regiones (gTS) para obtener un segmento de menor luma y un segmento de mayor luma en la imagen decodificada (IM_INTRM); y

65 - aplicar una transformación de color (PD_BLACK) como se codifica en el mapeo de luma (TOM), transformando al menos valores de luma de los colores de píxel, a píxeles en el segmento de menor luma, y aplicar una transformación de color (PM_BLACK) como se codifica en el mapeo de luma (TOM), transformando al menos valores de luma de los colores de píxel, a píxeles en el segmento de mayor luma, para obtener una segunda representación de imagen (IM_RC_HDR), con lo que una de la representación de imagen y la segunda representación de imagen corresponde a una imagen HDR maestra graduada (IM_MSTR_HDR) y la otra es una imagen LDR.

14. Un producto de programa informático que comprende código de software de todas las respectivas etapas de método que habilitan que un procesador ejecute el método de la reivindicación 12 o la reivindicación 13, cuando se ejecuta en un ordenador.
- 5 15. Una señal de imagen que codifica una imagen (Im_1) de colores de regiones de una escena de alto rango dinámico, que comprende palabras de código de N bits que codifican al menos luminancias de los colores de regiones; un valor de gris de diferenciador de regiones (gTS), indicando, en un sistema de código usado para codificar las palabras de código de N bits que codifican al menos las luminancias de los colores de regiones, una demarcación entre al menos una primera región geométrica de píxeles que representa un primer objeto de mayor luminancia en al menos un bloque de la escena de alto rango dinámico o mayores valores de las palabras de código de N bits que codifican los mismos, y al menos una segunda región geométrica de píxeles que representa un segundo objeto de menor luminancia en el al menos un bloque de la escena de alto rango dinámico o menores valores de las palabras de código de N bits que codifican los mismos, y un mapeo de luma (TOM) para cambiar luma de píxeles de al menos uno del primer y el segundo objeto para transformar la imagen (Im_1) a una imagen de diferente rango dinámico (IM_RC_HDR), en la que la imagen (Im_1) y la imagen de diferente rango dinámico son una imagen LDR y una imagen HDR con un brillo máximo fijo preestablecido cuando se crea la señal de codificación de imagen.
- 10
- 15
16. Una memoria extraíble que almacena una señal de acuerdo con la reivindicación 15.
- 20 17. Un sistema de distribución de video a través de cualquier tecnología de red, que emplea cualquier codificador de imagen de acuerdo con las reivindicaciones 1-7.

