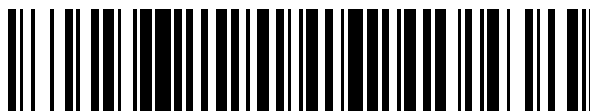


19



OFICINA ESPAÑOLA DE  
PATENTES Y MARCAS

ESPAÑA



11 Número de publicación: **2 728 153**

51 Int. Cl.:

**H04N 19/46** (2014.01)

**H04N 19/136** (2014.01)

**H04N 19/154** (2014.01)

**H04N 19/85** (2014.01)

**G09G 3/34** (2006.01)

12

## TRADUCCIÓN DE PATENTE EUROPEA

T3

86 Fecha de presentación y número de la solicitud internacional: **09.09.2011 PCT/IB2011/053950**

87 Fecha y número de publicación internacional: **22.03.2012 WO12035476**

96 Fecha de presentación y número de la solicitud europea: **09.09.2011 E 11776535 (4)**

97 Fecha y número de publicación de la concesión europea: **13.03.2019 EP 2617195**

54 Título: **Aparatos y métodos para mejorar la codificación de imágenes**

30 Prioridad:

**16.09.2010 EP 10177155**

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

**22.10.2019**

73 Titular/es:

**KONINKLIJKE PHILIPS N.V. (100.0%)**

**High Tech Campus 5**

**5656 AE Eindhoven, NL**

72 Inventor/es:

**DAMKAT, CRIS;**

**DE HAAN, GERARD;**

**MERTENS, MARK, JOZEF, WILLEM;**

**MUIJS, REMCO;**

**HAMMER, MARTIN y**

**NEWTON, PHILIP, STEVEN**

74 Agente/Representante:

**ISERN JARA, Jorge**

ES 2 728 153 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 para mejorar la codificación de imágenes

5 Campo de la invención

La invención se refiere a aparatos y métodos y productos resultantes como productos de almacenamiento de datos para una codificación mejorada de imágenes, en particular, permitiendo un mejor manejo de las imágenes codificadas por pantallas.

10 Antecedentes de la invención

Recientemente se han producido nuevos desarrollos con respecto a la codificación de imágenes/vídeo (ya sea de escenas capturadas o gráficos de ordenador), es decir, es deseable capturar mejor todo el rango de luminancias y colores que ocurren en la naturaleza, lo que se denomina HDR (alto rango dinámico) de codificación. Como ambas cámaras y pantallas están obteniendo rangos nativos cada vez más grandes, se requiere un mejor estándar para transmitir la información de la imagen entre ellas. Por otro lado, todavía existe una gran cantidad de dispositivos de rango inferior (por ejemplo, pantallas antiguas, impresoras, etc.), y estos también están presentes en algunas cadenas de sistemas de imágenes. Normalmente, un dispositivo de bajo rango dinámico (LDR), como una cámara de baja calidad, codifica en palabras de datos de 8 bits (píxeles) un rango medio de valores interesantes (por ejemplo, colores de cara bien iluminados), al coste de colores fuera de este rango [Cabe señalar que donde no se sacrifica la comprensión, podemos usar el término color incluso si en una tripleta de codificación de colores su luminancia es el factor más importante para la presente discusión].

25 Si un humano mira una imagen, hay un número de factores que influyen en la calidad. En primer lugar, está el brillo del blanco más blanco que puede reproducirse. En segundo lugar, existe el negro más oscuro que aún se puede reproducir, y quizás reproducir razonablemente, por ejemplo, con poco ruido u otra interferencia. El blanco y negro determinan el rango dinámico del dispositivo. Pero para una imagen real, esos no son los únicos parámetros que influyen en el aspecto. También hay parámetros que determinan dónde deberían estar idealmente los grises intermedios. El primero es el contraste, que es una medida relacionada con la claridad de los diferentes objetos en la imagen. Si hay al menos algunos objetos de los diferentes grises posibles entre un buen blanco y negro, se dice que la imagen tiene un buen contraste global. Pero también el contraste local puede ser importante, por ejemplo, entre un objeto y su entorno. Incluso cambios de luminancia muy locales, como la nitidez, influyen en el contraste percibido. Es por mirar a, por ejemplo, una escena real que los espectadores ven tiene un contraste realmente impresionante (por ejemplo, en contraste con una imagen proyectada adyacente de 6 bits). Pero en segundo lugar, también la ubicación de los objetos/regiones en el eje de blanco a negro tendrá un impacto, especialmente en la naturalidad (o aspecto artístico). Por ejemplo, se supone que las caras (bien iluminadas) tienen un cierto porcentaje de reflexión de la luz en comparación con el blanco. Una cara demasiado blanca puede parecer extrañamente brillante, o el espectador puede malinterpretar la imagen porque piensa que la luz está iluminada por alguna luz adicional. En tercer lugar, la precisión de los colores asignados puede ser importante, no tanto en texturas complejas, sino, por ejemplo, En gradientes faciales. Muchos espectadores parecen preferir las mejoras de calidad relacionadas con el brillo (incluida la saturación de color relacionada) sobre los otros aspectos, y esta aplicación se centrará principalmente en problemas relacionados con la luminancia.

45 El propósito de una pantalla es mostrar una renderización de calidad para un espectador. Idealmente, esta sería una renderización precisa (fotorrealista), pero como todavía está muy lejos en el futuro, se pueden usar otros criterios de calidad como, por ejemplo, Reconocimiento de la imagen, naturalidad aproximada (por ejemplo, ausencia de artefactos) o efecto/impacto visual, etc.

50 Una pantalla HDR popular que emerge actualmente es una pantalla LCD con retroiluminación LED en un patrón bidimensional, que permite la atenuación bidimensional. El rango dinámico de tales pantallas está influenciado por varios factores.

55 En primer lugar, los LCD se están volviendo cada vez más brillantes debido a la mejora de la retroiluminación. Donde hace un par de años, 200nit blanco era típico, ahora 500nit es típico, los próximos años 1000nit serán típicos, y más tarde incluso 2000nits o más. Sin embargo, esto plantea graves limitaciones técnicas en la televisión o el monitor, como el coste y el uso de potencia.

60 En segundo lugar, con respecto a los negros, los LCD tienen un problema con la fuga de luz (especialmente en ciertas condiciones como la visión de ángulo grande), lo que significa que un LCD puede tener un contraste intrínseco (celda de LCD abierta/cerrada) de 100:1, aunque la investigación está haciendo que los LCD sean mejores. Una solución a esto es cambiar la cantidad de luz que viene desde atrás a través de la válvula de LCD. De este modo, teóricamente, las pantallas de atenuación 2D pueden lograr un contraste muy alto, ya que si la luz detrás de la celda LCD tiene una luminancia cero, aparte de la fuga, una luminancia cero saldrá localmente de esa región de la pantalla. Se han informado rangos dinámicos por encima de 10000:1 o incluso 100000:1. Sin embargo, en la práctica, un factor importante que limita la renderización en negro de la pantalla es la luz del entorno que se refleja en el cristal frontal de

la pantalla. Esto puede reducir el rango dinámico a un 100:1 más realista o incluso a menos de 20:1 para ambientes brillantes. Sin embargo, también en un entorno oscuro, la luz puede escapar debido a todo tipo de razones, por ejemplo, interreflexiones en el vidrio frontal de una región más brillante a una región más oscura.

- 5 Por último, por supuesto, el ojo humano también es importante, y principalmente su estado de adaptación, pero también el complejo análisis de imágenes que tiene lugar en el cerebro. El ojo se adapta en una combinación de iluminación de la habitación por un lado y el brillo de la pantalla en el otro (en realidad, las imágenes mostradas). Estos dos factores pueden estar relativamente en sintonía, por ejemplo, televisores de 500 nit bajo visualización normal de la habitación de estar, pero también pueden estar muy alejados en otros escenarios de renderización. No solo se verá
- 10 afectado el detalle que se ve en negro, sino también la apariencia de las regiones brillantes. Por ejemplo, la comodidad de la visualización se verá influenciada por la configuración de la pantalla en particular, es decir, el cansancio de los ojos o incluso los efectos psicológicos, como el disgusto de la renderización de la imagen. La retina es muy compleja, pero puede resumirse simplemente de la siguiente manera. Sus conos tienen un proceso bioquímico que siempre trata de hacer que la sensibilidad del ojo (por medio de cantidades de moléculas sensibles a la luz) sea óptima para cualquier escena. Esto funciona porque cualquiera que sea la iluminación (que puede cambiar entre la luz de la luna
- 15 llena de 0.1 lx, para cielo nublado o habitaciones de 100 lx no muy bien iluminadas, para dirigir la luz del sol brillante de 100000lx, es decir, tiene un rango de más de un millón de factores de diferencia), las reflexiones de los objetos suelen variar de 1-100%, y es esa pantera oscura en el arbusto oscuro que la visión humana necesita discernir localmente. El ojo debe hacer frente a un rango dinámico de escena más grande, teniendo en cuenta los efectos de
- 20 iluminación como las sombras o la iluminación artificial, que normalmente puede ser de 10000:1. Las células retinianas adicionales como las células ganglionares hacen un uso más inteligente de la combinación de todas estas señales primarias, y así lo hacen, por ejemplo, cambiando el nivel de una respuesta local dependiendo de las luminancias de su entorno, etc.
- 25 Finalmente, un factor muy importante en la conversión mediante el análisis de este campo de imagen en bruto preprocesado es la corteza visual. Volverá a determinar, por ejemplo, el color de un parche amarillo una vez que se dé cuenta de que este parche no es un objeto separado, sino parte de otro objeto amarillo, o cambie el color del césped que se ve detrás de una ventana de vidrio una vez que comprende el reflejo de color que se superpone a esa región local. Genera lo que podríamos llamar "apariencia" del color final y, en teoría, este factor es el factor que tanto los
- 30 fabricantes de pantallas como los creadores de contenido están interesados. Por lo tanto, cualquier tecnología que se ajuste más a lo que necesita la visión humana es deseable (en particular cuando se toman en cuenta otras restricciones técnicas).

- Aunque todavía no existe un estándar generalmente reconocido para codificar imágenes HDR (especialmente para
- 35 vídeo), los primeros intentos de codificar imágenes (típicamente capturados al estirar los límites de los sistemas de cámara, por ejemplo, utilizando exposiciones múltiples y esperando que la lente no impida el esfuerzo demasiado) hicieron esto al asignar palabras de bit grandes (por ejemplo, 16 bits, permitiendo una codificación lineal de 65000:1, y más para la codificación no lineal) a cada píxel (por ejemplo, el formato exr). Luego, el mapeo de una cantidad variable de luz que se refleja (a la que el ojo se adapta parcialmente pero en gran parte) en los objetos de la escena
- 40 a un sistema de procesamiento de imágenes que comprende un módulo de válvula LCD y una retroiluminación se puede hacer, por ejemplo, técnicas de estimación de iluminación como en el documento EP1891621B [Hekstra, stacked display device]. Un algoritmo simplista para realizar la salida de luminancia = luminancia de retroiluminación x Transmisión LCD es tomar la raíz cuadrada de la entrada de 16 bits de HDR, asignando así una imagen de fondo multiplicativa de 8 bits que se puede submuestrear para los LED (conforme a las técnicas de codificación de relación).
- 45 También hay otros métodos para codificar claramente los valores de luminancia de la escena que aparecen simplemente como son en las formas clásicas, por ejemplo, el documento EP2009921 [Liu Shan, Mitsubishi Electric] que utiliza un enfoque de dos capas para codificar los valores de píxeles.

- Sin embargo, los inventores se han dado cuenta de que, si uno va por una nueva codificación, además de tal mera
- 50 codificación de los píxeles de la imagen de la escena (y utilizando esta codificación principal, única para toda la cadena), es deseable una codificación adicional, ya que mejorará en gran medida la comprensión y, por tanto, la utilidad de las acciones de imagen.

- El documento WO2008/095037 enseña una tecnología de ahorro de energía con retroiluminación para mostrar
- 55 imágenes LDR en pantallas LCD LDR con retroiluminación, utilizando una información de control de brillo sincronizado (SSBC) del sistema. Este sistema no se puede utilizar para crear apariencias de imagen HDR controladas, por ejemplo, derivándolas de las imágenes de entrada LDR, entre otras cosas porque involucra el procesamiento de imágenes solo en el rango de luminancia o luminancia LDR (y que está estrechamente vinculado como la inversa de la cantidad de atenuación de la retroiluminación), y también porque no enseña el tiempo control a través de metadatos TMA, que
- 60 pueden comunicarse con los datos de vídeo.

#### Sumario de la invención

- De manera ideal, una cadena de codificación de vídeo es una cadena simple, y solo hay errores menores en
- 65 comparación con la representación ideal, que por lo tanto se pueden descartar. Así fue como se realizó la codificación de la señal de televisión en el pasado (por ejemplo, NTSC, y los estándares basados en esos principios, como

MPEG2). Se define una pantalla estándar/de referencia (con fósforo EBU, un gamma de 2.2, ciertas condiciones de visualización), y esto al menos permite que se capture alguna codificación de los colores de la escena fijamente definida de la escena a capturar. Luego, la cámara se diseñará sobre la base de esta pantalla (resultado de señales codificadas relacionadas con la pantalla, por ejemplo, YCrCb). Las habilidades del operador de la cámara, la postproducción, etc., ajustan los datos para que estén más cerca del espacio de color de la pantalla final (generalmente, al ver el resultado final en un monitor de referencia). Sin embargo, esta era una situación que estaba bien al principio de la reproducción de la imagen, cuando solo había un tipo de pantalla, y ya estábamos contentos de tener cualquier sistema en funcionamiento. Sin embargo, en la actualidad, las pantallas de televisión abarcan tecnologías tan variadas como los teléfonos móviles en condiciones de luz solar o la proyección de cine en casa, y también los productores de televisión ofrecen cada vez más funciones de procesamiento de imágenes en sus pantallas. Por lo tanto, se puede plantear una pregunta interesante acerca de quién controlaría la mayor parte de las apariencias de color finales: el creador del contenido (es posible que Hollywood desee que al menos algunos tengan voz en los límites de cómo un televisor puede cambiar los colores y el brillo de sus películas), el fabricante de la pantalla (generalmente a través de la mejora automática de la imagen u otro procesamiento relacionado con la pantalla), o el usuario final (a través de los controles ofrecidos por la pantalla). Al formular un nuevo estándar de televisión, uno puede involucrar tales consideraciones para definir qué se puede prescribir (al menos opcionalmente) en dicho estándar.

Por ejemplo, mientras que en los próximos años no habrá una coincidencia perfecta entre lo que le gustaría mostrar al creador del contenido y lo que puede mostrar cualquier visualización real en particular (y el entorno de visualización) (por ejemplo, el creador del contenido puede querer acercarse) a una escena oscura, pero en el lado de la renderización puede iluminarse), se pueden ofrecer mejores opciones para controlar el comportamiento de las mismas (por ejemplo, permitir que la pantalla haga una mejora más inteligente de la imagen, o en configuraciones técnicas generales para su funcionamiento, como valores de manejo para los diversos componentes de la pantalla).

Esto puede ser útil para el espectador (por ejemplo, para proporcionar una cierta cantidad de renderización o efecto de imagen (color/luminancia), dado el hardware de visualización pero también las preferencias del usuario, por ejemplo, en función de su edad, carácter, estado de ánimo, etc.), pero al menos la información adicional en la codificación de vídeo (más allá de los simples colores de píxel) también se puede usar para manejar restricciones de visualización física como el consumo de potencia, problemas térmicos, envejecimiento, etc. De manera interesante, algunas codificaciones de datos adicionales son ventajosamente tan genéricas, que pueden proporcionar un mayor valor en toda la cadena. Un creador de contenido (o postcreador, que podría comprender un servicio adicional basado en humanos, o incluso un análisis de vídeo automático, por ejemplo para una transcodificación) puede, por ejemplo, usar los datos codificados adicionalmente para crear una mejor descripción de su película, y las intenciones reales que tenía con ella, lo que permite una mejor renderización en el lado de la pantalla. Un fabricante de pantallas puede controlar mejor el comportamiento en tiempo de ejecución de su pantalla (dada la entrada de imagen muy variable). Un usuario final/espectador puede, si así lo desea, ajustar mejor el vídeo a sus propias preferencias, y verlo como más le guste (por ejemplo, en caso de que encuentre alguna programación parpadeante demasiado molesta, él puede sintonizar esos fragmentos de programas abajo).

Varios de estos problemas y consideraciones de la necesidad de una mejor codificación de vídeo se tomaron como entrada cuando se contemplan las diversas realizaciones de acuerdo con la presente invención.

Para abordar al menos algunas de esas preocupaciones, proponemos un método de codificación, además de datos de vídeo (VID), datos adicionales (DD) que comprenden al menos un instante de cambio de tiempo (TMA\_1) que indica un cambio en el tiempo de un luminancia característica (CHRLUM) de los datos de vídeo, cuya luminancia característica resume un conjunto de luminancias de píxeles en una imagen de los datos de vídeo, el método comprende:

- generar sobre la base de los datos de vídeo (VID) datos descriptivos (DED) de la variación de luminancia característica del vídeo, los datos descriptivos que comprenden al menos un instante de cambio de tiempo (TMA\_1), y

- codificación y salida de los datos descriptivos (DED) como datos adicionales (DD).

