

19



OFICINA ESPAÑOLA DE  
PATENTES Y MARCAS

ESPAÑA



11 Número de publicación: **2 667 497**

51 Int. Cl.:

**H04N 5/225** (2006.01)

**G06T 3/40** (2006.01)

**G06T 3/00** (2006.01)

12

TRADUCCIÓN DE PATENTE EUROPEA

T3

86 Fecha de presentación y número de la solicitud internacional: **13.07.2010 PCT/JP2010/004545**

87 Fecha y número de publicación internacional: **05.05.2011 WO11052114**

96 Fecha de presentación y número de la solicitud europea: **13.07.2010 E 10826254 (4)**

97 Fecha y número de publicación de la concesión europea: **21.03.2018 EP 2498486**

54 Título: **Programa, dispositivo y método de procesamiento de video, y dispositivo de imagen montado con dispositivo de procesamiento de video**

30 Prioridad:

**02.11.2009 JP 2009252148**

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

**11.05.2018**

73 Titular/es:

**SONY COMPUTER ENTERTAINMENT INC.  
(100.0%)  
2-6-21, Minami-Aoyama  
Minato-kuTokyo 107-0062, JP**

72 Inventor/es:

**OHBA, AKIO;  
SEGAWA, HIROYUKI y  
INADA, TETSUGO**

74 Agente/Representante:

**LEHMANN NOVO, María Isabel**

ES 2 667 497 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**

Programa, dispositivo y método de procesamiento de vídeo, y dispositivo de imagen montado con dispositivo de procesamiento de vídeo.

5

**CAMPO DE LA INVENCION**

La presente invención se refiere a una técnica para la transmisión, a un terminal concentrador, una imagen creada por un elemento de creación de imágenes que se proporciona en una cámara.

10

**ANTECEDENTES DE LA INVENCION**

En la técnica relacionada, se conocen juegos en los que la imagen de una parte de un cuerpo del usuario, tal como una cabeza, se captura por una videocámara, en donde se extrae, a continuación, una zona predeterminada como un ojo, una boca, una mano, o similar, y en donde la zona se sustituye con otra imagen para su visualización en una pantalla (a modo de ejemplo, documento de patente nº 1). También es conocida una interfaz de usuario que recibe, como una instrucción para poner en práctica una aplicación, el movimiento de una boca o mano que se captura por una videocámara.

15

Documento de patente nº 1, la Patente Europea nº 0999518, de Adelson E H et. al., titulada: "Métodos piramidal en procesamiento de imágenes", RCA Engineer, Cherry Hill, NJ, US, vol. 29, nº 6, 1 noviembre 1984, páginas 33-41, da a conocer métodos para la generación de una imagen piramidal. La imagen piramidal incluye un conjunto de imágenes de paso bajo o pasabanda, incluyendo cada una información de modelo correspondiente que se puede utilizar para el análisis de imágenes. En el documento de Rastislav Lukac, titulado: "Creación de imágenes de sensor único en cámaras digitales de consumo: una encuesta de recientes avances y direcciones futuras", Revista de Procesamiento de Imágenes en Tiempo Real, vol. 1, nº 1, 1 marzo 2006, páginas 45-52.

20

25

**SUMARIO DE LA INVENCION**

30 Problema a resolverse mediante la invención

30

En una técnica tal como la anteriormente descrita, es necesaria una imagen en alta resolución para extraer una zona predeterminada para una boca, o mano de un usuario, o similar. Sin embargo, a medida que aumenta el rendimiento de un elemento de imagen de una videocámara, la realización de un proceso de compresión conlleva más tiempo cuando la videocámara transmite una imagen a un concentrador, tal como un dispositivo de juegos o un ordenador personal después de la compresión. Por lo tanto, existe un problema cuando aumenta la latencia entre la captura de la imagen por la cámara y la generación de un vídeo por el concentrador. Surge un problema cuando el uso de la cámara como una interfaz de usuario trae consigo el hecho de que el aumento en la latencia reduce, drásticamente, la usabilidad. Tal como se describe, incluso cuando se aumenta el rendimiento de un elemento de imagen de una videocámara, se puede reducir el rendimiento total de un sistema.