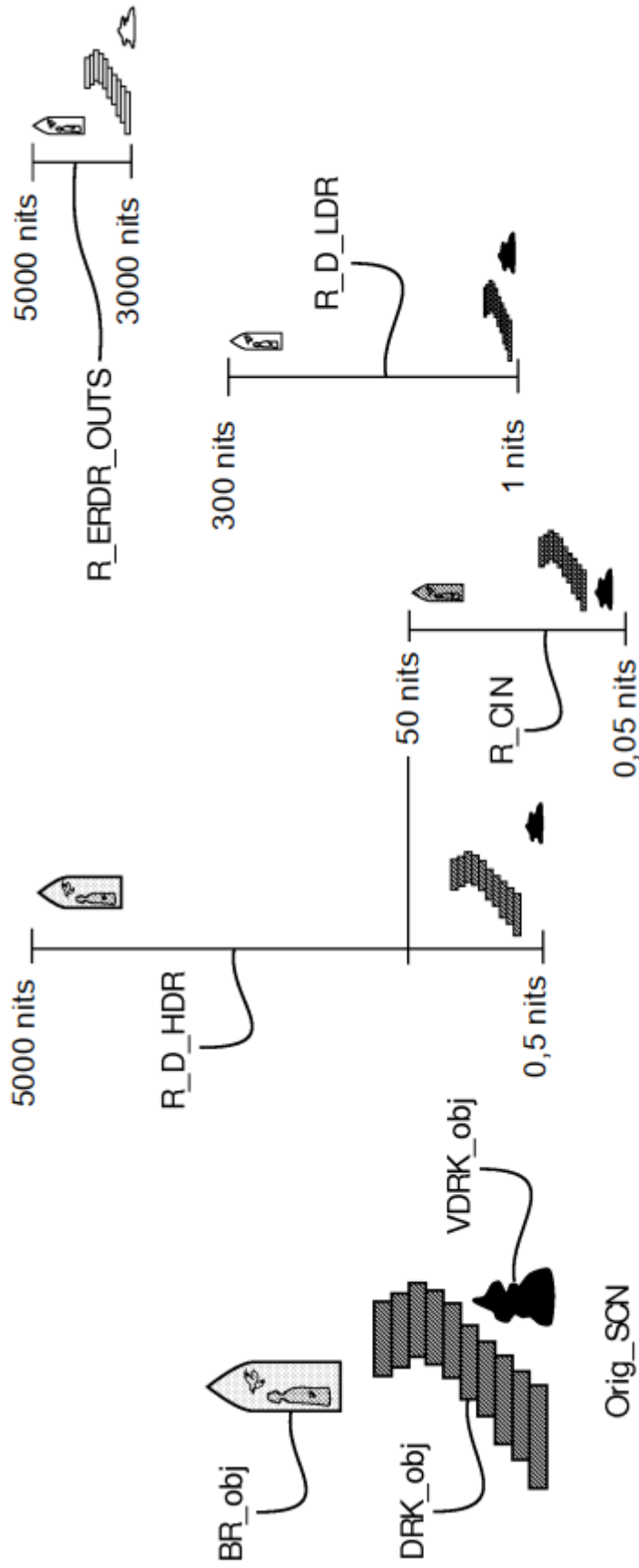


FIG. 1a

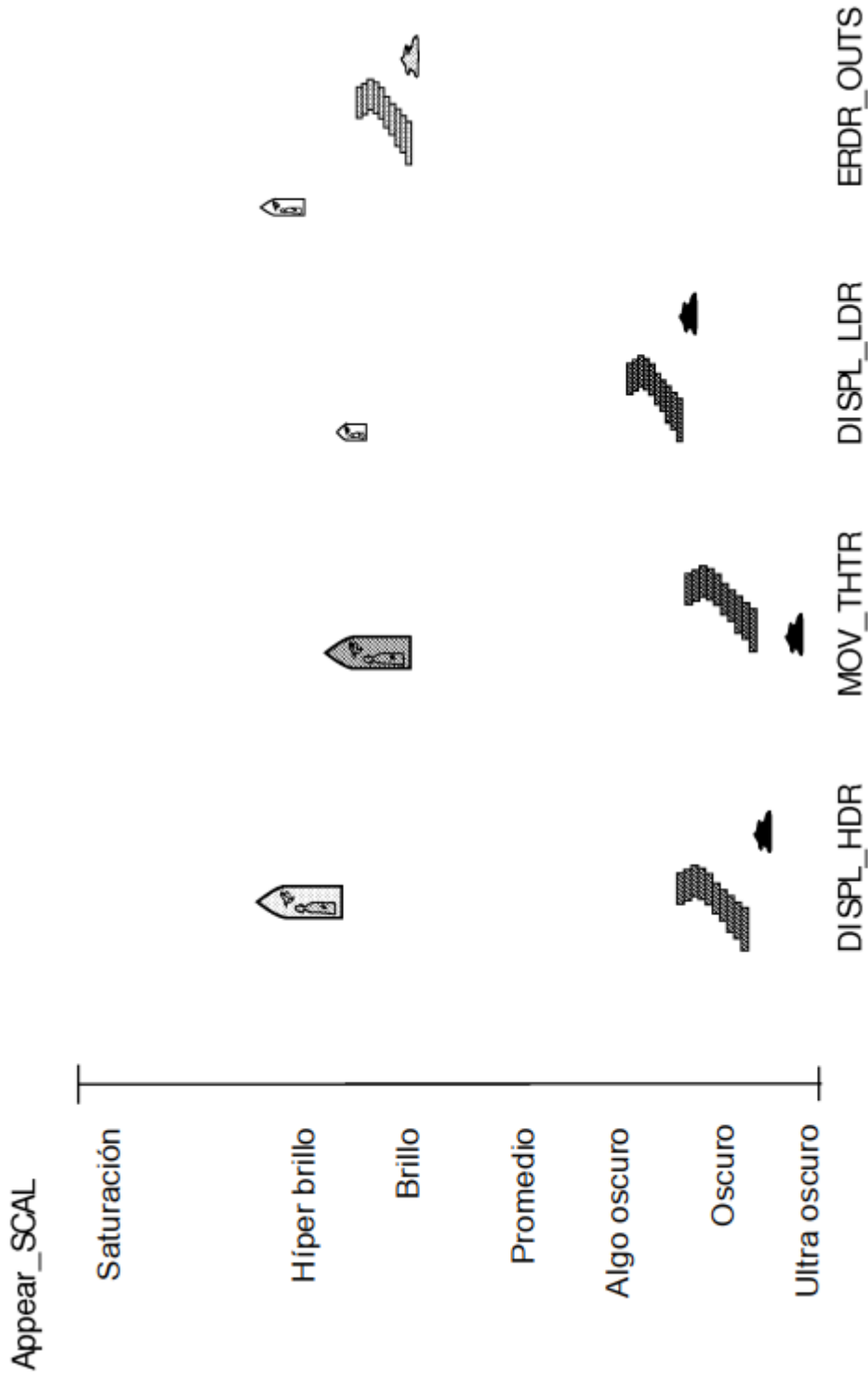