Estos instantes de tiempo de cambio proporcionan información adicional muy importante en el vídeo, y se pueden usar para procesar y/o renderizar el vídeo en un dispositivo receptor como un televisor, y en particular, mejor sintonizado para cada televisor en particular, la preferencia actual del espectador (potencialmente dependiente de cada subsegmento de un vídeo en particular), etc. Convencionalmente, la filosofía de la codificación de vídeo siempre ha sido que este conjunto de imágenes se puede codificar satisfactoriamente mediante la codificación de imágenes separadas con técnicas de codificación de imágenes de píxeles. Sin embargo, al mirar una escala más gruesa, también hay información importante en la estructura temporal de la escala más gruesa del vídeo. En principio, se podría esperar que esta información pueda derivarse al tener disponibles esas imágenes. Sin embargo, puede haber factores en esa información que pueden no ser fácilmente derivables, por ejemplo, mediante un aparato automático de análisis de vídeo en un extremo receptor. Por ejemplo, es posible que el componente de análisis no tenga recursos suficientes, tales como: puede que no tenga algoritmos de análisis suficientemente complejos, o no tenga acceso a suficientes

imágenes de un futuro en comparación con un instante de tiempo particular, como un instante de tiempo para la visualización. Además, un creador de contenido puede querer transmitir algo especial sobre alguna evolución temporal de la señal de la imagen, en particular sus luminancias de píxeles. Por ejemplo, el creador puede haber creado una sucesión de imágenes codificadas que contienen una explosión, que puede tener valores de píxeles dependientes de las limitaciones físicas del sistema de codificación (por ejemplo, puede haber tenido que comprometerse para asignar la explosión a los mejores valores LDR de 8 bits disponibles). Además de eso, es posible que desee transmitir información adicional, por ejemplo, que se supone que esto es una "explosión muy poderosa", mientras que una segunda explosión más adelante, aunque sus valores de píxel pueden no ser tan diferentes debido a las limitaciones de codificación (por lo que es muy difícil para el dispositivo de análisis juzgar automáticamente la diferencia), se supone que es una "explosión menos potente". En el lado de la creación de contenido, todavía suele haber un artista humano presente, por lo que además de determinar la codificación óptima de las imágenes de píxeles, puede codificar datos adicionales (por ejemplo, cambiar los valores de los píxeles de las imágenes, pero describirlos con datos complementarios en los datos adicionales).

Los interesantes datos adicionales que modelan mejor la naturaleza temporal del vídeo de acuerdo con las presentes realizaciones, pueden derivarse en base a un concepto de luminancia característica (CHRLUM). Esto resume las luminancias presentes globalmente en al menos una imagen, y muchas veces en imágenes sucesivas (potencialmente también promediadas sobre algunas imágenes). Por ejemplo, un movimiento de la cámara desde el interior de una región de sombra a una vista esencialmente soleada se mostrará en la luminancia promedio (de todos los píxeles) de la imagen sombreada que es diferente de la luminancia promedio de la imagen de la vista soleada. En particular, la luminancia característica se ve seriamente afectada si la variación es tan grande que cambia una cantidad considerable del rango de vídeo LDR, o si la luminancia característica está así formulada, se caracteriza por los niveles o variaciones del rango HDR, es decir, por ejemplo, una explosión contiene varios píxeles con luminancias muy altas en comparación con un nivel de luminancia promedio o esperado o deseado (o viceversa con entornos oscuros). Se puede generalizar ese concepto de nivel grueso de luminancia característica considerando solo algunos cambios de luminancia local de nivel aproximado (aunque, por ejemplo, al observar solo una región que contiene una luz brillante en una imagen, la luminancia característica es más local que el promedio de una imagen completa, si la caracterización se realiza sobre las regiones/acciones de luminancia principal del disparo actual, es esencialmente una caracterización de nivel aproximado). Por ejemplo, si un conjunto sucesivo de imágenes contiene una llama de explosión localizada, se puede derivar la luminancia característica solo promediando los píxeles de la llama (sin necesitar los píxeles del edificio circundante, por ejemplo). Esto puede hacerse promediando sobre la primera imagen que tiene la llama, o tomando una integral característica de píxeles de la llama en varias imágenes seleccionadas que contienen la llama, sin embargo, uno puede asignar el momento del cambio en el primer instante en que aparece la llama. Los datos descriptivos (DED) resultantes del análisis de las imágenes pueden ser de acuerdo con las realizaciones de la presente invención, realizadas de manera diversa según lo entienda la persona experta (por ejemplo, uno puede ser un punto de partida, o la coinformación final codificar un modelo de brillo elipsoidal de la bola de llama de explosión) sin embargo, siempre contendrá al menos un instante de cambio de tiempo (TMA\_1) mientras que la unidad de análisis y/o el operador humano consideran que el cambio en ese cambio de luminancia característico se produce durante el vídeo (esto puede ser la primera imagen exacta con la explosión, o aproximadamente, en algún lugar, por ejemplo, al comienzo de la explosión). Los datos descriptivos finalmente se codifican además de una codificación clásica de datos de vídeo (que puede tener un contenido de información menor si parte del HDR está codificado en los datos adicionales) como datos adicionales DD, que puede ser solo una copia de los datos descriptivos DED, o comprenden un subconjunto y/o transformación de esos datos, pero es lo que se requiere en una estación adicional de una cadena de imágenes de acuerdo con los requisitos predeterminados.

La invención se define en las reivindicaciones independientes.

Ahora esto permite que el procesador o la pantalla de renderizado realicen varios procesamiento de imagen específicos en torno a los instantes de cambio de tiempo, en lugar de lo que normalmente haría a ciegas. Esto puede ser a la ligera ("la pantalla hace lo que tú quieras") o una estrategia más o menos precisa de lo que debería hacer una pantalla, pero preferiblemente todo puede ajustarse para tener en cuenta las características específicas de la pantalla y el entorno, pero también permite cierto control por parte de la creación, es decir, hacer que la pantalla siga al menos en cierta medida una sugerencia codificada (si el procesamiento debería, podría, no podría, etc., ocurrir). Una pantalla LCD con retroiluminación puede por ejemplo, considerar (ligeramente) modificar el manejo de la retroiluminación en comparación con lo que se consideraría una renderización exacta (es decir, una luminancia de píxel de salida producida con un porcentaje óptimo de transmisión LCD, y de la luminancia de retroiluminación, para obtener exactamente los valores de píxel deseados como se describe en, por ejemplo, una renderización de imagen HDR de 16 bits). Esto puede llevar a una imagen renderizada de manera algo diferente (diferentes colores de salida/luminancias), sin embargo, esto puede ser deseable. Además, las pantallas que tienen un único elemento de visualización por píxel, como por ejemplo, los OLED pueden usar la misma teoría algorítmica utilizando una "pseudoretroiluminación", es decir, permitiendo la modulación de su señal de manejo total, definiendo algún componente básico y una variación típicamente multiplicativa de la misma.

El reprocesamiento normalmente implicará una transformación funcional, por ejemplo, mapear los colores/luminancias de píxeles anteriores para al menos algunas regiones de un conjunto de imágenes sucesivas a nuevos colores/luminancias de píxeles. El cambio en la luminancia característica se puede reformular en diversas

realizaciones de reprocesamiento (por ejemplo, para la reducción del el parpadeo) como un cambio en las estrategias o parámetros de transformación, en particular, que comprende un momento de cambio deseado de transformación (tener en cuenta que, en principio, el tiempo de cambio en el instante TMA\_1 de cuando se consideró que se produjo el cambio en la luminancia característica puede ser diferente de un instante de tiempo TP\_1 en el momento en que comienza un reprocesamiento deseado [por ejemplo, atenuando un segmento de retroiluminación], pero a menudo pueden considerarse iguales, por ejemplo, si es necesario, definiendo la función de reprocesamiento o algoritmo para no tener impacto en las primeras imágenes, por ejemplo, para una función multiplicativa que le da los principales 1. Las indicaciones de las estrategias de procesamiento pueden ser diversas, desde niveles muy altos hasta muy estrictos. Por ejemplo, se puede indicar que se permite cualquier procesamiento, por ejemplo, para la toma actual, o no (si se debe renderizar de forma crítica debido a su calificación crítica). O se puede indicar si se permite un tipo de procesamiento (por ejemplo, una mera reducción de la luminosidad), o si solo se permite el procesamiento del tipo para tratar de renderizar de manera óptima el aspecto (por ejemplo, una escena oscura) dadas las consideraciones del lado de la pantalla frente a si también por ejemplo, se permite el procesamiento específico de la pantalla, como el ahorro de potencia, lo que puede reducir la calidad de la imagen. O incluso se puede prescribir una función específica para aplicar alrededor del instante de cambio de tiempo. Cabe señalar que el reprocesamiento no necesita ser uno fijo, sino que puede ajustarse, por ejemplo, depende de los ajustes preestablecidos deseados del espectador, pero aún puede construirse sobre la base de al menos un instante de cambio de tiempo (por ejemplo, con funciones de reprocesamiento paramétrico).

También es útil un método de codificación de datos adicionales (DD), que comprende un paso de codificación de un código de reprocesamiento particular (MULT) a partir de un conjunto de códigos acordados prefijados.

Además, es útil un método para codificar datos adicionales (DD), que comprende un paso de codificación en los datos adicionales (DD) de una estrategia de desviación, tal como, por ejemplo, un perfil temporal codificado (PROF) o un algoritmo matemático para calcular una estrategia de desviación, para reprocesar durante un intervalo de tiempo DTI las luminancias de píxeles de los datos de vídeo (VID), en comparación con las luminancias iniciales (Lin\*), en las que se puede basar el reprocesamiento en un modelo psicovisual, o en las características físicas de la pantalla y/o el entorno de visualización, etc.

Es decir En este caso la indicación se ha convertido más en una prescripción específica. Uno puede por ejemplo comenzar desde las luminancias iniciales Lin\* tal como se codificaron en la señal de vídeo VID, y aplicarlas a un perfil multiplicativo, lo que disminuye de manera suave/imperceptible las luminancias a lo largo del tiempo para esta toma. El perfil puede ser aditivo, multiplicativo, solo una indicación, por ejemplo, un promedio de nivel aproximado de lo que debería ser el perfil de luminancia final (salida) a lo largo del tiempo (y la televisión puede procesar, sin embargo, para obtenerlo aproximadamente), etc.

También es útil un método para codificar datos adicionales (DD) en donde el reprocesamiento es de un tipo que comprende determinar una imagen de iluminación para una retroiluminación (ILIM), y el paso de codificación comprende datos de codificación para influir en la determinación de la imagen de iluminación para una retroiluminación (ILIM) durante un intervalo alrededor del instante de cambio de tiempo (TMA\_1), tal como una función temporal compuesta por contribuciones de funciones de base elemental para al menos una región espacial de posiciones de una matriz bidimensional (MAP). Entonces se puede sugerir o controlar renderizaciones específicas actuando más directamente en la parte de la retroiluminación, de una manera espacio-temporal. Por ejemplo, uno puede caracterizar (parte de) un efecto HDR, tal como una explosión, componiéndolo a partir de un conjunto de funciones, tales como algunas oscilaciones locales, funciones de potencia decreciente, descomposiciones gaussianas, etc., que se definen al menos en la parte basada en el instante de tiempo (por ejemplo, una ventana de muestreo sobre la función, la ubicación del modo de un gaussiano se determina en comparación con TMA\_1, o el punto de inicio de una función decreciente, etc.)

También es útil un método de codificación de datos adicionales (DD) que comprende un paso en la codificación de la información de datos adicionales (DD) de luminancias características de un futuro del instante de cambio de tiempo (TMA\_1) y/o información de luminancias esperadas de una imagen de iluminación para una retroiluminación (ILIM) de una pantalla de referencia.

Teniendo lo más preciso posible un conocimiento del futuro del vídeo, especialmente un resumen de las próximas luminancias de píxeles de imagen puede hacer que la pantalla o el procesador de renderizado, o cualquier dispositivo que utilice los datos adicionales codificados, tome decisiones inteligentes con respecto a su procesamiento actual, por ejemplo, maximizando el impacto visual, potenciar la sensibilidad de la retroiluminación en vista del uso futuro de la potencia, etc. Para algunas aplicaciones tales como la administración de potencia, esta caracterización de las luminancias características futuras puede ser de un nivel muy aproximado ya que solo se necesita saber aproximadamente cuánta luz será necesaria (es decir, un promedio de la luminancia característica en los próximos 10 segundos y adicionalmente, o alternativamente, se pueden codificar dos minutos; una jerarquía temporal de tales caracterizaciones permite al lado receptor realizar predicciones más inteligentes, por ejemplo, sobre la potencia para gastar actualmente), sin embargo, para una realización precisa del impacto psicovisual, puede ser necesario un conocimiento más detallado de las modulaciones temporales. Ya sea para una pantalla con retroiluminación o sin retroiluminación, uno puede codificar de manera equivalente las variaciones características en una codificación de

imagen total (tal como VID), o en un componente (virtual) de la misma, tal como una contribución de retroiluminación, y el lado receptor puede obtener cualquier variante requerida, por ejemplo mediante el uso de un algoritmo de división multicomponente prefijado o codificado.

5 También es útil un método de codificación de datos adicionales (DD) que comprende un paso de codificación en los datos adicionales (DD) y una indicación de importancia (IMPLEV) para al menos un instante de cambio de tiempo (TMA\_1). Esto permite un reprocesamiento de desviaciones muy versátil, tal como por ejemplo un tratamiento jerárquico (por ejemplo, disminuyendo) de la renderización en varios intervalos de tiempo relacionados (por ejemplo, varios efectos de alto brillo relacionados). En el caso de que el lado de la pantalla tenga dificultades para reproducir  
10 todos los efectos, puede basarse en la importancia de hacer solo los más importantes, o puede diseñar un reprocesamiento teniendo en cuenta la jerarquía de importancia, etc.

También es útil un aparato (524) de codificación de vídeo dispuesto para codificar, además de datos de vídeo (VID), datos adicionales (DD) que comprenden al menos un instante de cambio de tiempo (TMA\_1) que indica un cambio en el tiempo de una luminancia característica (CHRLUM) de los datos de vídeo, cuya luminancia característica resume  
15 un conjunto de luminancias de píxeles en una imagen de los datos de vídeo, sobre la base de datos descriptivos (DED) con respecto a la variación de luminancia característica del vídeo.

Otros aparatos (524) de codificación de vídeo son útiles para codificar, además de datos de vídeo (VID), datos adicionales (DD) de acuerdo con cualquiera de los principios descritos anteriormente o descritos a continuación, en particular con codificadores incorporados específicamente, formateadores, etc. para diferentes especificaciones de lo que el lado receptor puede realizar como reprocesamiento de imágenes en momentos específicos.

También es útil un método de decodificación de datos adicionales (DD) de datos de vídeo (VID), los datos adicionales que comprenden al menos un instante de cambio de tiempo (TMA\_1) que indica un cambio en el tiempo de una luminancia característica (CHRLUM) de los datos de vídeo, cuya luminancia característica resume un conjunto de luminancias de píxeles en una imagen de los datos de vídeo, el método comprende además emitir al menos un instante de cambio de tiempo (TMA\_1).

30 Típicamente, el método de decodificación analizará la señal entrante y encontrará paquetes específicos, campos de datos, etc., reconocerá los datos codificados, posiblemente hará una extracción, transformación, refundición en un formato útil para el aparato, etc. puede generar los instantes de tiempo en los que puede o debe ocurrir alguna acción específica. Un aparato conectado a un decodificador que usa dichos datos adicionales puede prescribir otras formas de entrega (o incluso extraer solo datos particulares) dependiendo de su uso de los datos. Por ejemplo, si el aparato  
35 solo necesita conocer los instantes de tiempo de cambio de la luminancia característica, puede ser suficiente tener solo esos, pero un aparato de procesamiento de imágenes puede solicitar a la unidad de decodificación que realice un método de decodificación que también convierte los índices codificados en transformaciones previamente acordadas en un formato más manejable, por ejemplo, funciones sobre un segmento de tiempo finito para atenuación multiplicativa. Es decir, todos los datos adicionales serán emitidos en formatos acordados, ya sea fijos predefinidos, o negociados sobre la marcha con el aparato receptor, ya sean instantes de tiempo, indicaciones de reprocesamiento, datos adicionales que especifiquen la naturaleza temporal de la señal tales como medidas dependientes de la imagen, orientaciones de visualización, u orientaciones al aspecto de la película, etc.

Además, es útil un método para decodificar datos adicionales (DD) de datos de vídeo (VID), comprendiendo además el método de decodificar y generar al menos una de las entidades de datos codificados como se describe en este texto.