35

40

Dentro de este contexto, un objetivo de la presente invención es dar a conocer una técnica de procesamiento de imágenes para reducir la latencia asociada con una transmisión de imágenes desde una cámara a un concentrador mientras se está utilizando un elemento de creación de imágenes de alto rendimiento.

45

Medios para resolver el problema

Una forma de realización de la presente invención da a conocer un programa de procesamiento de imágenes en movimiento según se estipula en la reivindicación 1.

50

En conformidad con esta forma de realización, solamente una parte específica, que es una parte de una imagen no procesada, y una imagen reducida especificada, se transmiten desde el dispositivo de procesamiento de imágenes en movimiento, al terminal concentrador. De este modo, mientras una parte específica se pueda procesar, libremente, en el terminal concentrador, se reduce la cantidad de datos transmitidos al terminal concentrador en comparación con la transmisión de una imagen completa no procesada. En consecuencia, se puede reducir la latencia asociada con la transmisión de imágenes mientras se mantiene la cantidad de libertad para procesar una imagen en el terminal concentrador.

55

Otra forma de realización de la presente invención da a conocer, además, un programa de procesamiento de imágenes en movimiento según se estipula en la reivindicación 5.

60

De conformidad con la forma de realización, se puede realizar un proceso de interpolación cromática en el terminal concentrador sobre una parte específica de una imagen bruta capturada por el elemento de creación de imágenes. Por lo tanto, se puede realizar un proceso de interpolación cromática de imagen de alta calidad utilizando un recurso informático del terminal concentrador, cuya capacidad de procesamiento es normalmente más alta a la de un dispositivo de captura de imágenes. Puesto que una imagen reducida se amplía para uso para una parte distinta de

65

la parte específica, la cantidad de datos recibidos desde el dispositivo de captura de imágenes se puede reducir. De este modo, se puede suprimir la latencia asociada con la comunicación de imagen.

5 Las formas de realización de la invención, en la forma de métodos, sistemas, programas informáticos, y medios de grabación que almacenan programas informáticos, se pueden poner en práctica, además, como modos operativos adicionales de la presente invención. Además, se pueden realizar formas de realización preferidas de conformidad con las reivindicaciones subordinadas.

#### 10 VENTAJA DE LA PRESENTE INVENCION

De conformidad con la presente invención, se puede obtener una imagen parcial de alta resultado utilizando una imagen bruta, para una parte específica, mientras se reduce la latencia asociada con la comunicación de imagen desde un dispositivo de captura de imágenes a un terminal concentrador.

#### 15 BREVE DESCRIPCION DE LOS DIBUJOS

La Figura 1 es un diagrama que ilustra la configuración completa de un sistema de cámara de baja latencia de conformidad con una forma de realización de la presente invención;

20 La Figura 2 es un diagrama que ilustra la configuración de una cámara de conformidad con una forma de realización;

La Figura 3 es un diagrama que ilustra la configuración de una unidad de transmisión de imágenes de la cámara ilustrada en la Figura 2;

25 La Figura 4 es un diagrama que ilustra una función de la unidad de transmisión de imágenes;

La Figura 5 es un diagrama que ilustra la configuración de una parte de un terminal concentrador que es responsable para la generación de una imagen de baja latencia;

30 La Figura 6 es un diagrama de flujo que ilustra la operación del sistema de cámara de baja latencia;

La Figura 7 es un diagrama de flujo que ilustra la operación del sistema de cámara de baja latencia; y

35 La Figura 8 es un diagrama que ilustra la descripción de un proceso de imagen cuando el sistema de cámara de baja latencia se pone en práctica para una aplicación de videoconferencia.

#### DESCRIPCION DETALLADA DE LA INVENCION

40 La Figura 1 ilustra la configuración completa de un sistema de cámara de baja latencia 10 de conformidad con una forma de realización de la presente invención.