FIG. 1b

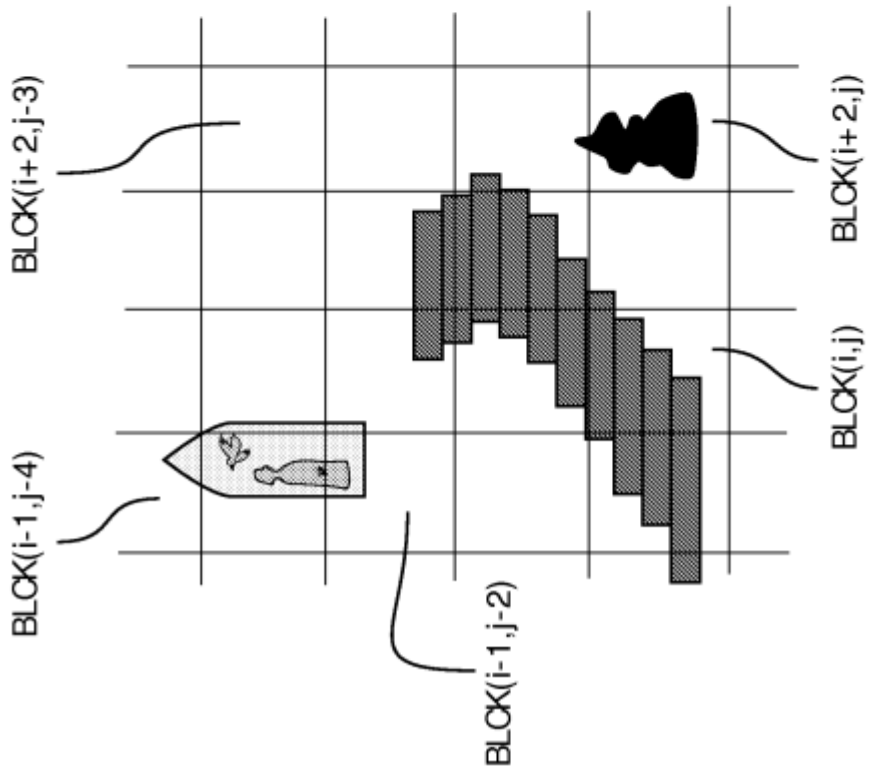