También es útil una señal de datos (NDAT) asociada con datos de vídeo (VID), que comprende al menos un instante de cambio de tiempo (TMA\_1) que indica un cambio en el tiempo de una luminancia característica (CHRLUM) de los datos de vídeo, cuya luminancia característica resume el conjunto de luminancias de píxeles en una imagen de los datos de vídeo.

También es útil un aparato (600) de decodificación de vídeo dispuesto para decodificar, relacionado con datos de vídeo (VID), datos adicionales (DD) que comprenden al menos un instante de cambio de tiempo (TMA\_1) que indica un cambio en el tiempo de una luminancia característica (CHRLUM) de los datos de vídeo, cuya luminancia característica resume un conjunto de luminancias de píxeles en una imagen de los datos de vídeo, y emite a través de una salida (650) al menos un instante de cambio de tiempo (TMA\_1).

También es útil un aparato (600) de decodificación de vídeo dispuesto para decodificar al menos una de las entidades de datos codificados especificados en cualquier parte de este texto, y además dispuesto para comunicarse con un segundo aparato (100) capaz de renderizar el vídeo (VID) esas al menos una de las entidades de datos codificados, con el fin de influir en la renderización por esas al menos una de las entidades de datos codificados.

Típicamente, diversas realizaciones del aparato de decodificación tendrán diversas subunidades, tales como un IC dedicado (parte de alguno) o un firmware o software dedicados que se ejecutan al menos temporalmente en un IC, por ejemplo, observan una parte específica de los datos adicionales, que comprende, por ejemplo, un código de

reprocesamiento, que aísla el código de reprocesamiento y lo envía sin procesar a una salida IC, o lo envía a una subunidad de conversión para convertirlo en un valor digital (o analógico) o conjunto de datos que es más útil para un dispositivo conectado. La persona experta entenderá que los mismos datos también pueden enviarse de diversas maneras un número de veces a través de diferentes salidas.

5 También es útil una disposición (110+100) que comprende un aparato (600) de decodificación de vídeo y una pantalla (100) en la que la pantalla está dispuesta para cambiar su renderización en base a al menos un instante de un cambio (TMA\_1).

10 Breve descripción de los dibujos

Estos y otros aspectos del método y el aparato de acuerdo con la invención serán evidentes y se explicarán con referencia a las implementaciones y realizaciones descritas a continuación, y con referencia a los dibujos adjuntos, que sirven simplemente como ilustraciones específicas no limitativas que ilustran el concepto más general, y en donde se usan guiones para indicar que un componente es opcional, los componentes sin guiones no son necesariamente esenciales. Los guiones también se pueden usar para indicar que los elementos, que se explican como esenciales, están ocultos en el interior de un objeto, o para cosas intangibles tales como, por ejemplo, selecciones de objetos/regiones (y cómo se pueden mostrar en una pantalla).

20 En los dibujos:

La figura 1 ilustra esquemáticamente una disposición de recepción de vídeo de ejemplo capaz de usar los datos adicionales DD de acuerdo con al menos algunas de las realizaciones descritas en el presente texto;

25 La figura 2 ilustra esquemáticamente una renderización de cómo cambian las luminancias de las imágenes en un vídeo para explicar algunos ejemplos de cómo funcionan al menos algunas realizaciones de la presente invención;

La figura 3 ilustra esquemáticamente los procesos que se aplicarían a ese vídeo, para obtener una renderización más satisfactoria en una pantalla;

30 La figura 4 ilustra esquemáticamente un poco más de procesamiento, para un ejemplo específico de renderización de una escena oscura;

La figura 5 ilustra esquemáticamente un entorno de creación, para crear datos adicionales DD;

35 La figura 6 ilustra esquemáticamente un aparato de decodificación para decodificar los datos adicionales DD;

La figura 7 ilustra esquemáticamente las matemáticas detrás del uso de los datos adicionales DD;

40 La figura 8 ilustra esquemáticamente una aplicación de la presente codificación de datos adicionales en un escenario/disposición de optimización de potencia; y

La figura 9 ilustra esquemáticamente un ejemplo de una codificación de los datos adicionales en relación con los datos de vídeo.

45 Descripción detallada de los dibujos

La figura 1 describe una posible disposición de visualización de vídeo en el hogar, que comprende un televisor 100 LED (o una pantalla general compatible con HDR, o incluso una pantalla LDR, especialmente si es más fácil de ajustar en su renderización que la única renderización del vídeo VID) , con un panel 101 LCD iluminado por un número de LED 102 (blancos o de colores), que pueden mostrar imágenes HDR o imágenes LDR (estándar, bajo rango dinámico) [en cuyo caso puede haber una cierta cantidad de procesamiento de vídeo, al menos para asignar valores de accionamiento de LCD y LED] de acuerdo con los principios descritos anteriormente y en la técnica anterior. Cabe señalar que la persona experta entenderá que los principios de la presente invención también se pueden asignar a otras pantallas, por ejemplo, un proyector con una iluminación segmentada y un DMD, pantallas OLED, etc.

En una realización de ejemplo (que usaremos como ejemplo para describir nuestros principios técnicos), el televisor/pantalla recibe sus señales de televisión o de imagen a través de una conexión (por ejemplo, con cable/HDMI o inalámbrica) 113, forman un reproductor basado en memoria, por ejemplo, un reproductor 110 BD (pero, por supuesto, alternativamente, las señales podrían provenir, por ejemplo, de un servidor a través de Internet, etc.). Este reproductor 110 BD obtiene el vídeo codificado de un disco 115 bluray, sobre el cual se codifica una pista 116 adicional con los datos adicionales DD de acuerdo con cualquiera de las realizaciones de la invención descritas a continuación (por supuesto, dichos datos también pueden codificarse de acuerdo con muchos principios diferentes, por ejemplo, dentro de la codificación de vídeo, por ejemplo, en campos antes de grupos de bloques, sin embargo, un conjunto separado de elementos de datos permite la codificación a través de otro canal, por ejemplo, Internet, para ser cosuministrada).



La figura 2 muestra un perfil temporal a lo largo del eje  $t$  del tiempo de una película (que puede ser cualquier sucesión temporal de imágenes relacionadas, por ejemplo, una película de Hollywood o un vídeo de una cámara de seguridad) con una luminancia característica  $YC$  para cada imagen (en el tiempo  $t$ ). Esta luminancia característica se deriva de todas las luminancias de píxeles presentes en esa imagen, por ejemplo, puede ser un promedio ponderado (ya que las cámaras normalmente también usan promedios pesados para determinar sus ajustes que manejan a la distribución de la luminancia de píxeles codificada en la imagen, por lo que esto se reflejará parcialmente en su grabación de la escena capturada), pero el análisis de histograma más inteligente puede estar involucrado. Por ejemplo, las luminancias medidas en los percentiles más altos pueden contribuir al menos parcialmente a  $YC$ , de modo que uno puede juzgar, por ejemplo, el alto rendimiento clave de ambientes al aire libre, saturación de sensores, grandes regiones locales muy luminosas (también se pueden tener en cuenta las propiedades espaciales en el algoritmo para derivar  $YC$ , tales como las relaciones de áreas claras y oscuras, o incluso histogramas de regiones oscuras dentro de áreas claras) por ejemplo, para analizar una captura a contra luz de una persona (parcialmente) delante de una ventana), etc. En cuanto a la determinación temporal de la luminancia característica, se puede determinar por imagen o calcular cualquier fórmula de acumulación matemática sobre cualquier número de imágenes sucesivas (por ejemplo, como en la figura 2, que proporciona la misma luminancia característica a todas las imágenes de una toma; entre límites de cambio de luminancia de disparo o característica). Cabe señalar que un ser humano que caracteriza/anota los cambios en la luminancia característica puede usar diversas indicaciones, y también puede delimitar los límites entre regiones temporales que se supone que tienen luminancias características diferentes (al menos en la imagen de salida procesada (por ejemplo, HDR o pseudo HDR) para la pantalla), pero las diferencias pueden ser difíciles de calcular con un algoritmo automático (por ejemplo, especificar un conjunto de intentos de renderización alternativos, por ejemplo, para diferentes pantallas).

Esta luminancia característica se puede usar para determinar dónde se produce una diferencia en las características de captura de escena, lo que, por ejemplo, tiene que traducirse en una renderización diferente en la pantalla, en particular, un manejo diferente de los LED de retroiluminación. Tomando el ejemplo de contraluz, la persona experta puede entender que el procesamiento de la pantalla (ya sea un simple cambio de la imagen de entrada basado en software, o el procesamiento relacionado con el hardware, tal como el manejo óptimo de la retroiluminación) puede ser tal que mejore esta imagen (teniendo en cuenta implícitamente (caso promedio) o explícitamente todos los aspectos relacionados con la visualización y el visor, la visualización puede, por ejemplo, renderizar las diferentes luminancias del objeto persona en la imagen de una manera más visible), o, especialmente con las nuevas pantallas HDR, la renderización puede empeorar la calidad visual.

Una primera parte de la película, una escena de conversación  $SCN\_CONVRS$ , describe tal una situación. Consiste en alternar los primeros disparos  $SHT\_WND$  de una primera persona sentada en una parte más luminosa de una habitación, y los segundos disparos  $SHT\_WLL$  de una segunda persona sentada en una parte más oscura de la habitación (o para los fines de esta explicación, un escenario similar y relacionado el procesamiento técnico puede ocurrir al intercalar una secuencia de tomas interiores y exteriores). Tener en cuenta las condiciones de iluminación de la escena artística y la configuración de la exposición de la cámara (humana o automática) puede mitigar parcialmente la diferencia (haciendo que ambas tomas estén bien/promedio expuestas), pero también puede mantener una cierta diferencia (por ejemplo, el director de fotografía quiere un aspecto particular contrastando los dos). Sin embargo, cuando se mapean todas las luminancias de píxeles subyacentes a una pantalla HDR (por ejemplo, el estiramiento involucrado en el mero mapeo de la señal LDR  $\{0, 255\}$  a una señal HDR  $\{0, 1024\}$  y a un rango HDR  $\{0, 2000nit\}$  de las luminancias de salida de la pantalla (región/píxel) en lugar de un  $\{0, 500\}$  rango de visualización LDR), no solo el aspecto particular puede verse comprometido, sino que incluso las regiones con ventanas brillantes pueden dañar los ojos del espectador, o al menos disgustar algunas espectadores. La situación se ilustra esquemáticamente en la figura 2 por medio de una matemática derivada que caracteriza el rango dinámico bajo  $PS\_R\_LDR$  dentro del espacio de manejo para una pantalla HDR (particular)  $R\_HDR$ . Esto podría ser un rango LDR  $\{0, 255\}$ , pero generalmente se corresponderá con un rango en donde uno representaría luminancias de escenas normales (tales como reflejos de objetos bien expuestos interiores o exteriores) y efectos HDR aún no optimizados, tales como explosiones, luces, etc. Uno puede querer que este rango no se incremente demasiado en HDR, pero mantenerlo mitigado, como LDR. Antes de renderizar en la pantalla HDR, la imagen descrita en el rango dinámico bajo que caracteriza a  $PS\_R\_LDR$ , tendrá un procesamiento para asignar típicamente los efectos a un rango de efectos HDR  $R\_UPEFF$  (véase figura 3) y un rango inferior  $R\_L\_HDR$  para objetos normales. Cabe señalar que este es solo un posible ejemplo esquemático para ilustrar las presentes invenciones y sus realizaciones. La imagen de entrada también podría estar ya codificada en HDR -por ejemplo,  $\{0, 1024\}$  o  $\{0, 65536\}$  con cualquier mapeo de tonos u otra imagen de significado de luminancia, o una imagen codificada de medio alcance, que aún puede necesitar procesamiento de acuerdo con las presentes invenciones. En realidad, la persona experta debería ver esta imagen esquemática como si la tecnología descrita solo se aplicara a una luminancia promedio (o mediana) para la imagen. En realidad, se puede aplicar cualquier operación compleja a cualquiera de los píxeles presentes en la imagen de entrada (especialmente para el análisis de la imagen en cuanto a dónde se producen los cambios de luminancia característicos, pero también para la (re) renderización de esos), pero para simplicidad de la explicación, describiremos solo los cambios (por ejemplo, la escala multiplicativa) de las luminancias, y describiremos esto como una escala de la retroiluminación (es decir, también los valores de manejo de la pantalla LCD pueden cambiar en correspondencia con los cambios de manejo del LED, pero actualmente ignoraremos eso en la descripción).

Un algoritmo de análisis humano o automático ha identificado momentos de tiempo a lo largo de la película, en los que la luminancia característica cambia (y, por lo tanto, la retroiluminación necesita o puede cambiar), tal como un momento importante de cambio TMA\_1 cuando la luminancia característica de la escena de conversación SCN\_CONVRS comienza, y los instantes de tiempo de cambio menor TMI\_1 y TMI\_12 (etc.) en esa escena para los cambios entre los disparos más claros y más oscuros SHT\_WND y SHT\_WLL (n.b. algoritmos automáticos más simples pueden estar limitados a determinar solo los instantes de cambio de tiempo importantes). La mayoría de las realizaciones simples de la presente invención solo codificarán dichos instantes de tiempo, y si se permite cualquier procesamiento de HDR (por ejemplo, a través de un ALCORR booleano, que prohíbe o fuerza a un escenario muy básico el procesamiento del lado de visualización de HDR si es igual a 0, pero permite, por ejemplo, una estrategia de impulso inteligente si es igual a 1). Esto permite que la pantalla (o el aparato de preprocesamiento, tal como el reproductor de bluray, decodificador, ordenador, etc.) pueda aplicar un procesamiento inteligente en lugar de aplicar ciegamente su algoritmo único, independientemente de cual sea el contenido de la película actual, sus intenciones artísticas de los creadores o su contenido futuro.

En la figura 3 se ilustran esquemáticamente algunos procesamientos de imagen deseados, así como varias posibles realizaciones de codificación de datos.

Psicovisualmente, tenemos diferentes necesidades de renderización para la escena "estática" (estado estable) SCN\_CONVRS y una segunda escena SCN\_WLK en la que una de las personas entra primero en un corredor oscuro, y luego (alrededor del tiempo de cambio TMA 2) llega a la luz del sol brillante. Los artistas pueden desear que el renderizado cree un aspecto final específico, pero no tiene suficiente control con solo los píxeles de la imagen que se codifican por sí mismos. En particular, los valores capturados de la cámara después de configurar la exposición de la cámara pueden ser realmente similares para escenarios con diferentes intentos de renderizado, como los dos ejemplos anteriores, especialmente para vídeos heredados. En ese ejemplo, la ventana que entra y sale de la vista es más una molestia que un efecto deseado; sin embargo, dependiendo de las capacidades de visualización, es posible que desee hacer algo emocionante con la persona que camina entre el interior y el exterior. Pero si la combinación de operador de cámara y niveladora es una codificación más directa de la escena en lugar de hacerlo de forma brillante, es posible que todavía se quede con valores de color/luminancia de píxeles bastante similares en ambos escenarios. Agregando el problema de cómo una pantalla intenta lidiar con eso ("a ciegas") parece conveniente tener un mecanismo adicional de codificación de la información y, de preferencia, el control. En SCN\_CONVRS, aunque los artistas pueden desear mostrar la diferencia en la iluminación hasta cierto punto (que puede incluir diferentes modos de retroiluminación, así como diferentes histogramas de píxeles para la imagen para la pantalla LCD, y como codificación de entrada o bien una señal HDR total diferente, o diferentes sugerencias para el ajuste de la retroiluminación, además de una codificación de imagen estándar), lo hará bajo el supuesto de que el ojo se ajusta en gran medida a la situación para ambos tipos de disparos intercalados. Es decir, el hecho de que el espectador cuando está fuera durante un cierto tiempo, o mirando a la persona que está frente a la ventana durante un cierto tiempo, haya ajustado sus características de sensibilidad retiniana, debe reflejarse en la codificación de la imagen que se va a renderizar, pero lo más importante en la propia imagen renderizada. En particular, una luminancia característica para la imagen de visualización emitida (y típicamente una luminancia a la que el ojo responderá con su bioquímica, por ejemplo, intentos de estabilizar y codificar diferencias en un estado de equilibrio constante, por ejemplo, luminancia promedio), debe ser tal que la escena no es molesta, parpadeante, visualmente fatigante, etc. Esto no fue tanto un problema con pantallas antiguas con sus rangos limitados, sino con pantallas brillantes emergentes (incluso con contenido LDR), y se convertirá especialmente en un punto de atención para futuras pantallas HDR. Por lo tanto, la pantalla puede querer mantener la diferencia en la luminancia característica limitada para escenarios como SCN\_CONVRS, por ejemplo, no realiza demasiado (las señales y, por lo tanto, su diferencia) de manera demasiado excesiva (es decir, usa un subrango pequeño del rango total de HDR R\_HDR para todos o la mayoría de los píxeles de vídeo en todas las imágenes de esa escena), o incluso reduzca su diferencia (es decir, como salida de pantalla) en luminarias, en comparación con lo que un PS\_R\_LDR característico de bajo rango dinámico daría si se emitiera, por ejemplo, la emulación de PS\_R\_LDR en la pantalla HDR muestra cómo una pantalla LDR de 500 nits renderizaría esa parte temporal de la película, mediante el mapeo dentro de ese rango [una realización de ejemplo de mapeo de la gama adaptada temporalmente]. Por ejemplo, la renderización o el aparato de cálculo de señal puede reducir el rango local, o la dispersión de luminancia de algunas partes de los píxeles de vídeo que caen dentro de ese rango (por ejemplo, los píxeles que más contribuyen al resumen de luminancia característica), por ejemplo, puede disminuir las luminancias de al menos algunos de los píxeles (por ejemplo, la vista recortada del mundo exterior). Además, puede aumentar la luminancia de las partes más oscuras del entorno, al menos la región en la que reside la segunda persona. La persona experta entenderá cómo las modificaciones de luminancia por píxel modifican la luminancia característica y viceversa, lo que puede hacerse mediante varios métodos simples o estadísticos.