100 y se pone en práctica un proceso de imagen sobre la imagen en movimiento en un terminal concentrador 20. A continuación, se ilustra una imagen del usuario en una pantalla 4 o se transmite a un destino de comunicación predeterminado a través de una red 12, tal como la red Internet o una red LAN (Red de Área Local).

45 La cámara 100 es una videocámara digital provista de un elemento de creación de imágenes tal como un CCD (Dispositivo Acoplado por Carga) o un CMOS (Semiconductor Complementario de Óxido Metálico) y se proporciona en la parte superior de una carcasa de la pantalla 4, según se ilustra en la Figura. La pantalla 4 es, a modo de ejemplo, una televisión de cristal líquido, una televisión de plasma, un monitor de PC, o similar. Normalmente, el usuario 6 se coloca delante de la pantalla 4, de modo que una imagen de la totalidad o una parte del cuerpo del usuario se capture por la cámara 100. Una imagen ilustrada en la pantalla 4 varía dependiendo de una aplicación que se ejecuta en el sistema 10. A modo de ejemplo, cuando el sistema 10 se utiliza como una interfaz de usuario (UI) que reconoce el movimiento o expresión del usuario 6, e interpreta el movimiento o la expresión como algún tipo de instrucción de operación, una imagen 8 ilustrada en la pantalla 4 es una imagen de una parte del usuario 6, tal como de un rostro o de una mano, o del cuerpo completo. Cuando el sistema 10 se utiliza para la videoconferencia, la imagen 8, mostrada en la pantalla 4, es una imagen de un rostro de un socio de chat, y una imagen del usuario 6 se ilustra en una pantalla del compañero de chat a través de la red 12.

60 A la luz de un modo de uso tal como los descritos anteriormente, la cámara 100 se proporciona más adecuadamente en la parte superior de la pantalla 4. Sin embargo, siempre que una imagen de la totalidad o una parte del cuerpo del usuario 6 pueda capturarse, la cámara 100 se puede colocar en una ubicación distinta de la proximidad de la pantalla 4, p.ej., en la proximidad del terminal concentrador 20 o alrededor del usuario. En lugar de proporcionar la cámara 100 como una estructura única, la cámara 100 puede estar incrustada en la carcasa de la pantalla 4, o de modo similar. En lugar de usar un elemento de creación de imagen en la cámara 100, una imagen analógica puede convertirse de analógica a digital (A/D) para su uso.

65

El terminal concentrador 20 es un terminal informático, tal como un ordenador personal o un dispositivo de juegos, que se proporciona con una función de procesamiento de imágenes. El terminal concentrador 20, toma, de forma secuencial, en orden cronológico, imágenes en movimiento obtenidas mediante la captura de imágenes de la cámara 100 del usuario 6 y puesta en práctica un proceso de imagen predeterminado. En el caso de una aplicación de videoconferencia, se transmite una imagen del usuario 6 en la que se ha realizado el proceso de imagen a un compañero de chat por intermedio de la red 12 en tiempo real. En el caso de una aplicación de interfaz de usuario, se realiza, además, un proceso de espejo, y la imagen se envía, a continuación, a la pantalla 4 en tiempo real. El proceso de espejo es un proceso de creación de una imagen en una manera invertida a la derecha e izquierda, que permite al usuario utilizar el sistema como si el usuario se estuviera mirando en un espejo. Además del proceso de imagen anterior, el terminal concentrador 20 es capaz, además, de poner en práctica varias aplicaciones y luego, reproducir la imagen sintetizada en la pantalla 4.

En una aplicación para un juego, conversación de chat, etc., que utiliza una cámara convencional, la cámara asume una función de un proceso de reconocimiento de imagen y un proceso de compresión de imagen. La duración de los procesos suele ser más larga debido a los deficientes recursos informáticos de la cámara en comparación con un terminal concentrador. Por lo tanto, el rendimiento en tiempo real se ve afectado, a menudo, a modo de ejemplo, en situaciones en las que toma tiempo reconocer el movimiento del usuario y cuando hay un retardo en la visualización de las imágenes que se muestran en la pantalla. Esta tendencia se vuelve más pronunciada debido a un aumento en el número de pixels de un elemento de creación de imágenes que se proporciona en la cámara.