FIG. 2a

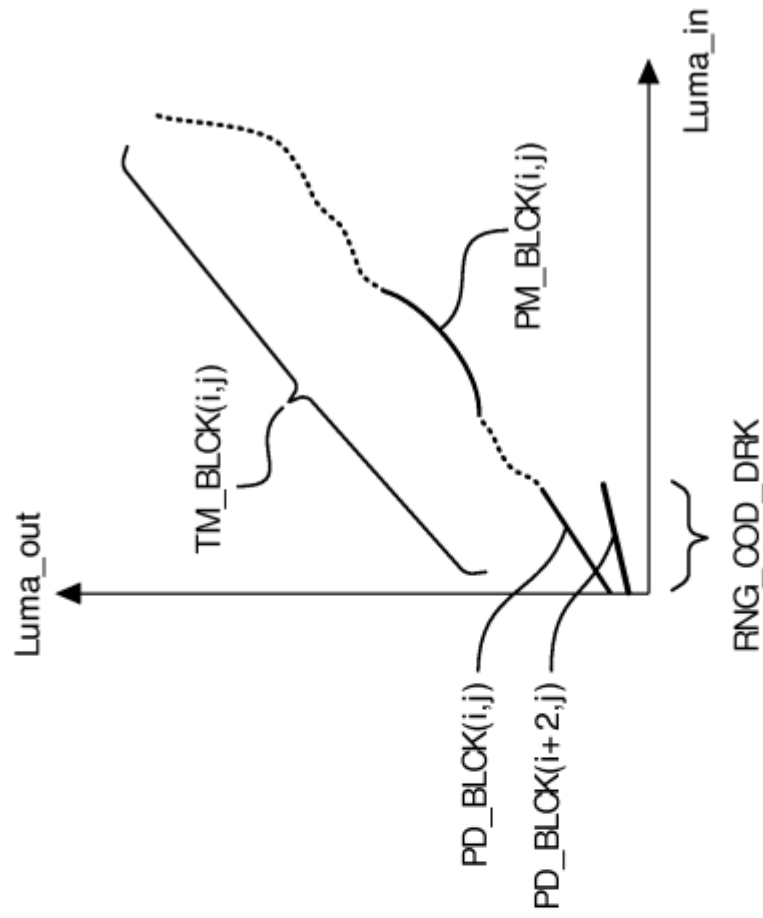


FIG. 2b