Al contrario, para SCN\_WLK, el artista desea generar una variación de brillo dinámico. La persona que camina por el corredor primero sufre de ceguera por insensibilidad retiniana (porque el ambiente es más oscuro que su estado de adaptación), y después de haberse adaptado, cuando sale, se ve cegada por la luz brillante del exterior (ceguera por sobreexposición). Es posible que el artista ya haya simulado esto hasta cierto punto (o incluso la autoexposición de la cámara, pero supondrá para la discusión actual que, al menos para películas de calidad y no para la televisión en vuelo, el artista tiene eso bajo control), incluso con una señal LDR {0, 255}, por ejemplo al hacer que la imagen sea alta (muchas regiones brillantes, tal vez incluso sobreexpuestas con bajo contraste, es decir, un histograma que se encuentra en gran parte en la mitad superior de {0, 255}). Sin embargo, tal imagen/vídeo puede verse mejor (la

renderización de alto brillo da un aspecto diferente a la precisión de codificación de ciertas regiones, por ejemplo, debido a la dependencia de las diferencias humanas apenas perceptibles JND en la luminancia local) o al menos más convincente, cuando (en Además, o incluso predominantemente en su lugar) hay un cambio de brillo real. Esto se puede realizar, por ejemplo, por lo general al aumentar la retroiluminación (dejando la señal LCD sin cambiar, por ejemplo, la señal de entrada {0, 255} que se utiliza como estimación de las reflectancias del objeto, o adaptada, por ejemplo, reprocesada de manera óptima, coincidiendo con el cambio de la retroiluminación [que puede ser diferente para diferentes pantallas con diferentes capacidades]). De manera similar, hacer que la señal sea oscura puede emular la incapacidad visual (solo se pueden ver las diferencias más grandes antes de la adaptación, por lo que se puede emular codificando menos valores), pero el impacto visual real tendrá lugar cuando la retroiluminación también se oscurezca, o en general, la pantalla HDR utiliza su estrategia de renderización óptima para este tipo de tomas oscuras.

Entonces, para SCN\_CONVRS uno quiere que la pantalla "no haga nada" (o al menos no aplique un impulso serio, o incluso que el estiramiento se vincule inevitablemente al mapeo predeterminado en rangos de luminancia de salida más grandes) mientras que para SCN\_WLK uno quiere usar al máximo las capacidades de la pantalla (HDR), mediante la aplicación de un (efecto de visualización) efecto visual (por simplicidad aquí descrita como principalmente un cambio de manejo de retroiluminación). De manera similar, para una tercera escena SCN\_EXPL con explosiones, uno quiere aplicar un efecto diferente, y la renderización también debería ser preferiblemente diferente para los diferentes tipos de explosiones capturadas (ver más abajo).

La diferencia limitada deseada entre los valores de píxeles de las imágenes respectivas de tipo SHT\_WND frente a SHT\_WLL se puede especificar en un número de maneras, y si la operación de procesamiento de imagen para lograrlo es del tipo de cambio de luminancia (multiplicativa o aditiva) en general, se puede especificar de manera similar en una representación de luminancia característica (es decir, en la práctica uno simplemente incorporará las enseñanzas actuales en las operaciones de procesamiento de imágenes clásicas que funcionan en conjuntos de píxeles).

Por ejemplo, el artista (o el algoritmo de anotación automática (supondremos en la explicación adicional de que todas las codificaciones de las realizaciones inventadas están determinadas por un humano, pero la mayoría de ellas también pueden determinarse automáticamente aplicando el análisis de imágenes) puede especificar un rango R\_CORR para las luminancias características que deben ocupar (posiblemente aumentadas con especificaciones adicionales en los histogramas de los objetos de la imagen, tales como un intervalo de luminancias, o un rango para que caigan las luminancias superiores y/o inferiores, etc.), que puede determinarse en relación con un rango de visualización de referencia, por ejemplo, un rango dinámico bajo particular PS\_R\_LDR, o un rango dinámico alto de referencia, etc. Las pantallas que tienen un rango dinámico de visualización real pueden entonces hacer su procesamiento para que las luminancias de salida se vean como conformes con el rango de especificación posible, por ejemplo, una pantalla con un rango dinámico más alto puede asignar un subintervalo para emular el rango dinámico bajo de referencia, o en general cualquier pantalla puede aplicar un procesamiento que resulte en una salida que se desvíe mínimamente del aspecto/rango deseado. La similitud en la luminancia característica (y los histogramas de píxeles subyacentes) también se pueden especificar de otras maneras, por ejemplo, como un porcentaje de cambio permitido o preferido (arriba y/o abajo) que puede usarse para el procesamiento multiplicativo. "Permitido" se refiere a que el artista no permite mayores desviaciones que una cierta cantidad, y el procesamiento por parte de la pantalla debe al menos tratar de ajustarse a esto, o completamente para que sea compatible, mientras que "preferido" da una preferencia por artista, y es posible que la pantalla solo desee tener en cuenta dichas indicaciones al realizar su propio procesamiento [por ejemplo, calculando nuevos valores de manejo para el entorno de visualización actual, las preferencias del espectador, etc.], de modo que al menos intente obtener un aspecto similar, pero puede desviarse de eso.

Para el ejemplo de SCN\_WLK, la pantalla puede querer aplicar un perfil, determinado al menos parcialmente por el instante de tiempo codificado. Por ejemplo, se sabe que el ojo humano se adapta temporalmente aproximadamente de acuerdo con una curva exponencial, por lo que las mediciones como JND también seguirán esta curva. La pantalla puede aumentar la retroiluminación, por ejemplo, con una función exponencial ascendente EXPU, u otra función que primero exagera la luminosidad, pero luego se relaja nuevamente a una luminancia característica más baja, que por un lado simula la adaptación del espectador a un brillo de pantalla cómodo, pero por el otro también pone el manejo en algún lugar en una región central del rango total R\_HDR, de modo que todavía haya suficiente espacio en el rango de visualización sin límites para la renderización de otros entornos, por ejemplo explosiones. Esto funciona porque psicovisualmente el efecto es mayor cuando se produce por primera vez, y luego el ojo del espectador comienza a compensarlo parcialmente, por lo que no hay necesidad de gastar esta potencia de retroiluminación adicional, ya que aporta menos a la experiencia visual.

De manera similar, para el corredor oscuro, se puede aplicar una EXPD exponencial descendente.

En la actualidad, dado que la mayor parte del enfoque de la calidad de imagen (e incluso HDR) ha estado en el brillo adicional, la renderización de escenas oscuras ha recibido menos de la atención necesaria. La figura 4 elabora con un posible ejemplo de cómo con la presente invención puede hacer una renderización mejorada en pantallas para tales escenas (que ahora en gran medida caen por debajo de lo que es visible, y mucho menos que se logra el efecto de renderización deseado). Por ejemplo, la capa oscura de una persona en un entorno algo más brillante pero aún oscuro en la imagen, solo se reproducirá fielmente, no solo si las condiciones de visualización actuales (visualización

en pantalla y entorno de visualización) son buenas, sino también si el espectador está correctamente adaptado. Para ello, las imágenes sucesivas renderizadas en el pasado pueden prepararse para ese estado de adaptación, al disminuir las luminancias de los píxeles en esas imágenes previas, preferiblemente de manera gradual, de modo que no sea demasiado perceptible, o al menos no demasiado objetable (para el espectador o el artista). La pantalla podría hacerlo automáticamente al saber qué nivel de luminancia característica CHRLUM debería alcanzar en el futuro para la escena oscura SCN\_DRK, o se pueden especificar las disminuciones exactas o preferidas/aproximadas en o junto a la señal de vídeo (por ejemplo, por el artista). Es posible codificar instantes de tiempo particulares TMSK\_1, TMSK\_2 en o durante los cuales dicha disminución ocurre preferiblemente para hacerla menos perceptible, por ejemplo, en los límites del disparo, o para los píxeles envolventes que rodean una cara cuando se espera que el espectador vea la cara, etc. Además, el rango alto R\_EFF1 para objetos brillantes, tales como las explosiones (en la noche) se puede reducir gradualmente (de modo que, por un lado, el espectador ya no se adaptará demasiado a ellos, pero por otro lado tampoco se destacará demasiado de las otras luminancias, por ejemplo, lo que lleva a una experiencia de visualización exagerada o incluso cegadora). El algoritmo de análisis del contenido temporal (ya sea hecho por la pantalla receptora por sí misma, o un análisis de imagen para codificar de manera más precisa en datos adicionales cómo debe renderizarse la pantalla receptora) también puede ver un presupuesto de brillo visual de ciertos segmentos de tiempo, que codifica tales aspectos como: qué tan brillantes son los objetos (por ejemplo, las explosiones), qué tan grandes son, cuánto duran, cuántos hay en la sucesión, cómo contrastan con los subperíodos más oscuros, etc. Pequeñas iluminaciones cortas y brillantes pueden por ejemplo, asignarse aún a un rango superior R\_EFF2, mientras que las regiones brillantes que tienen una mayor influencia en la adaptación visual se renderizarán en el rango alto decreciente R\_EFF1. También puede cambiar la relación entre la luminancia característica CHRLUM y los valores de píxel de la imagen subyacente dentro del rango SP\_LUM. Por ejemplo, uno puede derivar de la entrada que codifica una estimación de las reflexiones aproximadas del objeto, produciendo una imagen de salida, y sobre la base de esa imagen (o cualquier derivación de la imagen de entrada) se aplica una transformada inclinada que hace que los objetos brillantes se vuelvan un poco más oscuros y posiblemente también más oscuros objetos. En realidad, como una aproximación simple, puede considerarse la luminancia característica como se describió anteriormente como una determinación de valor 1 de un rango de luminancias (digamos el promedio), pero otras realizaciones podrían codificar adicional o alternativamente otras medidas que caracterizan las luminancias en (un subconjunto de) una imagen, por ejemplo un rango (el procesamiento de un rango bajo a un rango alto aumentado mayor, por ejemplo, puede especificarse basándose en los límites respectivos de esos dos rangos). En general, llamaremos conceptualmente a todas estas codificaciones de resumen posibles características de luminancia, pero para mantener la explicación simple, nuestra explicación se limita a las caracterizaciones de 1 valor. Este sistema funciona particularmente bien en cooperación con las lámparas 150 de iluminación envolvente controlables, como las lámparas de colores vivos de Philips. Estas lámparas pueden equiparse con un controlador 151 de manejo (por ejemplo, que se comunica de manera inalámbrica), que puede ser accionado por cualquier unidad de la disposición (por ejemplo, controlada por la pantalla) en función de los datos adicionales codificados de acuerdo con la invención. Por ejemplo, en el primer instante, TMSK\_1 puede apagarse o atenuarse al 50%, o más inteligentemente, puede atenuarse de acuerdo con la luminancia característica CHRLUM. En general, las luces pueden configurarse de manera óptima dependiendo de las caracterizaciones de la luminancia del vídeo temporal.

Volviendo a la figura 3, un perfil puede ser derivado por la propia pantalla, o preferiblemente, está codificado por el lado de producción de contenido. La señal puede incluir varias clases predefinidas bajo un código PROF de perfil (que el lado receptor, su IC de decodificación y su IC de procesamiento pueden comprender y manejar), por ejemplo, MULT multiplicativo, lo que significa que durante todo el disparo (o, de hecho, un período temporal entre dos códigos instantáneos de tiempo) solo una escala multiplicativa (ya sea de la propia señal HDR, o de la parte descompuesta correspondiente a la retroiluminación, o una combinación de ellas) puede ser aplicado. El perfil de modificación se puede parametrizar aún más, por ejemplo, siendo P1 la cantidad a disminuir (por ejemplo, 80%) y P2 la cantidad a aumentar (por ejemplo, 120%). Diferentes parámetros aún permiten que diferentes pantallas elijan una u otra opción. Para SCN\_WLK, el tipo de perfil es un EXP exponencial, que el codificador de contenido puede complementar con parámetros tales como una amplitud de inicio A y un tiempo de decaimiento TD. En general, un aparato del lado receptor también puede determinar por sí mismo el intervalo de tiempo DTI durante el cual una desviación de un primer valor (tal como una representación de luminancia característica del vídeo de entrada, o de acuerdo con un primer conjunto de algoritmo calculado de manejo de retroiluminación para esa entrada de vídeo) es necesaria, por ejemplo teniendo en cuenta la información sobre el futuro del vídeo (ver más abajo). Tal exponencial también se puede usar para disminuir la luminancia de una explosión que dura mucho tiempo, por ejemplo, porque está congelado artísticamente en una representación de tiempo estirado. Aunque la señal original puede codificar esa explosión en todos sus detalles, incluida su luminancia capturada original (porque así es como la cámara con sus ajustes de exposición la siguió grabando), incluida la exponencial permite reducir la luminancia de esa explosión gradualmente, sin afectar negativamente a la calidad visual, sin embargo permitiendo por ejemplo para una reducción de potencia (un equivalente temporal, impulsado por el contenido, de lo que uno haría de forma estática).

Además, con algunos de estos perfiles básicos (por ejemplo, segmentos exponenciales o lineales), el proveedor de contenido o el codificador (de material codificado previamente) puede producir perfiles temporales complejos. Estos se pueden utilizar, por ejemplo, para aplicar efectos HDR a material heredado. Por ejemplo, una película heredada puede contener una escena de una supernova con anillos de gas caliente que se ondulan hacia el exterior, que sin embargo se codificaron claramente en {0, 255}. En lugar de aplicar una predicción (o renderización) de gráficos de ordenador completa para llegar a una codificación HDR {0, 65K} de esa escena, se pueden aplicar los perfiles

temporales de la presente invención (típicamente, pero no exclusivamente, para activar la retroiluminación) en un momento determinado, por ejemplo, TMA\_3 después de lo cual se requiere el efecto HDR. Al permitir la codificación de dicho perfil temporal (casi) arbitrario, y también variable en el espacio, se puede, por ejemplo, definir un perfil multisinusoidal espaciotemporal ondulado hacia afuera en el componente de imagen destinado a la retroiluminación del LED en sincronización aproximada con la nube de gas más luminosa en la imagen original (o procesada) {0, 255} para activar la pantalla LCD. Para las caracterizaciones espaciales, SPATPROF de los perfiles temporales se puede codificar, por ejemplo, una forma multiplicativa, como una constante multiplicativa definida en un círculo con origen (x, y) y radio r1. Pero de manera más interesante, las codificaciones de mapas temporales MAP pueden ser codificadores, que pueden ser bidimensionales o tridimensionales. Esto se puede hacer, por ejemplo, tomando una resolución de referencia para una imagen de retroiluminación (por ejemplo, 50x50, que puede comprender diversos rangos de aspecto desde el lector de datos 2:3 posicionado verticalmente, hasta el televisor 21:9), que se puede remuestrear para obtener una retroiluminación de pantalla real. Este mapa puede incluir, por ejemplo, números binarios para las regiones (1 = carga pesada durante un período de tiempo próximo, 0 = carga menos severa), o manejo local acumulado (que se puede usar para predecir y contrarrestar el calentamiento local, el envejecimiento, etc.). En ese caso, se codifica una matriz bidimensional de números, por ejemplo, {10, 10, 10, 40, 55, 45, ...} son las luminancias integradas de los LED de referencia hasta el siguiente instante de tiempo codificado, o los siguientes 10 minutos fijados en un intervalo acordado, etc. Un MAP tridimensional puede incluir mucho más interesante datos espaciotemporales locales (parámetros o una función temporal espacial local real), como los que se pueden utilizar, por ejemplo, para la codificación de efectos. En el primer caso, el mapa solo contiene datos de medición, que pueden utilizarse como datos informativos interesantes para la optimización de la pantalla, por ejemplo, en relación con, por ejemplo, su imagen de la gestión del calor se procesa, mientras que en este último caso puede guiar o incluso ordenar el manejo, por ejemplo, los LED de retroiluminación del mapa de retroiluminación. Cabe señalar que cualquier estrategia para modular temporalmente una retroiluminación también se puede convertir en un único manejo HDR (por ejemplo, para una pantalla OLED) y viceversa, por lo que cualquier realización de codificación también se puede emplear para el reprocesamiento (guiado) en el espacio de color (HDR) para el manejo de la pantalla.