De este modo, en la presente forma de realización, se preparan en la cámara una pluralidad de imágenes cuyas resoluciones varían en etapas, y solamente se transmite una parte de imagen con una resolución necesaria desde la cámara al terminal concentrador, en función del tipo de una aplicación que se ejecuta en un sistema, de modo que se ponga en práctica un proceso de alta calidad en el terminal concentrador que tiene suficientes recursos informáticos.

La Figura 2 ilustra la configuración de la cámara 100 de conformidad con la forma de realización. Estas configuraciones pueden ponerse en práctica, igualmente, mediante configuraciones tales como una CPU (Unidad Central de Procesamiento), una memoria RAM (Memoria de Acceso Aleatorio), una memoria ROM (Memoria de Solamente Lectura) y un circuito de presentación en términos de hardware, y por programas para proporcionar varias funciones tales como una entrada de datos, almacenamiento de datos, procesamiento de imágenes, y representación en términos de software. La Figura 2 ilustra bloques funcionales puestos en práctica mediante la cooperación de dichos componentes. Por lo tanto, estos bloques funcionales pueden realizarse en una diversidad de formas mediante combinaciones de hardware y software. Con el fin de facilitar la explicación, la Figura 2 incluye un diagrama esquemático que ilustra una parte de imagen procesada en cada bloque funcional.

La cámara 100 incluye una unidad de adquisición de imágenes 102, una unidad de interpolación cromática 104, una unidad de transmisión de imágenes 150, una unidad de filtro piramidal 170, y una unidad de comunicación 108. La unidad de adquisición de imágenes 102 efectúa la lectura, en un tiempo predeterminado (p.ej., 60 veces/segundo), una imagen expuesta por un dispositivo de creación de imágenes, tal como un sensor CCD o CMOS. En la siguiente explicación, se supone que esta imagen tiene un ancho de "h" pixels en una dirección horizontal. Esta imagen se denomina una imagen RAW (bruta). La unidad de adquisición de imágenes 102 transmite la imagen RAW a la unidad de interpolación cromática 104 y la unidad de transmisión de imágenes 150, cada vez que se complete la exposición para una fila horizontal única de la imagen RAW.

La unidad de interpolación cromática 104 tiene una memoria intermedia FIFO (Primero en Entrar, Primero en Salir) 105 que tiene una capacidad de h pixels y una unidad de procesamiento de interpolación cromática simple 106. La información de pixel de una fila horizontal única de la imagen RAW se introduce y memoriza en la memoria FIFO 105 hasta que se introducen pixels para una fila horizontal posterior en la unidad de interpolación cromática 104. A la recepción de los pixels para las dos filas horizontales, la unidad de procesamiento de interpolación cromática simple 106 pone en práctica, utilizando los pixels recibidos, un proceso de interpolación cromática de información de color complementario para cada pixel, sobre la base de los pixels que lo rodean, y luego creando una imagen a todo color. Como es conocido para un experto en esta técnica, existen numerosos métodos para este proceso de interpolación cromática. Sin embargo, en este caso es suficiente un proceso de interpolación cromática simple que utiliza solamente los pixels para dos filas horizontales. A modo de ejemplo, cuando un pixel para el que ha de calcularse un valor YCbCr correspondiente tiene solamente un valor G, se obtiene un valor RGB utilizando un valor R obtenido tomando una media de los valores R respectivos de pixels adyacentes a derecha e izquierda, el valor G, y un valor B de un pixel por encima o debajo del pixel para un valor R, un valor G, y un valor B del mismo, respectivamente, y luego, se sustituyen en una ecuación de conversión predeterminada con el fin de calcular el valor YCbCr. Puesto que dicho proceso de interpolación cromática es bien conocido, se omite aquí una explicación más detallada.