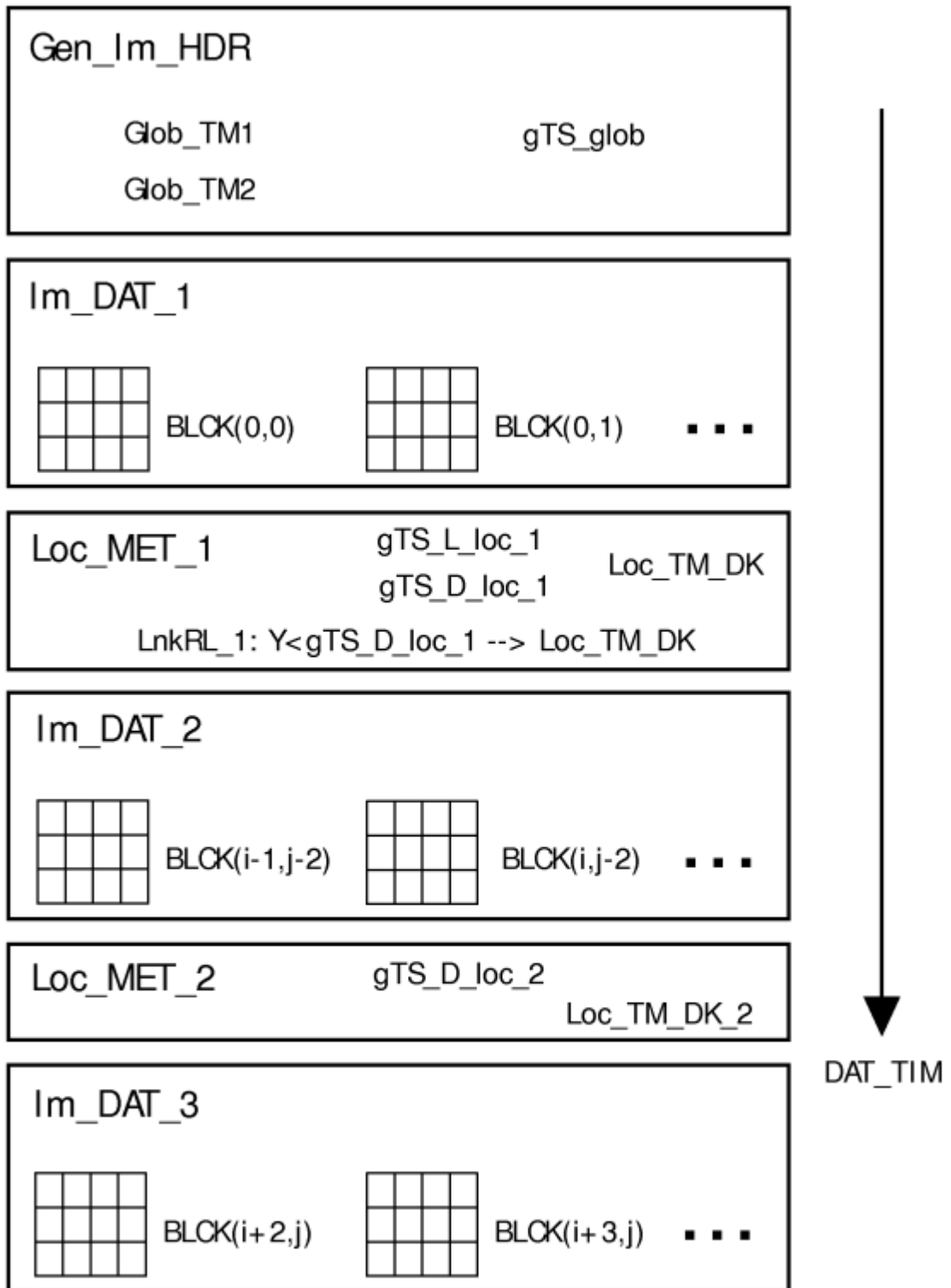


FIG. 3

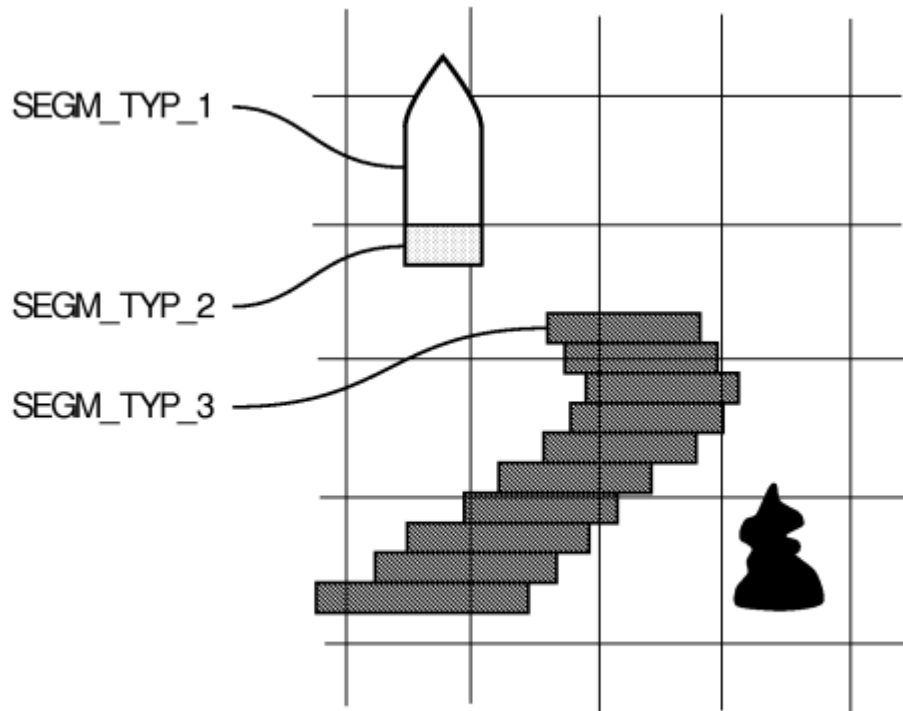