Un número de tipos de renderización HDR TIPO (cambiando la iluminación de escena ILLUMCHNG, cambiando el entorno de filmación local ENVCHNG, efectos como explosiones EXPL, etc.) también se pueden acordar de manera estándar para la comunicación entre la producción de contenido y la renderización de la pantalla, y el vídeo puede contener una secuencia de comandos que si la pantalla necesita o quiere, por ejemplo, reducir la retroiluminación para ahorrar potencia (por ejemplo, en un modo ecológico de menor interés para el espectador), omita los efectos de caminar desde un entorno oscuro a uno luminoso, pero no las explosiones (o más o todos los cambios del entorno antes de que comience la manipulación de las explosiones). O bien, se pueden poner límites en las renderizaciones del tipo de iluminación de escena, etc. Parámetros adicionales pueden ayudar en el procesamiento por parte de una unidad en el lado de la pantalla, por ejemplo, las estadísticas locales LOCSTATS puede indicar que el mayor problema de una luminancia de píxeles demasiado alta se encuentra en una región recortada (ventana) por encima del código de luminancia de 250 píxeles, que se puede aplicar una mayor deformación de color para valores de píxeles codificados originales por encima de 200, etc.

Otra realización útil permite determinar una jerarquía de renderizaciones temporales (por ejemplo, efectos tales como explosiones). Para ello se puede codificar un nivel de importancia IMPLEV. Mirando las tres explosiones sucesivas de SCN\_EXPL podemos ver dos cosas. En primer lugar, muchas explosiones tras otra pueden no tener un impacto tan grande en el espectador (y este impacto dependerá en gran medida de la pantalla y del entorno de visualización, por ejemplo, en un visor de películas móvil con un sonido envolvente brillante, uno podría tener dos explosiones brillantes bien espaciadas [tal vez incluso con mayor duración], con una modulación oscura más profunda en el medio, que tres explosiones concatenadas casi idénticas, agregando solo una diferencia perceptiva entre sí, cuyo efecto solo puede verse placenteramente en las pantallas del extremo superior y bajo mejores circunstancias de visualización). En segundo lugar, puede haber un consumo excesivo de potencia e incluso sobrecalentamiento cuando se estira la pantalla hasta su límite con tantas explosiones una tras otra, es decir, el contenido del vídeo puede estar en discordancia con las restricciones físicas de la pantalla.

Las explosiones aumentan en la luminancia característica (por ejemplo, la luminancia promedio sobre la bola de fuego, o una luminancia de una muestra característica de la bola de fuego). En el rango de codificación de la imagen de entrada original (o cualquier derivación de la misma) puede que ya no haya demasiado espacio para codificarlos. Normalmente, las luminancias capturadas cerca del límite del rango de una codificación se codifican progresivamente de forma no lineal (recorte suave). Dicha función puede estar codificada o estimada en el lado de decodificación (o transcodificación), aunque sea aproximadamente. De todos modos, las luminancias finales para la salida de la pantalla pueden estar más separadas, si hay un amplio rango para los efectos (R UPEFF). Sin embargo, en vista de la disminución de la sensibilidad y el impacto para el espectador humano, es posible que se requiera un aumento adicional de las explosiones, y que una gran cantidad de explosiones sucesivas ya no se ajusten al rango disponible R UPEFF. Un concepto útil es una "diferencia bien notable" WND. Esto se puede definir, por ejemplo, como un número de JND, y forman una base para una escala de impacto para la codificación por el artista. Los procesos a aplicar pueden hacer uso de impactos codificados, por ejemplo, como una guía para un número de WND entre explosiones sucesivas. Esto se puede hacer a través de la codificación PROF de perfil, o más explícitamente mediante codificaciones de procesamiento permitido ALWDPROC, por ejemplo, un mapeo de tonos en la mitad más brillante de la imagen.

Pero también, el nivel de importancia permite eliminar, o decolorar seriamente ciertas renderizaciones temporales. Mientras que la primera y la tercera explosión tienen IMPLEV=1, la segunda tiene IMPLEV=2. Esto significa que puede caerse (o deformarse) para hacer espacio en el rango de luminancia para proporcionar una experiencia visual más  
 5 óptima con la explosión inicial y final. Además, si por otros motivos se requieren cambios en la renderización, como la atenuación de la retroiluminación (local), la pantalla puede comenzar con el nivel de importancia más alto, por ejemplo, IMPLEV=3, luego IMPLEV=2 intervalos de tiempo, etc. Para no reducir o deformar completamente el impacto visual de la segunda explosión, lo que se pierde en el manejo de la retroiluminación puede compensarse parcialmente al hacer los valores de píxeles de la imagen para la pantalla LCD. excesivamente brillante. Esto se puede hacer  
 10 automáticamente en la pantalla mediante una compensación de imagen de LCD aproximada, o codificada explícitamente por procesamiento de mapeo de tonos específicos ALWDPROC. También se puede simular el impacto visual cambiando localmente los parámetros cromáticos de la pantalla LCD o la imagen de la retroiluminación, por medio de un COL de especificación de color, que puede comprender, por ejemplo, una diferencia de matiz para el objeto principal o región (en este caso, la explosión) HM, una diferencia de saturación para el objeto principal o región  
 15 SM, y una diferencia de matiz y saturación para una región circundante, por ejemplo, el resto de las imágenes.

Son interesantes para el procesamiento de imágenes basado en restricciones físicas tales parámetros relacionados con la luminancia característica futura NXTLD, como el tiempo hasta la siguiente luminancia característica excesiva  
 20 DT (por ejemplo, con una retroiluminación de referencia predicha que supera el 80%), una duración de luminancia característica excesiva DUR, una luminancia o potencia promedio gastada en un intervalo de tiempo en el futuro PAV, etc. Dicha información, por ejemplo, el tiempo para cambiar el intervalo de luminancia característica, se puede usar para determinar los perfiles temporales mediante la visualización, mediante el modelado de fórmulas, por ejemplo, Visión humana o comportamiento energético de la pantalla. Por ejemplo, uno puede calcular un perfil de atenuación de la retroiluminación basándose en una especificación final derivada de la carga de la retroiluminación, por ejemplo,  
 25 los próximos 30 segundos o 5 minutos, y escalar una exponencial basada en, por ejemplo, un número de valores o clases de la especificación final.

La figura 5 muestra esquemáticamente una realización de ejemplo de una disposición para un graduador 599 de color de postproducción de película (o puede estar operando una anotación semiautomática de vídeo heredado), dispuesta  
 30 para codificar los diversos datos de anotación de acuerdo con la presente invención (no ayudará la concisión enumerando todas las posibilidades, pero la persona experta puede determinarlas por analogía a partir de los ejemplos descritos). Notamos que, en principio, también los aparatos automáticos pueden encarnar las presentes realizaciones, por ejemplo, un dispositivo de preprocesamiento en el hogar para optimizar para una pantalla en particular durante la noche una película obtenida, sin embargo, vamos a ejemplificar los conceptos con graduación humana. El graduador de color tiene un aparato 500 de clasificación, que comprende medios 501 de entrada del usuario para especificar  
 35 diversas selecciones, parámetros, etc., típicamente con botones con significados fijos como "próximos 10 segundos de vídeo", "mostrar/ocultar la jerarquía actual del tiempo menor instantes", "agregar una marca de tiempo instantánea", un teclado alfanumérico, botones giratorios para determinar un tono o avanzar en una selección temporal de imágenes de vídeo clave sucesivas, etc. También tiene diversas pantallas, por ejemplo una pantalla 511 de referencia HDR y  
 40 una pantalla 512 para el análisis temporal de la película. Por ejemplo, se muestra un número de fotogramas clave, con un perfil de luminancia característico precalculado, y el clasificador de color puede basar este inserto en sus instantes de tiempo, hacer doble clic en ellos para abrir una página para codificar datos adicionales, escribir su tipo ILLUMCHNG y datos adicionales como los parámetros estadísticos, que pueden suministrarse fácilmente mediante aplicaciones auxiliares que se ejecutan en pantallas adicionales, como un plano de color. Todos los precálculos automáticos se  
 45 realizan mediante una unidad 520 de análisis de imagen, que determina diversos parámetros descritos anteriormente, por ejemplo, un perfil de luminancia característico, si se desea instantes de tiempo iniciales para cambios en el perfil de luminancia característico, codificaciones iniciales de otras codificaciones como los perfiles preferidos que se aplicarán en el lado de renderizado. El graduador de color puede entonces aceptar o descartar fácilmente las propuestas, y en el segundo caso vienen con sus propias versiones provistas por el hombre. Una unidad 522 de  
 50 aplicación aplica todas las codificaciones aplicables actuales, y puede enviarlas a través de una subunidad 5221 de visualización a las diferentes pantallas (por ejemplo, una selección de imágenes clave óptimas alrededor de los instantes de tiempo determinados TMA\_1, TMI\_1... para ver en la pantalla 512, y un aspecto final para ver en la pantalla 511 de referencia HDR. Uno de los botones en los medios de entrada del usuario está reservado para alternar rápidamente entre diferentes pantallas de referencia típicas de lo que hay en las habitaciones de estar del consumidor,  
 55 por ejemplo, una pantalla de 500 nits simulada en la pantalla 511, una pantalla de 1000 nits simulada en 511, etc. Estas simulaciones pueden incluir varios escenarios (por ejemplo, el peor de los casos) del procesamiento de la imagen que posiblemente se aplique a una pantalla en particular, como un modo ecológico o una nitidez. El graduador de color puede ver rápidamente el impacto de todas sus decisiones, ya sea que codifique de forma sencilla una guía única que permita que las pantallas apliquen aún una cantidad muy variable de procesamiento, lo que lleva a  
 60 renderizaciones finales muy diferentes, o si codifica con mayor precisión un conjunto de especificaciones para diferentes escenarios de referencia (por ejemplo, pantalla LDR antigua, pantalla de rango medio, .../oscuro frente a brillante envolvente...), que la pantalla debe tratar de cumplir con la mayor precisión posible seleccionando la especificación más adecuada. Finalmente, un codificador 524 codifica todos los datos de acuerdo con cualquier formato prescrito (por ejemplo, la señal de vídeo + datos codificados COD), y lo envía a través de una salida 530 a  
 65 una conexión 531, por ejemplo, a un dispositivo de almacenamiento, desde el cual luego, por ejemplo, se graba un BD o DVD, o qué codificación final (vídeo + datos adicionales) se envía por separado o en conjunto, por ejemplo, un

proveedor de contenido de cable a través de un satélite, etc. El codificador puede codificar los instantes de tiempo en formatos predefinidos (vea un ejemplo a continuación), y puede comprender además un formateador 5241 de indicación de estrategia de reprocesamiento para codificar lo que se puede hacer alrededor de los instantes de tiempo en el lado receptor en formatos predefinidos. Por ejemplo, el codificador puede escribir en un número de campos (por ejemplo, 10 campos reservados) un número de índice de un tipo de procesamiento que se puede realizar (por ejemplo, campo 1 = "1" significa disminución lineal de la luminancia de salida prevista actual con pendiente en el campo 2 = "x").

En el lado receptor, se puede incorporar una unidad 600 de manejo/decodificación de vídeo (como se muestra en la figura 6), por ejemplo, en un aparato de procesamiento de vídeo que comprende una unidad de lectura de disco (en el ejemplo de la figura 1, esta unidad de manejo de vídeo es un IC 112 en el reproductor 110 BD, pero también puede estar comprendido en el televisor, un ordenador en una red doméstica conectada a la pantalla, etc.). El IC 112 y/o el reproductor de BD pueden generar como salida una señal adecuada para controlar una pantalla, por ejemplo, una imagen de salida que codifica IMOUT que comprende un componente de imagen de manejo de retroiluminación ILIM y una imagen de manejo LCD LCDIM. La unidad de manejo de vídeo recibe vídeo codificado a través de la entrada 640 (por ejemplo, desde un lector de BD, un decodificador de conexión de cable, etc.) y comprende un decodificador 624 que puede decodificar el vídeo (típicamente compatible con versiones anteriores, es decir, codificado de acuerdo con un estándar MPEG como AVC), así como los datos adicionales de acuerdo con las realizaciones de la presente invención, como los instantes de tiempo TMA\_1... de los cambios en la luminancia característica, y la especificación adicional de las codificaciones de dichos intervalos de tiempo y el vídeo en él (tipo, estadísticas del vídeo, mapeo de perfiles a aplicar, etc.). La unidad 600 de decodificación de vídeo también suele recibir y decodificar información relacionada con el reprocesamiento de luminancia/color alrededor o definido por o con respecto a los instantes de tiempo TMA (por ejemplo, TMA puede definir una operación para mucho más adelante). La unidad 600 de manejo de vídeo también puede estar dispuesta normalmente incluyendo un analizador 620 de vídeo para realizar su propio análisis del de vídeo decodificado VID, para aplicar su propio procesamiento de vídeo (por ejemplo, la pantalla puede preferir aplicar su propio efecto particular aumentando incluso ignorando las especificaciones de perfil de la presente invención, pero aún así esto puede ayudarse al menos conociendo instantes de tiempo interesantes TMA\_1; también se puede aplicar un procesamiento de vídeo posiblemente menos relacionado con la presente invención, como la mejora de la textura del césped). El procesamiento de vídeo final, basado parcialmente en el análisis propio del analizador 620 de vídeo, y parcialmente en los datos adicionales descodificados DD de acuerdo con cualquier realización de la presente invención, se realiza mediante el procesador 630 de vídeo, y la codificación de vídeo resultante (típicamente para el ejemplo de la figura 1, una imagen de manejo LCD y LED se envía a través de la salida 650 a la pantalla. También hemos mostrado esquemáticamente en guiones una pantalla 6000 conectada (por supuesto, la unidad 600 de manejo de vídeo con capacidad de decodificación podría estar conectada a otro aparato, por ejemplo, un transcodificador o dispositivo de almacenamiento, etc.). En caso de que se conecte una pantalla inteligente, por lo general, la unidad 600 de manejo de vídeo emitirá aún una gran cantidad de la información original DD (incluso si ya creó su propia señal de vídeo óptima), por ejemplo, una especificación SPECFUT de cómo las luminancias características van a cambiar en al menos uno o varios segmentos de tiempo futuros. La pantalla puede usar estos datos para llegar a su propia señal final para renderizar en su panel 6001 de pantalla, por ejemplo, puede comprender un optimizador 6002 de experiencia de visualización dispuesto para determinar un manejo de vídeo óptimo de acuerdo con las preferencias de la pantalla.

Los datos adicionales DD se pueden codificar en una señal de diferentes maneras. Por ejemplo, en el encabezado principal al comienzo del vídeo puede abarcar la mayor parte del campo, por ejemplo una lista de instantes de tiempo TMA\_1... con detalles, tales como por ejemplo, si el televisor permite y qué procesamiento permite, por ejemplo, un campo que comienza con la palabra clave COL y 4 parámetros detrás (HM a SS). O DD puede comprender una composición de segmentos lineales que caracterizan un perfil de luminancia característico u otro perfil para las próximas imágenes, un LUT 3D con posiciones espaciales y como datos de puntos de una tercera dimensión de una curva o una lista dispersa de parámetros, etc. Pero también, por ejemplo, el encabezado de una imagen o GOP o grupo de bloques puede contener (generalmente menos) datos de un futuro próximo, tales como el tiempo hasta el próximo cambio de luminancia característica y su tipo. Por lo tanto, los datos DD se pueden codificar dentro de lo que se ve como la señal de vídeo VID (por ejemplo, utilizando estructuras de datos reservadas predefinidas y abiertas en él, tales como las SEI) o fuera de él (por ejemplo, en un almacenamiento separado y a través de una ruta de señal separada), pero relacionado con ello. Esta codificación se puede utilizar, por ejemplo, en un servicio en donde se envía una identificación del vídeo VID (por ejemplo, un título + otras especificaciones, o una marca de agua) a un proveedor de servicios, que luego envía o proporciona acceso a los datos adicionales DD. Para otro vídeo, como por ejemplo el vídeo capturado por el consumidor, la señal de vídeo completa VID puede enviarse al proveedor, pero para tal escenario (donde no hay datos de vídeo generalmente conocidos como con una película), el DD se almacena preferiblemente (quizás fuera de VID, pero) en conexión cercana con VID, por ejemplo, en el mismo almacenamiento extraíble, si no en el mismo disco duro, entonces el mismo almacenamiento en la red doméstica, etc. Esto será especialmente cierto si uno de los aparatos del consumidor (por ejemplo, decodificador, ordenador portátil) realiza el análisis de vídeo y proporciona los datos adicionales DD.

La figura 7 ilustra con más detalle una realización de ejemplo de lo que sucede matemáticamente cuando un aparato utiliza los datos adicionales para llegar a una renderización de vídeo deseada, y la relación detrás de una luminancia característica, y las luminancias de píxeles subyacentes de las imágenes de vídeo, y en particular su histograma.