A continuación se describe el motivo por el que un proceso de interpolación cromática simple es suficiente. Para una parte en la que sea necesaria una imagen de alta calidad (la parte se refiere, en lo sucesivo, como una "parte específica"), el terminal concentrador 20 recibe una imagen RAW procedente de la cámara 100 y realiza un proceso. Por lo tanto, la calidad de una imagen no es importante para una parte distinta de la parte específica, o una parte distinta de la parte específica se utiliza solamente para detección de imagen, o similar. En consecuencia, la calidad

de una imagen después de la interpolación cromática no es de gran importancia. En una variación de un proceso de interpolación cromática simple, se puede utilizar un método para formar un valor YCbCr de un único pixel a partir de cuatro pixels de RGB. En este caso, se puede obtener una imagen interpolada de 1/4 del tamaño de la imagen RAW. Por lo tanto, no es necesario un primer filtro 110 descrito más adelante.

Tal como se ilustra en la figura, la unidad de procesamiento de interpolación cromática simple 106 convierte RGB de cuatro pixels, que son dos pixels horizontales y dos pixels verticales, en señales de color YCbCr, a modo de ejemplo. Un bloque constituido de estos cuatro pixels se transmite a una unidad de transmisión de imágenes 150 como una imagen interpolada de 1/1 y, además, se transmite a la unidad de filtro piramidal 170.

La unidad de filtro piramidal 170 tiene una función de jerarquización de una imagen dada en una pluralidad de resultados y de generación de la imagen jerarquizada. Un filtro piramidal está provisto de filtros de reducción de 1/4 sobre la base del número de niveles de resolución que, en general, son necesarios. En la forma de realización, el filtro piramidal tiene filtros de cuatro niveles: un primer filtro 110 hasta un cuarto filtro 140, inclusive. Cada filtro realiza un proceso de interpolación bilineal para pixels que son adyacentes entre sí y luego, calcula un valor medio de pixel de los cuatro pixels. En consecuencia, el tamaño de imagen obtenido después del proceso es 1/4 el tamaño de la imagen antes del proceso.

Antes del primer filtro 110, se proporciona una única memoria intermedia FIFO 112 para h pixels, en correspondencia para cada una de las señales Y, Cb y Cr. Estas memorias FIFO 112 tienen cada una de ellas, una función de almacenamiento de pixels YCbCr para una fila horizontal única hasta que se emiten pixels para una fila horizontal posterior desde la unidad de procesamiento de interpolación cromática simple 106. El tiempo de almacenamiento de los pixels se determina de conformidad con la velocidad de exploración de línea de un elemento de creación de imágenes. A la entrada de los pixels para dos filas horizontales, el primer filtro 110 realiza un promedio de los valores Y, Cb y Cr de pixel para cuatro pixels de dos pixels horizontales por dos pixels verticales. Mediante la repetición de este proceso, la longitud de una imagen objeto de interpolación 1/1 se hace de 1/2, tanto vertical como de forma horizontal, de modo que la imagen se convierta para tener un 1/4 del tamaño completo. La imagen objeto de interpolación de 1/4 convertida, se transmite a la unidad de transmisión de imágenes 150 y se transmite, además, al segundo filtro posterior 120.

Antes del segundo filtro 120, se proporciona una unidad memoria FIFO 122 para 2/h pixels, en correspondencia para cada una de las señales y, Cb y Cr. Estas memorias FIFO 114 tienen, cada una de ellas, una función de almacenamiento de pixels YCbCr para una fila horizontal única, hasta que se emiten pixels para una fila horizontal posterior, desde el primer filtro 110. Después de la entrada de los pixels para dos filas horizontales, el segundo filtro 120 realiza un promedio de los valores Y, Cb y Cr de pixel para cuatro pixels, de entre dos pixels horizontales por dos pixels verticales. Repitiendo este proceso, la longitud de la imagen objeto de interpolación de 1/4 se convierte en 1/2 tanto vertical como de forma horizontal, con lo que la imagen se convierte en 1/16 del tamaño completo. La imagen objeto de interpolación de 1/16 convertida se transmite a la unidad de transmisión de imágenes 150 y se transmite, además, al tercer filtro posterior 130.