FIG. 4

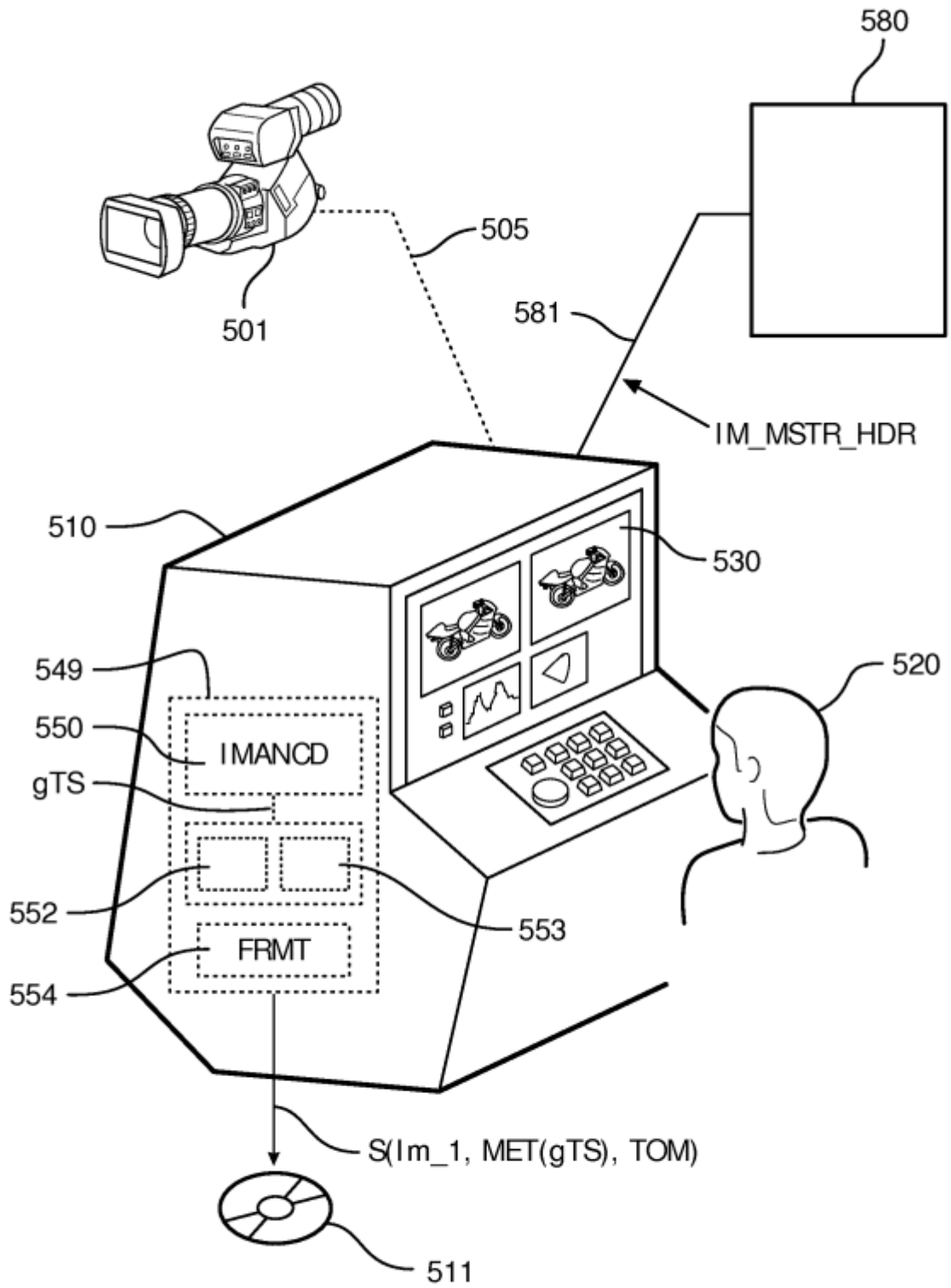


FIG. 5