Suponemos por simplicidad que la imagen de entrada (cuyo histograma se muestra en la parte inferior del gráfico, que aumenta la luminancia de los píxeles Lin a la izquierda entre 0 y 255) está codificada en LDR, y tiene una sola región brillante con histograma parcial BRLGHT. Esta imagen de entrada se puede caracterizar con una luminancia característica (que, como se dijo anteriormente, puede ser cualquier ecuación espacial/de valor en la distribución del valor espacial y/o de luminancia(/color) de los píxeles de la imagen de entrada que resume como se ilumina (física o perceptualmente) la imagen) CHRLUM\_i, que en este caso muestra que la imagen no es muy clara debido a su posición baja en el eje Lin (probablemente porque hay muchos píxeles oscuros, y las regiones brillantes no son predominantes, ni en cantidad ni en luminancia). Esta luminancia característica única, por lo tanto, define la imagen que es mayoritariamente oscura, aunque puede haber una región brillante (en general, se puede usar una caracterización más compleja, que comprende valores de luminancia característicos adicionales que describen la complejidad de la imagen actual o el disparo). Representar esta imagen original en el espacio de color HDR destinado a la visualización en pantalla con la luminancia de salida Lout (ya sea a través de una descomposición de la retroiluminación/transmisión, es decir, Lout posiblemente representa, por ejemplo, una imagen {0, 65K} totalmente codificada, o alternatively, el histograma de una la retroiluminación de la imagen de manejo), corresponde con una luminancia característica inicial (original, inicial) CHRLUM\_ini (por ejemplo, calculada con la misma ecuación que para la imagen de entrada en la imagen {0, 65K} resultante de un mapeo simple, tal como un simple estiramiento, o una función más no lineal que mapea las luminancias más oscuras aproximadamente a un rango de una representación estándar de 500nit [es decir, para manejo lineal (o compensada cualquiera que sea la gamma) que para una pantalla de 2000nit correspondería a una parte de la habitación más baja de los valores de manejo] y los objetos más brillantes se asignan a valores de luminancia más altos del rango HDR). La asignación inicial al rango HDR se conceptualizó como Lin\* (ejemplo mostrado para píxeles que no necesitan una desviación/aumento de luminancia). Sin embargo, queremos dar un ejemplo adicional impulso psicovisual a al menos las luminancias/regiones más brillantes, moviendo el histograma parcial BRLGHT hacia arriba en el eje Lout, que corresponde a una luminancia característica mayor CHRLUM\_o. Cabe señalar que aunque describimos todo lo relacionado conceptualmente con la luminancia característica para definir la invención, en realidad se puede realizar de varias maneras diferentes. Normalmente, el procesamiento de la imagen se corresponderá con operaciones tales como mapeos de tonos (locales) (TMO\_B), que generalmente varían con el tiempo (al menos en parte guiados por algunos datos recuperados de DD), como se puede ver claramente en el segundo histograma vertical para un tiempo (t2) posterior, para el cual el subhistograma parcial BRLGHT se ha movido un poco hacia abajo, correspondiendo a una luminancia característica más baja CHRLUM\_o (t2) [asumimos por simplicidad que el histograma de la imagen de entrada era el mismo en t2, de lo contrario también se reflejará en el histograma de salida, ya que generalmente solo la estrategia de mapeo de tonos que cambia en función del tiempo y los datos adicionales según lo prescrito por las realizaciones de la presente invención]. Como se dijo anteriormente, los mismos conceptos de procesamiento de acuerdo con esta invención también pueden caracterizarse adicionalmente o de manera similar, por ejemplo, observando, en los intervalos locales de los histogramas parciales SP\_I frente a SP\_O, etc. (es decir, donde se podría calcular una luminancia característica, una representación alternativa que mida lo mismo sería equivalente). Cualquier realización puede realizarse en operaciones simples y simplificadas, por lo que el reprocesamiento debe interpretarse en un sentido genérico de procesamiento. Cabe señalar que los instantes de tiempo TMA\_1, etc., también se pueden codificar en el vídeo a una velocidad más densa (y/o más equidistante), en cuyo caso les daríamos a algunos de ellos un código sin cambios, o al menos un código ALCORR "no permitido", o similar, ya que no son particularmente especiales (pero esto puede ser útil para tener una descripción más densa de algunas propiedades similares relacionadas con la luminancia o características similares, que es útil para controlar el procesamiento en una vecindad temporal, por ejemplo, por consideraciones energéticas, como el manejo a contraluz). Una concepción relacionada es la codificación de los cambios un poco antes de que realmente ocurran.

Debe entenderse que con el presente, uno puede realizar muchas renderizaciones de efectos ópticos, por ejemplo, efectos de brillo, etc. También se debe tener en cuenta que la presente invención se puede utilizar en relación con un de codificación de vídeo único VID (por ejemplo, aumento de una codificación LDR), pero también junto con varias codificaciones relacionadas (por ejemplo, una variante de LDR y HDR), y luego por ejemplo ser utilizados para relacionarlos. Por ejemplo, los instantes de tiempo pueden indicar segmentos de tiempo particularmente interesantes de similitud o disimilitud, y los perfiles de procesamiento de imágenes pueden ser tales para relacionarlos o hacerlos más o menos similares, o derivar nuevas renderizaciones en ambos, etc. Al menos algunas partes de los datos adicionales se pueden determinar, al menos en parte, en los datos de vídeo, o separados de ellos (aunque normalmente habrá alguna correlación, un humano puede prescribir alguna formulación específica). Además, la derivación de los datos adicionales, tales como instantes de tiempo, etc., se realiza preferiblemente a partir de cualquier señal HDR (por ejemplo, un grado HDR), pero también podría hacerse -por ejemplo como una estimación aproximada basada en codificaciones de vídeo derivadas.

Habiendo elaborado la presente invención con varias realizaciones, volvemos a la figura 1 para otras posibilidades con respecto a una disposición en la recepción de vídeo, y típicamente el lado de la pantalla. Varios otros aparatos pueden comprender al menos parte de los componentes de la invención, y contribuir a la invención, por ejemplo, un aparato 120 receptor de vídeo con almacenamiento puede conectarse a través de la conexión (inalámbrica o por cable) 121. Este aparato 120 receptor de vídeo puede aplicar su propio análisis y anotación de acuerdo con la invención, por ejemplo, desconectado durante la noche para descargar un programa de vídeo, por ejemplo, a través de una conexión 122 a internet 130, y que se verá más adelante, para crear una estrategia de manejo sensata para la retroiluminación de la pantalla conectada. Cabe señalar que a través de internet se pueden alcanzar ordenadores que contienen datos



de anotación de acuerdo con la presente invención (por ejemplo, de un proveedor de servicios fuera de línea), y el aparato 120 de recepción de vídeo puede incluso conectarse a través de internet a las fuentes de las cámaras 132 LDR o HDR.

La figura 8 describe una disposición de ejemplo con un primer aparato lateral (el primer lado todavía está típicamente en la misma ubicación que los otros aparatos de la disposición, por ejemplo, la casa de un consumidor, pero posiblemente operando en un momento diferente), y un segundo lado del aparato (por ejemplo, un televisor). En este ejemplo, incorporamos el primer aparato lateral como un aparato 800 de procesamiento de imágenes con una función de potencia, por ejemplo, puede ser un decodificador con almacenamiento que puede preprocesar una película (por supuesto, lo mismo puede ocurrir ya en la televisión, o en una parte de procesamiento en algún lugar del mundo, etc.).

Como se describe, por ejemplo, en el documento US7284874B [Jeong, LED backlight including cooling], las pantallas pueden calentarse, y especialmente si se muestran muchas imágenes brillantes, la retroiluminación puede calentarse mucho, en particular si el enfriador tiene que funcionar por encima de sus especificaciones, demasiado caliente. Sin embargo, se puede modelar cómo se elimina el calor de una región de la retroiluminación.

El aparato 800 de procesamiento de imágenes comprende un analizador 801 de vídeo, que está dispuesto para analizar el vídeo de una manera relacionada con el rendimiento térmico. Es decir, por lo general, tiene conocimiento de un modelo térmico y el impacto de un vídeo en particular, tales como explosiones o vistas brillantes en escenas al aire libre, en el rendimiento térmico de, por ejemplo, una caracterización de pantalla precargada (por ejemplo, un modelo térmico de la retroiluminación de un televisor conectado). Describimos una unidad de análisis algo más simple que simplemente envía caracterizaciones temporales "genéricas" que el lado receptor puede usar dentro de su propio modelado térmico, y alternativamente una unidad de análisis que ya determina en gran medida el comportamiento óptimo de manejo de la pantalla para la pantalla receptora. Una señal 820 de vídeo puede contener dos explosiones. Una caracterización temporal genérica puede describir al menos una explosión de este tipo, o en general un perfil de luminancia futura, con una función 821 de modelado particular. Por ejemplo, un pesaje aditivo lineal de luminancias características de un número de imágenes en el futuro (o regiones locales, etc.) puede ser calculado. Generalmente, dicho pesaje puede depender (para una pantalla de referencia) de la duración de una sobrecarga, ya que se supone que los periodos más largos tienen una mayor probabilidad de sobrecalentamiento. Es decir, el pesaje de explosiones de mayor duración puede ser mayor (la amplitud se incorpora de forma trivial). Los coeficientes de peso pueden ser recibidos, por ejemplo, desde la televisión/segundo lado. De todos modos, el televisor puede usar una caracterización temporal térmica del vídeo para determinar de manera más segura sus propios ajustes. Por ejemplo, un televisor que no tenga el beneficio de los datos adicionales presentados actualmente realizará su modulación de retroiluminación basada en el perfil 870 punteado. Simplemente seguirá el impulso, pero necesita atenuarse a la mitad porque se está sobrecalentando. No sabiendo que se avecina un segundo impulso, por razones térmicas se forzará a perder aún más brillo allí (haciendo que la segunda explosión sea menos brillante que la primera en lugar de más brillante). Con los datos adicionales, la pantalla puede utilizar una estrategia de manejo más inteligente, simbolizada con el perfil 871 de luminancia característico punteado. Es decir, puede atenuarse con menos molestia en la parte oscura antes de la explosión, y quizás algo en la primera explosión, para reservar un impulso para la segunda.

Alternativamente, el analizador 801 de vídeo puede simular modelos térmicos reales dados cuál sería el efecto de las modificaciones 822 reales (aproximadas), y prescribirlos como modelos de modificación, al menos para una guía tentativa de la pantalla. En cualquier caso, un optimizador 850 de manejo de la pantalla relacionada con la potencia determinará el manejo final de la pantalla, en función de los datos adicionales DD que obtenga. Realizaciones alternativas pueden especificar como datos adicionales DD, por ejemplo, una señal de advertencia que varía temporalmente, o un presupuesto térmico disponible, que especifica qué tan crítico (probable) es que una pantalla se sobrecaliente, etc.

La figura 9 proporciona una realización de ejemplo de la codificación de los datos adicionales presentes de acuerdo con la estructura SEI de MPEG4-AVC. Describimos el ejemplo de AVC como un ejemplo de transmisión desde un lado de "creación de contenido" a un lado de renderización de contenido, tal como un televisor para el consumidor, y un ejemplo CEA 861-D de una codificación entre, por ejemplo, dos aparatos de consumo, tal como un reproductor BD y un televisor, y el control o las posibilidades de información entre ellos.

El MPEG ha definido un contenedor de metadatos especiales específicamente para información de señalización adicional relacionada con el vídeo codificado. Este contenedor de metadatos se denomina mensaje de información de mejora suplementaria abreviado como mensaje SEI. El mensaje SEI se transporta en bloques de datos separados junto con los datos de vídeo en una secuencia (Unidad 901 SEI NAL).

Se construye un flujo h2.64 a partir de unidades NAL (Capa de abstracción de red). En h2.64 se definen varios tipos diferentes de unidades NAL, por ejemplo, una unidad NAL que contiene los datos de imagen codificados y una unidad NAL que contiene los mensajes SEI. Varias de estas unidades NAL forman una unidad de acceso. En una unidad de acceso, están disponibles todos los datos necesarios para comenzar a decodificar y presentar uno o más cuadros de vídeo.

Los instantes de tiempo de, por ejemplo, Las escenas excepcionalmente brillantes se pueden describir con valores PTS (sellos de tiempo de programación). Los valores DTS pueden indicar cuándo, suficientemente en el tiempo, el mensaje SEI debe decodificarse y enviarse a las subunidades que lo utilizan.

- 5 Una sintaxis de mensaje SEI de ejemplo para transportar un especificador 902 de aumento de brillo HDR puede tener el siguiente aspecto

Indicador de impulso HDR(tamaño de carga útil) {	No de bits	tipo
Bits de marcador	1	BSLBF
Tasa de cuadro	8	UIMSBF
Inicio de PTS	32	UIMSBF
Fin de PTS	32	UIMSBF
HDR_DTS	32	UIMSBF
Posición horizontal de la región	16	UIMSBF
Posición vertical de la región	16	UIMSBF
Ancho de la región	16	UIMSBF
Altura de la región	16	UIMSBF
Ganancia HDR	7	UIMSBF
Reservado para uso futuro	16	UIMSBF
}		

- 10 En este mensaje, las codificaciones tienen el siguiente significado:

Bits de marcador: Bits para indicar el inicio del mensaje SEI.

Tasa de cuadro: La velocidad de cuadros del vídeo asociado para calcular los valores de PTS para el reloj del sistema.

- 15 Inicio de PTS: El valor PTS del primer cuadro IDR que contiene las escenas excepcionalmente brillantes.

Fin de PTS: El valor PTS del último cuadro IDR que contiene las escenas excepcionalmente brillantes.

- 20 HDR\_DTS: Marca de tiempo que indica cuándo deben decodificarse los mensajes SEI.

Posición horizontal de la región: La posición horizontal de la región de los cuadros que son excepcionalmente brillantes.

- 25 Posición vertical de la región: La posición vertical de la región de los cuadros que son excepcionalmente brillantes

Ancho de la región: El ancho de la región

Altura de la región: La altura

- 30 Ganancia HDR: un código que define qué tan brillantes son los cuadros actuales, por ejemplo, en relación con un nivel de referencia que las pantallas pueden manejar sin sobrecalentamiento.

- 35 El siguiente ejemplo tiene el mensaje incorporado en la señalización a través de la interfaz de vídeo entre un aparato de entrega de contenido de vídeo y una pantalla. Los estándares actualmente de ejemplos para esto son HDMI y Displayport. La señalización en ambos estándares se basa en el estándar CEA 861-D. Esto define el contenido de un cuadro de información específico del proveedor, que consiste en una cantidad de bytes que se pueden transmitir durante los períodos de supresión vertical de la transmisión de vídeo.

- 40 Un ejemplo de HDR de bloque de datos específico del proveedor de HDMI puede tener este aspecto.

Byte del paquete #	7	6	5	4	3	2	1	0
PB0	No relevante						Longitud	
PB1	Identificador de registro IEEE de 24 bits							
PB2								
PB3								
PB4	No relevante							
PB5								
PB6	No relevante							
PB7								
PB8	No relevante			Impulso presente HDMI HDR				
PB9-15	No relevante							
PB16	Ganancia de HDR							
PB17	Impulso HDR de región LSB horizontal							
PB18	Impulso HDR de región MSB horizontal							
PB19	Impulso HDR de región LSB vertical							
PB20	Impulso HDR de región MSB vertical							
PB21	Impulso HDR de ancho de región LSB							
PB22	Impulso HDR de ancho de región MSB							
PB23	Impulso HDR de altura de región LSB							
PB24	Impulso HDR de altura de región MSB							
PB (longitud)								

Los componentes algorítmicos descritos en este texto pueden (en su totalidad o en parte) realizarse en la práctica como hardware (por ejemplo, partes de un IC específico de la aplicación) o como software que se ejecuta en un procesador de señal digital especial, o un procesador genérico, etc. Pueden ser semiautomáticos en el sentido de que al menos alguna entrada del usuario puede estar (por ejemplo, en la fábrica, o entrada del consumidor) presente.

La persona experta de nuestra presentación debe comprender qué componentes pueden ser mejoras opcionales y se pueden realizar en combinación con otros componentes, y cómo (opcional) los pasos de los métodos se corresponden con los medios respectivos de los aparatos, y viceversa. El hecho de que algunos componentes se divulguen en la invención en una cierta relación (por ejemplo, en una figura única en una determinada configuración) no significa que no sean posibles otras configuraciones como realizaciones bajo el mismo pensamiento inventivo que se describe para las patentes en este documento. Además, el hecho de que, por razones pragmáticas, solo se haya descrito un espectro limitado de ejemplos, no significa que otras variantes no puedan caer dentro del amplio alcance de las reivindicaciones. De hecho, los componentes de la invención pueden incorporarse en diferentes variantes a lo largo de cualquier cadena de uso, por ejemplo, todas las variantes del codificador pueden ser similares o corresponder a los decodificadores correspondientes y viceversa, y estar codificadas como datos de señal específicos en una señal para transmisión, o uso adicional tal como la coordinación, en cualquier tecnología de transmisión entre codificador y decodificador, etc. La palabra "aparato" en esta solicitud se usa en su sentido más amplio, es decir, un grupo de medios que permite la realización de un objetivo particular, y por lo tanto puede ser (una pequeña parte de) un IC, o un dispositivo dedicado (tal como un dispositivo con una pantalla), o parte de un sistema en red, etc. "Disposición" también debe utilizarse en el sentido más amplio, por lo que puede comprender, entre otros, un solo aparato, una parte de un aparato, una colección de (partes de) aparatos que cooperan, etc.