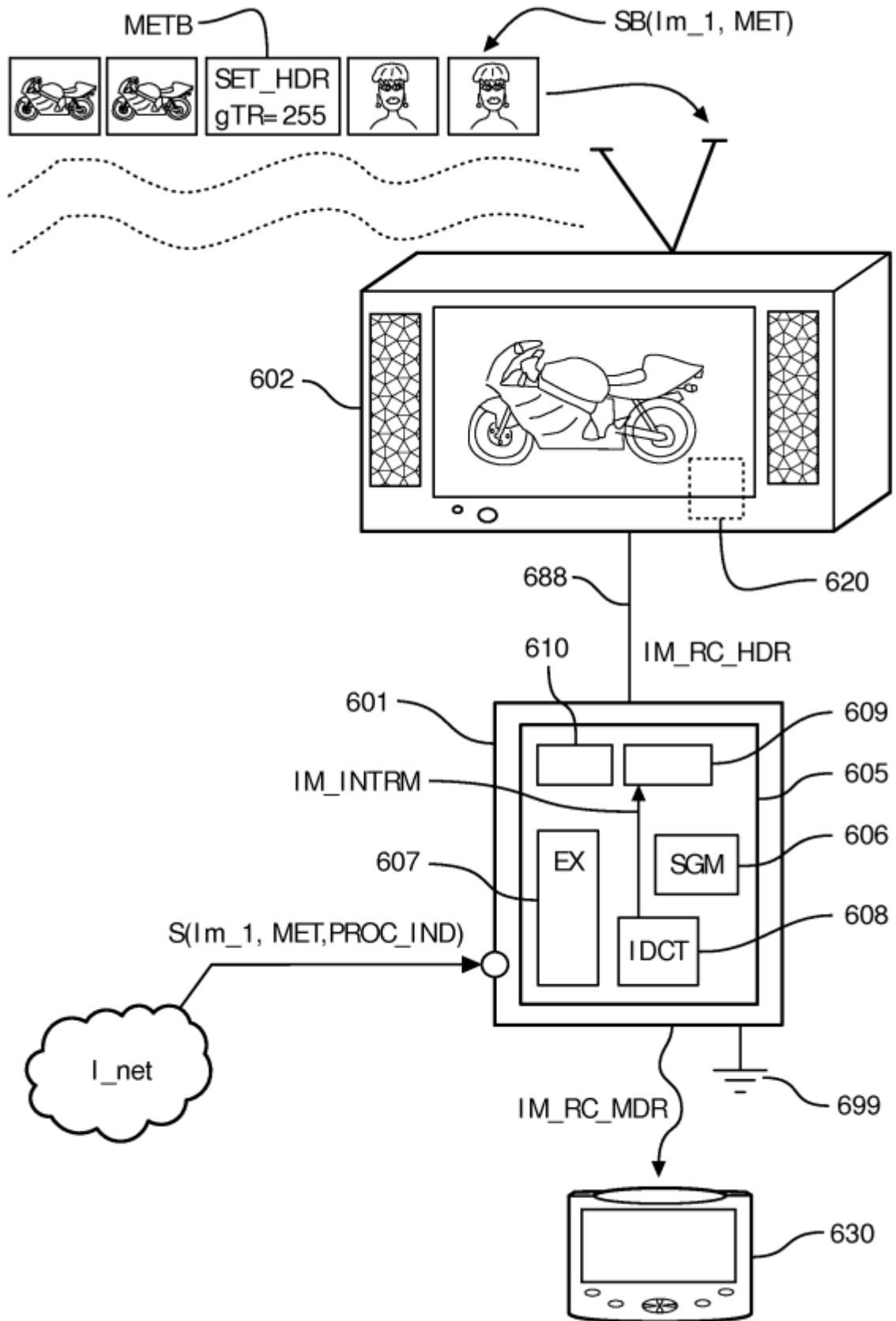


FIG. 6

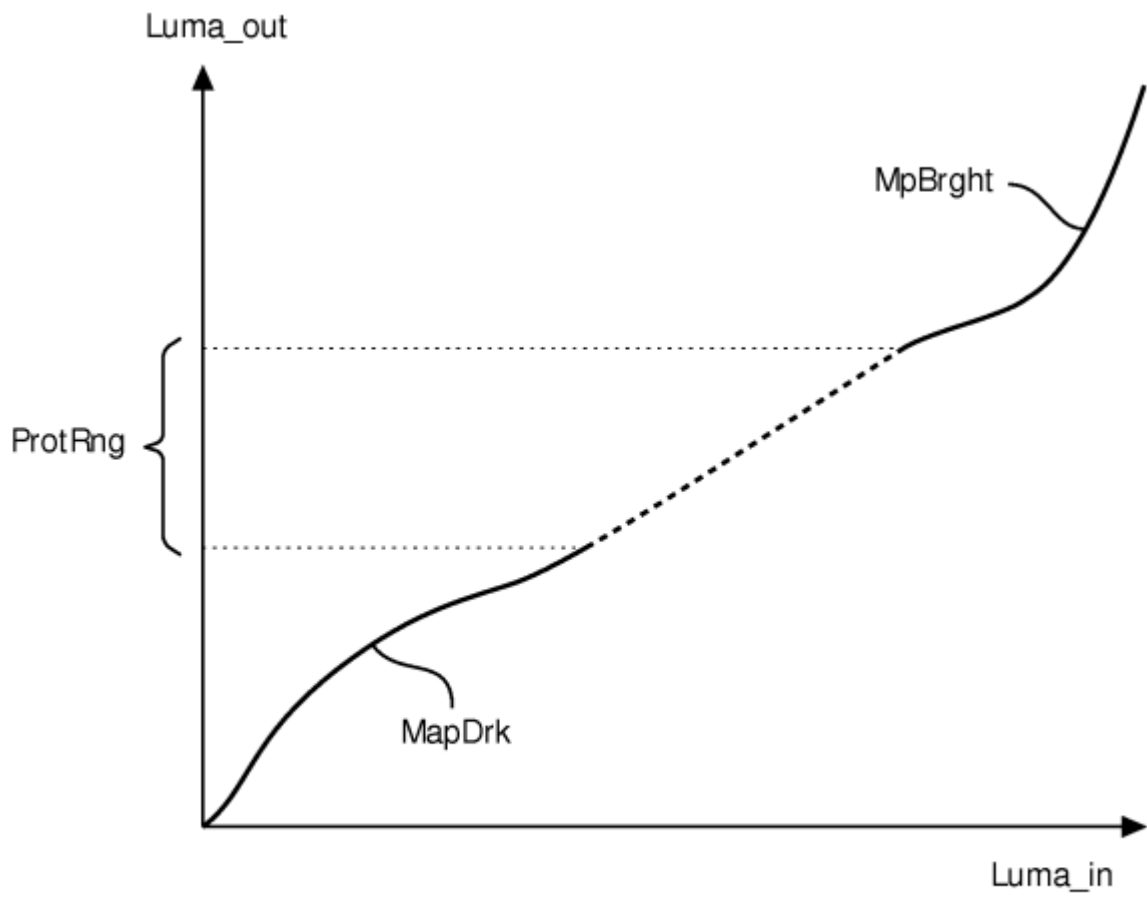


FIG. 7

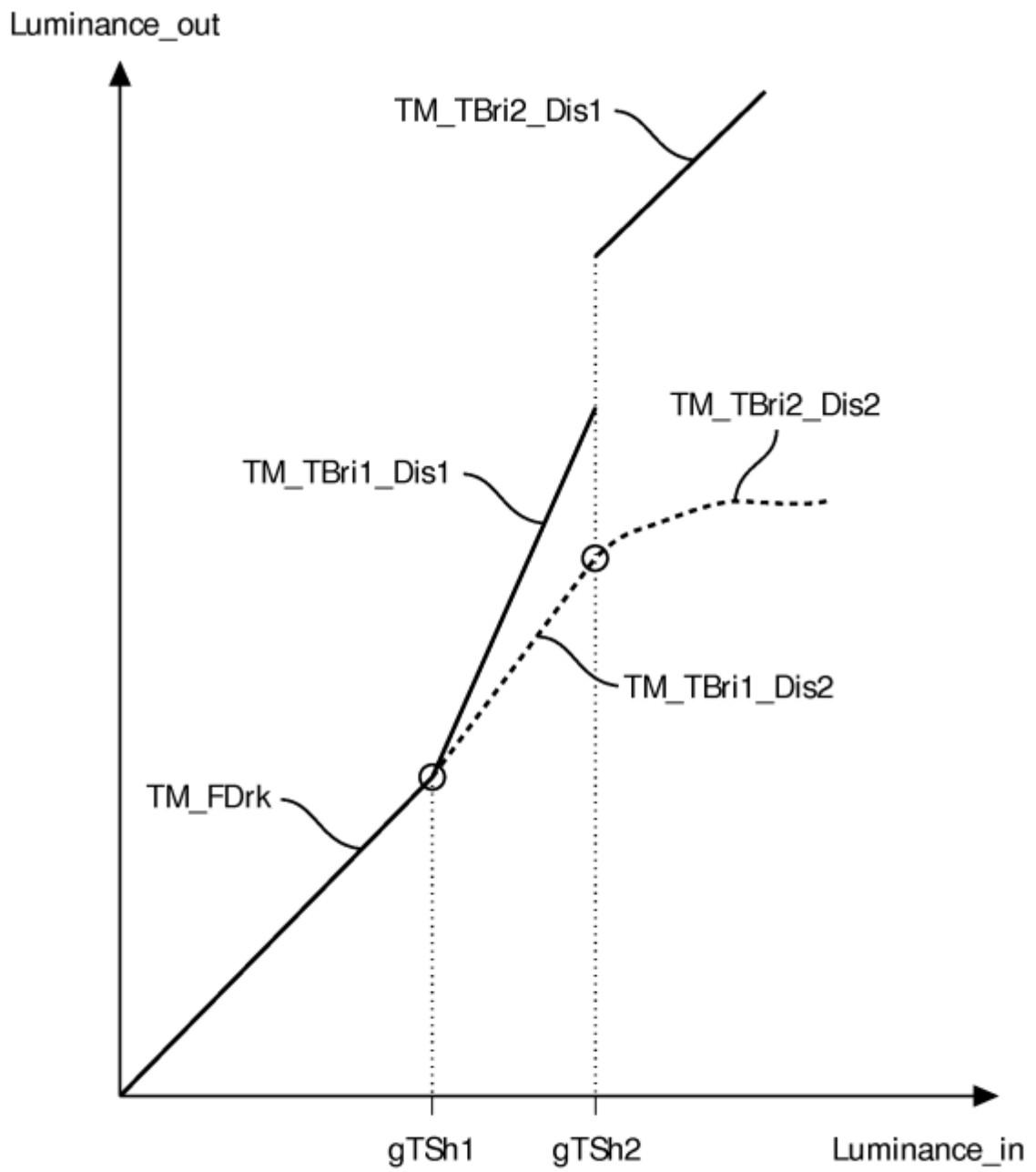


FIG. 8