Se debe entender que la denotación del producto del programa de ordenador abarca cualquier realización física de una colección de comandos que habilita un procesador genérico o de propósito especial, después de una serie de pasos de carga (que pueden incluir pasos de conversión intermedios, tal como la traducción a un idioma intermedio), y un lenguaje de procesador final) para ingresar 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 portador tal como, por ejemplo, un disco o cinta, datos presentes en la memoria, datos que viajan a través de una conexión de red (por cable o inalámbrica) o código de programa en papel. Además del código del programa, los datos característicos requeridos para el programa también pueden incorporarse como un producto de programa de ordenador. Dichos datos pueden ser suministrados (parcialmente) de cualquier manera.

Algunos de los pasos requeridos para la operación del método ya pueden estar presentes en la funcionalidad del procesador o en cualquier realización del aparato de la invención en lugar de estar descritas en el producto de programa de ordenador o cualquier unidad, aparato o método descrito aquí (con detalles específicos de las realizaciones de la invención), tales como pasos de entrada y salida de datos, pasos de procesamiento típicamente incorporados bien conocidos, tal como manejo de pantalla estándar, etc. También deseamos protección para productos resultantes y resultantes similares, como por ejemplo las nuevas señales específicas involucradas en cualquier paso de los métodos o en cualquier subparte de los aparatos, así como cualquier nuevo uso de dichas señales, o cualquier método relacionado.

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

Cualquier signo de referencia entre paréntesis en la reivindicación no pretende limitar la reivindicación, ni ningún símbolo en particular en los dibujos. La palabra "que comprende" no excluye la presencia de elementos o aspectos que no figuran en una reclamación. La palabra "un" o "alguno" que precede a un elemento no excluye la presencia de una pluralidad de tales elementos.

## REIVINDICACIONES

1. Un método de codificación de datos de vídeo de alto rango dinámico (VID) y datos adicionales (DD) que comprende al menos un instante de cambio de tiempo (TMA\_1) que indica un cambio en el tiempo de una luminancia característica (CHRLUM) de los datos de vídeo, que luminancia característica resume un conjunto de luminancias de píxeles en una imagen de datos de vídeo, el método comprende:
  - generar sobre la base de los datos descriptivos (DED) de datos de vídeo (VID) de la variación de luminancia característica del vídeo, los datos descriptivos que comprenden al menos un instante de cambio de tiempo (TMA\_1), y
  - codificar y emitir de los datos de vídeo (VID) que comprenden luminancias (Lin\*) iniciales;
  - Codificar en datos adicionales (DD) al menos una indicación (ALCORR, (TIPO)) de estrategias de reprocesamiento permitidas de al menos las luminancias iniciales de los píxeles de los datos de vídeo por un aparato (112, 110) utilizando los datos de vídeo y datos adicionales, para obtener luminancias de alto rango dinámico, que al menos una indicación de las estrategias de reprocesamiento permitidas comprende un algoritmo matemático para calcular una estrategia de desviación para cambiar las luminancias iniciales (Lin\*), prescritas en instantes de tiempo sucesivos después del al menos un instante de cambio de tiempo (TMA\_1);
  - codificar y emitir los datos adicionales (DD), que comprenden los datos descriptivos (DED), para permitir que un aparato (600) de decodificación de vídeo se aplique a instantes de tiempo sucesivos después del al menos un instante de cambio de tiempo (TMA\_1) del algoritmo matemático a las luminancias iniciales para obtener las luminancias de alto rango dinámico.
2. Un método para codificar datos adicionales (DD) como se reivindica en la reivindicación 1, en donde el método comprende mapear luminancias para ser procesadas en píxeles entre un rango de luminancia dinámica más bajo y un rango de luminancia dinámica más alto.
3. Un método para codificar datos adicionales (DD) como se reivindica en la reivindicación 1, que comprende un paso de codificación de un código de reprocesamiento particular (MULT) a partir de un conjunto de códigos acordados prefijados.
4. Un método para codificar datos adicionales (DD) como se reivindica en cualquiera de las reivindicaciones anteriores, en donde el reprocesamiento se basa en un modelo psicovisual que modela la luminosidad percibida de las luminancias de píxeles renderizadas.
5. Un método para codificar datos adicionales (DD) como se reivindica en cualquiera de las reivindicaciones anteriores, en donde el reprocesamiento se basa en una característica física de la pantalla y/o del entorno de visualización.
6. Un método para codificar datos adicionales (DD) como se reivindica en cualquiera de las reivindicaciones 1 a 6, en donde el reprocesamiento es de un tipo que comprende determinar una imagen de iluminación para una retroiluminación (ILIM), y el paso de codificación comprende datos de codificación para influenciar la determinación de la imagen de iluminación para una retroiluminación (ILIM) durante un intervalo alrededor del instante de cambio de tiempo (TMA\_1), tal como una función temporal compuesta por contribuciones de función de base elemental para al menos una región espacial de posiciones de una matriz bidimensional (MAP).
7. Un método de codificación de datos adicionales (DD) como se reivindica en cualquiera de las reivindicaciones anteriores, que comprende un paso de codificación en la información de datos adicionales (DD) de luminancias características de un futuro del instante de cambio de tiempo (TMA\_1) y/o información de las luminancias esperadas de una imagen de iluminación para una retroiluminación (ILIM) de una pantalla de referencia.
8. Un método de codificación de datos adicionales (DD) como se reivindica en cualquiera de las reivindicaciones anteriores, que comprende un paso de codificación en los datos adicionales (DD) y una indicación de importancia (IMPLEV) durante al menos un instante de cambio (TMA\_1).
9. Un aparato (524) de codificación de vídeo dispuesto para codificar, además de datos de vídeo de alto rango dinámico (VID), datos adicionales (DD) que comprenden al menos un instante de cambio (TMA\_1) que indica un cambio en el tiempo de una luminancia característica (CHRLUM) de los datos de vídeo, cuya luminancia característica resume un conjunto de luminancias de píxeles en una imagen de los datos de vídeo, con el aparato (524) de codificación dispuesto para:
  - codificar y emitir los datos de vídeo (VID) que comprenden las luminancias iniciales (Lin\*);
  - codificar en datos adicionales (DD) al menos una indicación (ALCORR, (TIPO)) de estrategias de reprocesamiento permitidas de al menos las luminancias iniciales de los píxeles de los datos de vídeo por un

aparato (112, 110) utilizando los datos de vídeo y datos adicionales, para obtener luminancias de alto rango dinámico, que al menos una indicación de las estrategias de reprocesamiento permitidas comprende un algoritmo matemático para calcular una estrategia de desviación para cambiar las luminancias iniciales (Lin\*), prescritas en instantes de tiempo sucesivos después del al menos un instante de cambio de tiempo (TMA\_1);

- codificar y emitir los datos adicionales (DD) que comprenden los datos descriptivos (DED), para permitir que un aparato (600) de decodificación de vídeo solicite instantes de tiempo sucesivos después del al menos un instante de cambio (TMA\_1) del algoritmo matemático a las luminancias iniciales para obtener las luminancias de alto rango dinámico.

10. Un aparato (524) de codificación de vídeo como se reivindica en la reivindicación 9 dispuesto para codificar, además de datos de vídeo (VID), datos adicionales (DD) que comprenden un código de reprocesamiento particular (MULT) de un conjunto de códigos acordados prefijados.

11. Un método de decodificación de datos de vídeo (VID) de alto rango dinámico y datos adicionales (DD) que es adicional a los datos de vídeo (VID) que comprenden luminancias iniciales (Lin\*), los datos adicionales (DD) que comprenden al menos un instante de cambio de tiempo (TMA\_1) que indica un cambio en el tiempo de una luminancia característica (CHRLUM) de los datos de vídeo, cuya luminancia característica resume un conjunto de luminancias de píxeles en una imagen de los datos de vídeo, y los datos adicionales (DD) que comprenden al menos una indicación de las estrategias de reprocesamiento permitidas que comprenden un algoritmo matemático para calcular una estrategia de desviación para cambiar las luminancias iniciales (Lin\*), prescritas en instantes de tiempo sucesivos después de al menos un instante de cambio de tiempo (TMA\_1), el método comprende además emitir al menos un instante de cambio de tiempo (TMA\_1) y utilizarlo para transformar las luminancias de píxeles de las imágenes después del instante de tiempo aplicando instantes de tiempo sucesivos después del al menos un instante de cambio de tiempo (TMA\_1) el algoritmo matemático a las luminancias iniciales para obtener las luminancias de alto rango dinámico, y la decodificación y salida de los datos de vídeo (VID) de alto rango dinámico.

12. Un método para decodificar datos de vídeo (VID) de alto rango dinámico y datos adicionales (DD) adicionales a los datos de vídeo (VID) como se reivindica en la reivindicación 11, el método comprende además un paso de decodificación a partir de la información de datos adicionales (DD) de luminancias características de un futuro del instante de cambio de tiempo (TMA\_1).

13. Un aparato (600) de decodificación de vídeo dispuesto para decodificar datos de vídeo (VID) que comprende luminancias iniciales (Lin\*) y dispuesto para decodificar, relacionado con los datos de vídeo (VID) de los datos adicionales (DD) que comprende al menos un instante de cambio de tiempo (TMA\_1) que indica un cambio en el tiempo de una luminancia característica (CHRLUM) de los datos de vídeo, cuya luminancia característica resume un conjunto de luminancias de píxeles en una imagen de los datos de vídeo, y está dispuesta para emitirse a través de una salida (650) al menos un instante de cambio de tiempo (TMA\_1), y el aparato (600) de decodificación de vídeo está dispuesto para decodificar a partir de los datos adicionales (DD) al menos una indicación de estrategias de reprocesamiento permitidas que comprende un algoritmo matemático para calcular una estrategia de desviación para cambiar las luminancias iniciales (Lin\*), prescritas en instantes de tiempo sucesivos después del al menos un instante de cambio de tiempo (TMA\_1); y el aparato (600) de decodificación de vídeo está dispuesto para transformar las luminancias de píxeles iniciales de las imágenes después del instante de tiempo aplicando el algoritmo matemático para obtener luminancias de alto rango dinámico.

14. Un aparato (600) de decodificación de vídeo como se reivindica en la reivindicación 13, dispuesto además para decodificar un código de reprocesamiento particular (MULT) a partir de un conjunto de códigos acordados prefijados.

15. Un aparato (600) de decodificación de vídeo como el reivindicado en la reivindicación 13, dispuesto además para decodificar al menos una especificación para determinar una imagen de activación para una retroiluminación (ILIM).

16. Un aparato (600) de decodificación de vídeo como se reivindica en la reivindicación 13, dispuesto además para decodificar al menos una especificación (SPECFUT) que resume las luminancias características de un futuro del instante de cambio (TMA\_1) y/o información de las luminancias esperadas de una imagen de iluminación para una retroiluminación (ILIM) de una pantalla de referencia

17. Una disposición (110+100) que comprende un aparato (600) de decodificación de vídeo como en cualquiera de las reivindicaciones anteriores de decodificación de vídeo, y una pantalla (100) en la cual la pantalla está dispuesta a cambiar su renderización en base a al menos un instante de cambio de tiempo (TMA\_1).

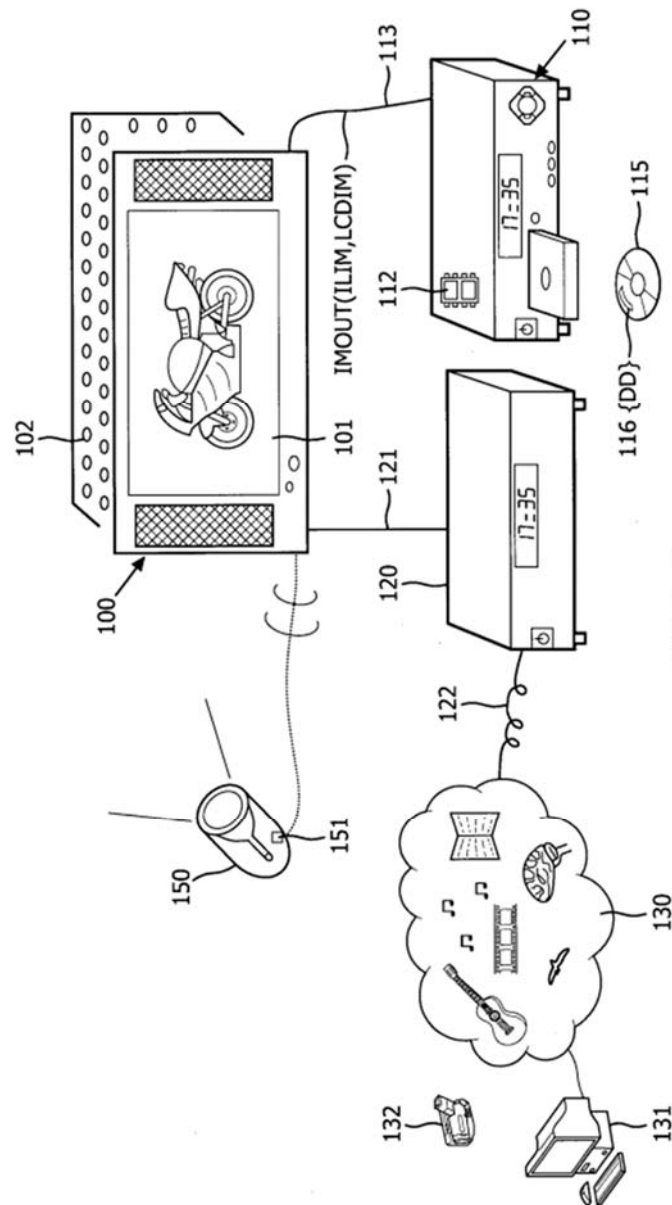


FIG. 1

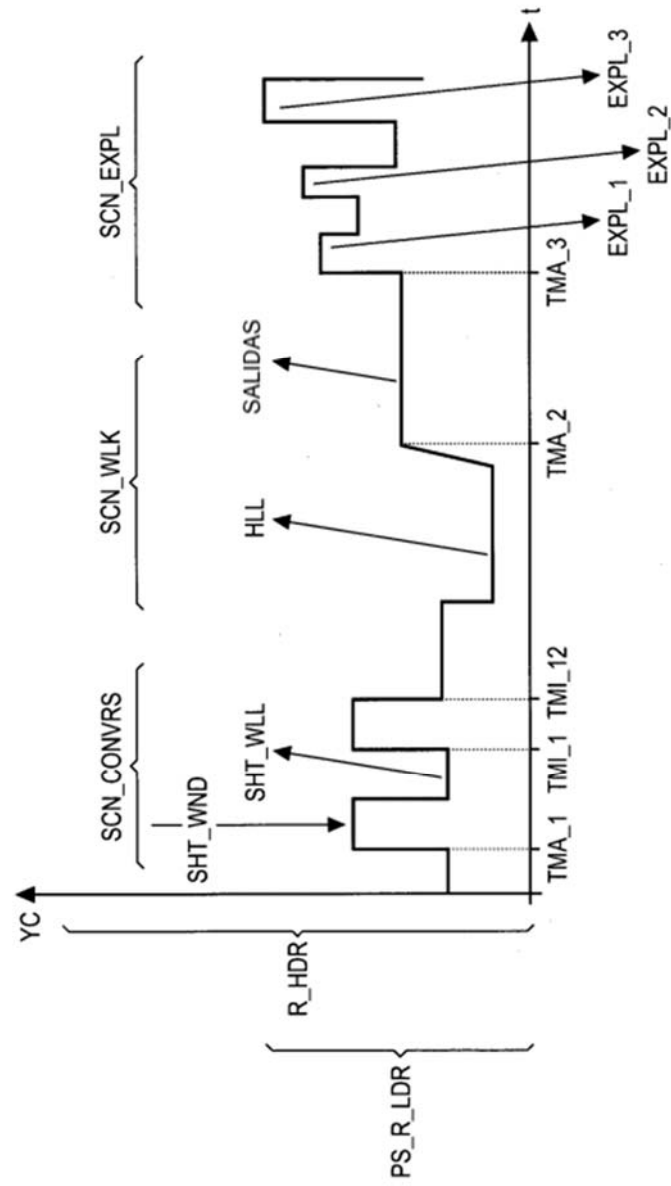


FIG. 2



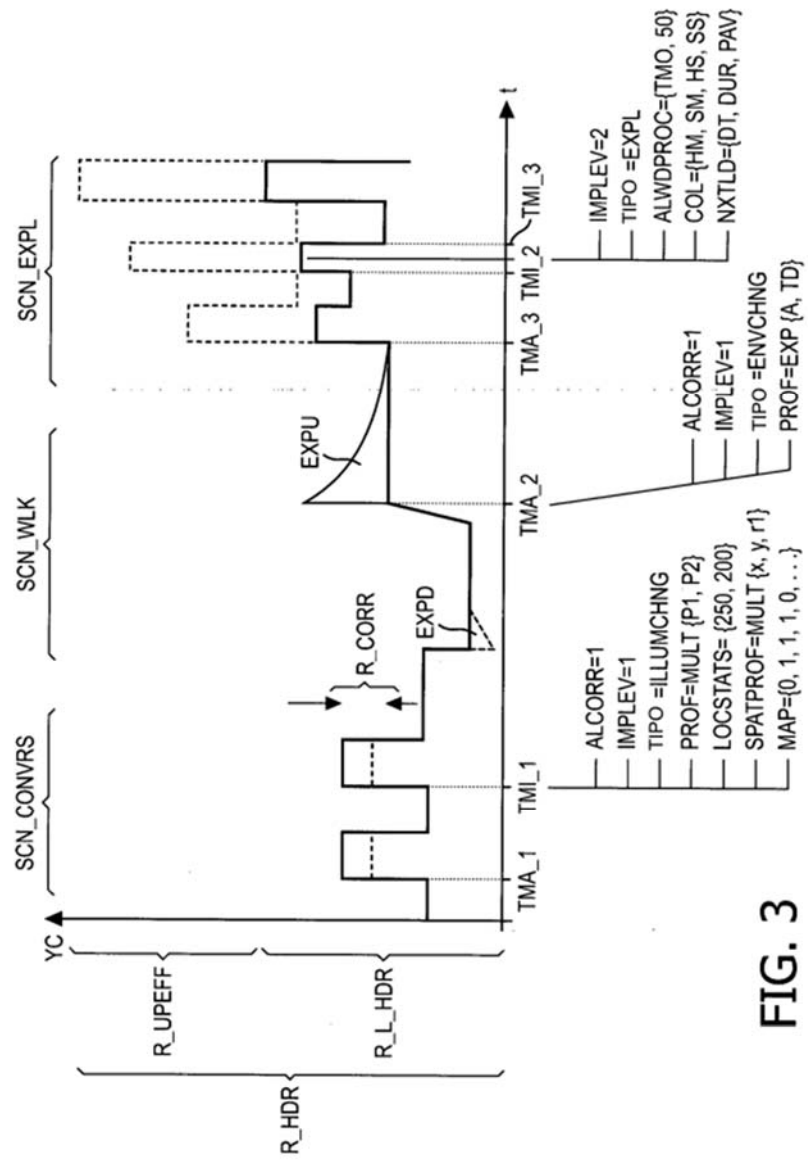


FIG. 3

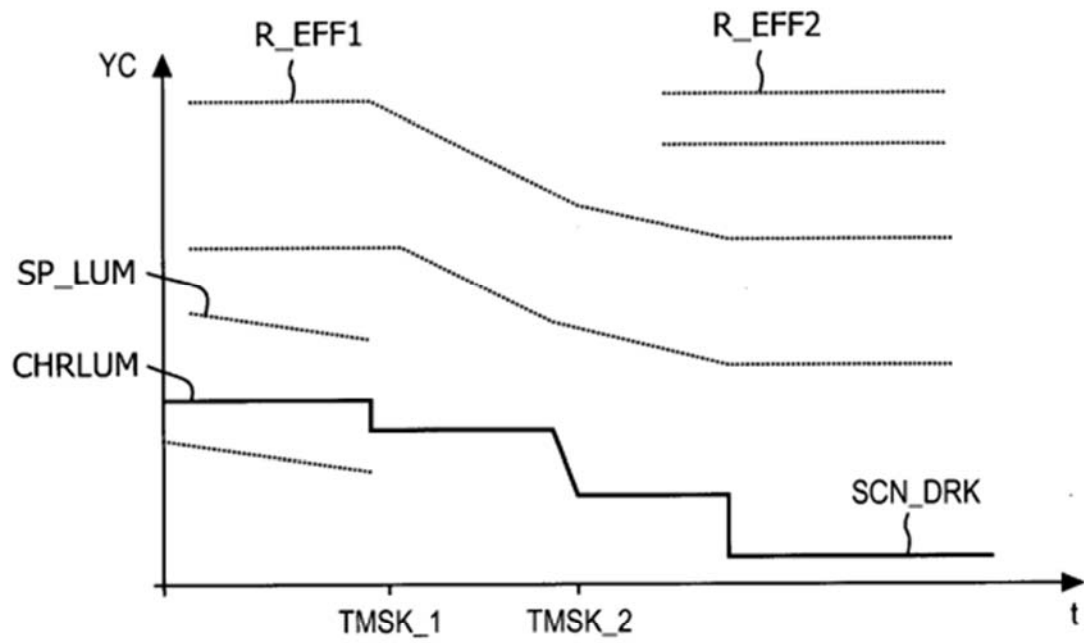


FIG. 4

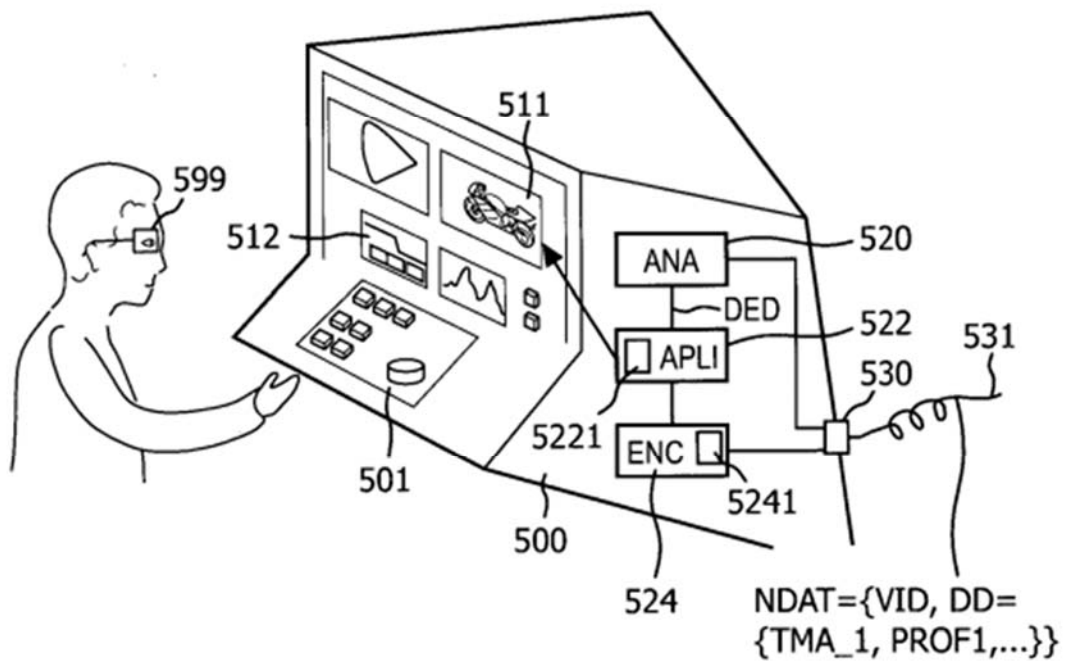


FIG. 5

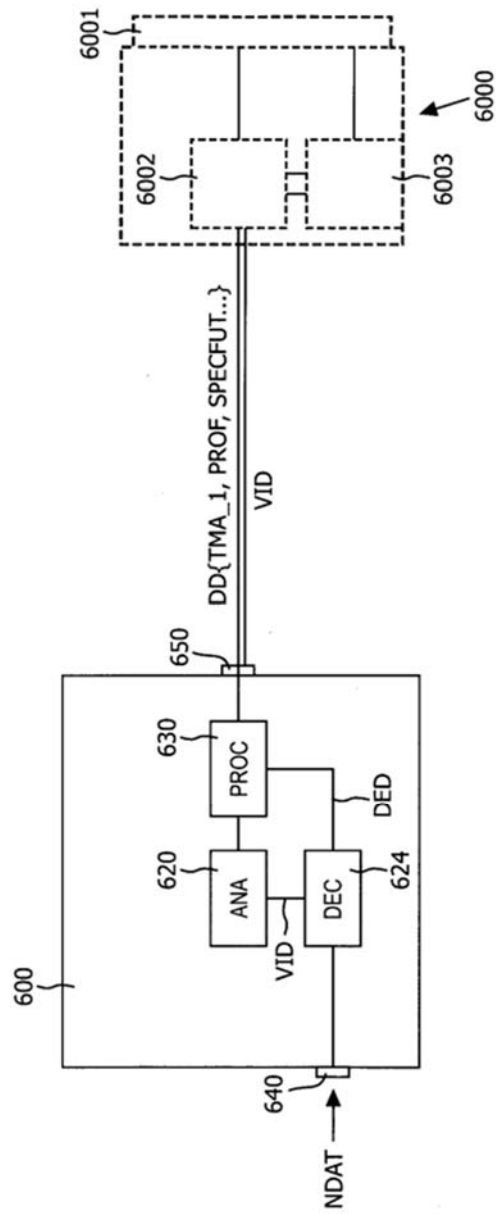


FIG. 6

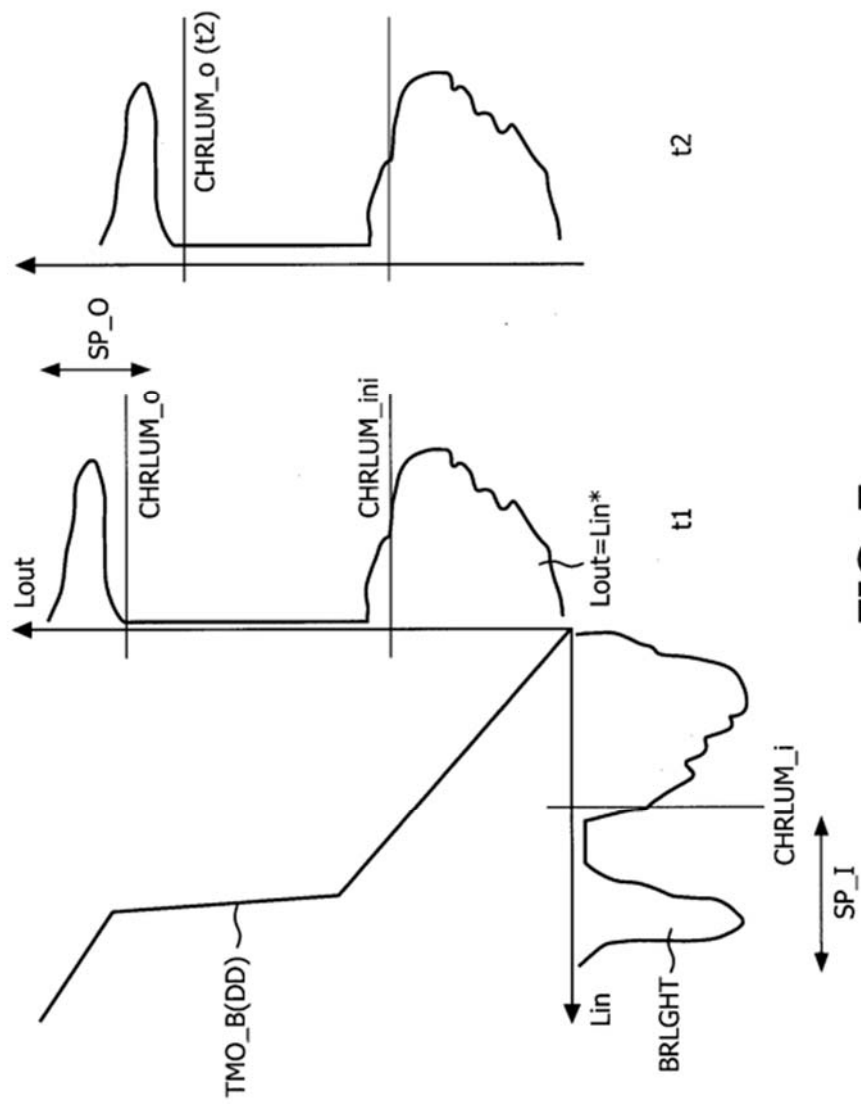


FIG. 7

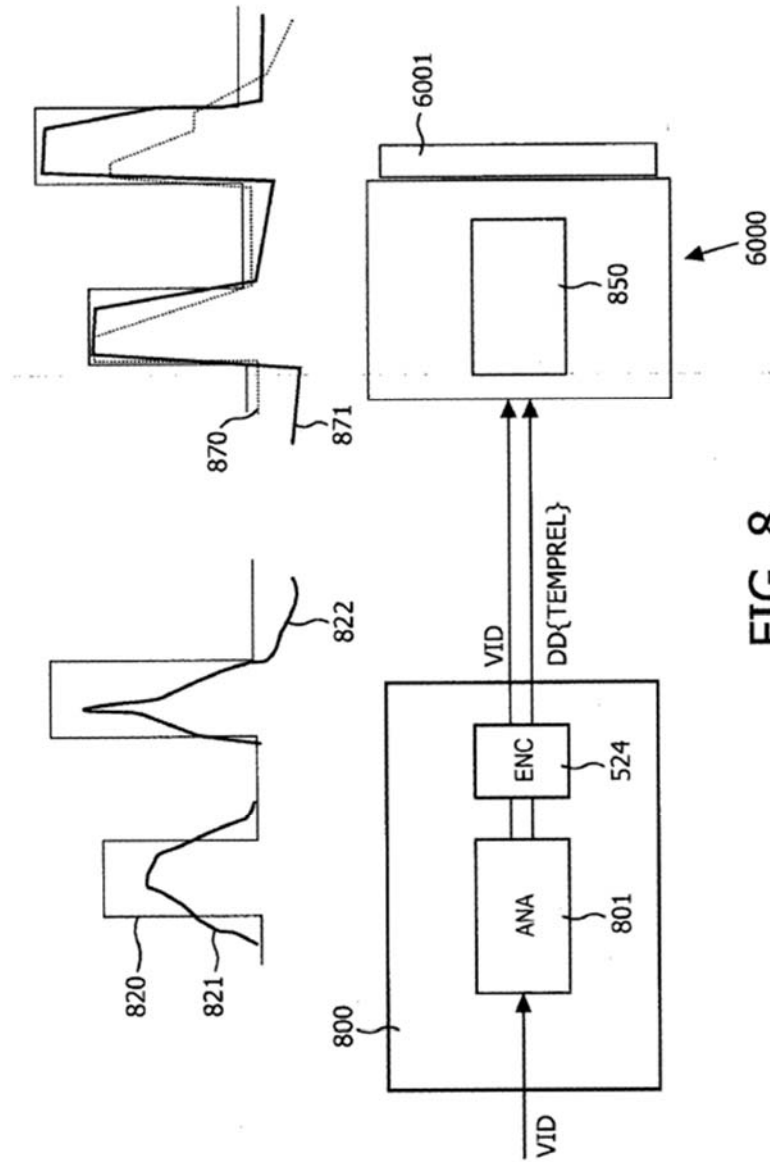


FIG. 8

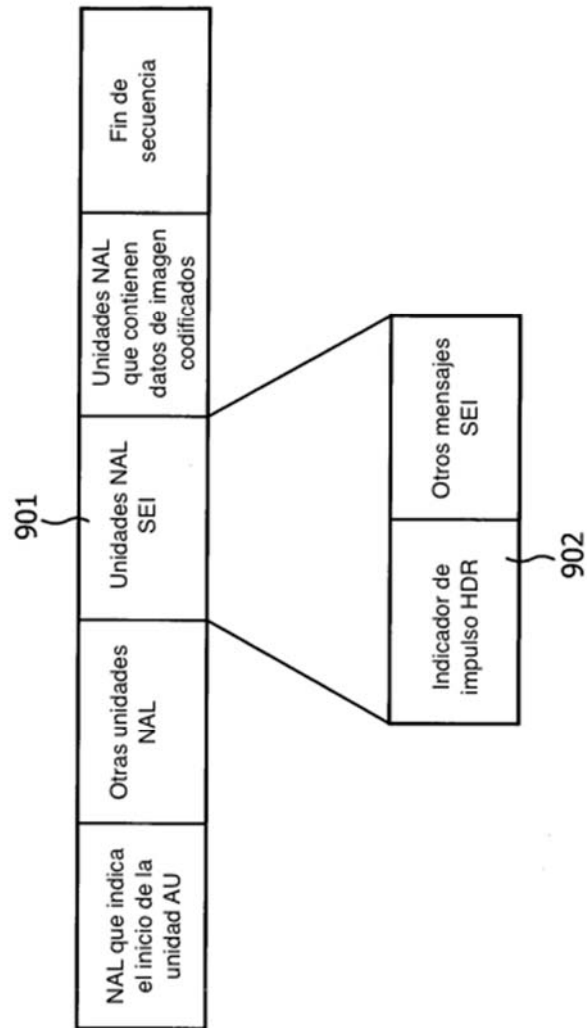


FIG. 9