Delante de cada uno de entre el tercer filtro 130 y el cuarto filtro 140, se proporcionan memorias intermedias FIFO 132 para h/4 pixels y memorias FIFO 142 para h/8 pixels, respectivamente. A diferencia de lo que antecede, se repite el mismo proceso que los descritos con anterioridad en el tercer filtro 130 y el cuarto filtro 140. Las imágenes objeto de interpolación de tamaño 1/64 y de tamaño 1/256, se transmiten, a continuación, a la unidad de transmisión de imágenes 150. Puesto que un filtro piramidal tal como los ilustrados anteriormente, es bien conocido, tal como se describe en el documento de patente nº 1, se omite una explicación más detallada en esta descripción.

Tal como se describe, las salidas de imágenes reducidas para ser de un tamaño de 1/4, una por una se introducen en la unidad de transmisión de imágenes 150 desde los filtros respectivos de la unidad de filtro piramidal 170. Como es evidente a partir de lo que antecede, el tamaño de una memoria FIFO, que es necesario antes de cada filtro, se reduce a medida que el número de filtros pasados se hace superior en la unidad de filtro piramidal 170.

De conformidad con una instrucción recibida desde el terminal concentrador 20, a través de la unidad de comunicación 108, la unidad de transmisión de imágenes 150 selecciona una imagen necesaria de entre la imagen RAW, recibida desde la unidad de adquisición de imágenes 102, la imagen interpolada de 1/1 recibida desde la unidad de interpolación cromática 104, y las imágenes interpoladas de 1/4 a 1/256, recibidas desde la unidad de filtro piramidal 170. La unidad de transmisión de imágenes 150 crea un paquete utilizando estas imágenes y luego, transmite el paquete a la unidad de comunicación 108.

La unidad de comunicación 108 transmite el paquete al terminal concentrador 20 de conformidad con, a modo de ejemplo, un protocolo predeterminado tal como USB 1.0/2.0, o similar. La comunicación con el terminal concentrador 20 no está limitada a una comunicación cableada. A modo de ejemplo, la comunicación puede ser una comunicación de red LAN inalámbrica, tal como IEEE 802.11a/b/g, o similar, o una comunicación por infrarrojos tal como IrDA, o similar.

La Figura 3 es un diagrama que ilustra la configuración de la unidad de transmisión de imágenes 150 de la cámara

100. La unidad de transmisión de imágenes 150 incluye una unidad de escritura de bloques 152, una memoria intermedia 154, una unidad de lectura de bloques 156, una unidad de codificación 158, una unidad de empaquetado 160, una memoria intermedia de paquetes 162 y una unidad de control 164.

5 Sobre la base de la instrucción procedente del terminal concentrador 20, la unidad de control 164 indica a la unidad de escritura de bloques 152 y a la unidad de lectura de bloques 156, los datos de imagen que han de transmitirse como un paquete entre varios conjuntos de datos de imagen. Tal como se describirá más adelante, solamente una parte de la imagen RAW, y las imágenes decodificadas que se han introducido en la unidad de transmisión de imágenes 150, son objeto de transmisión al terminal concentrador en la forma de realización.

10 Una imagen RAW se introduce en la unidad de escritura de bloques 152, procedente de la unidad de adquisición de imágenes 102, e imágenes de tamaños de 1/1 a 1/256 son objeto de entrada a través de la unidad de interpolación cromática 104 y la unidad de filtro piramidal 170. Una unidad de selección de tamaño 152A, de la unidad de escritura de bloques 152, escribe una parte de las imágenes interpoladas en la memoria intermedia 154, sobre la base de una instrucción procedente de la unidad de control 164. La unidad de escritura de bloques 152 recibe una imagen en una unidad de dos por dos pixels, tal como se ilustra en la Figura 2 y es operativa para realizar la escritura, de forma secuencial, de la imagen en la memoria intermedia 154. Una unidad 152B, para la selección de un bloque que ha de cortarse de la unidad de escritura de bloques 152 realiza la escritura, con respecto a una imagen RAW, solamente un bloque que contiene una parte específica, que se indica por la unidad de control 164, en la memoria intermedia 154. Esta parte específica es, a modo de ejemplo, una parte para la que es necesaria una imagen de alta calidad y una alta resolución de imagen, tal como una parte para un rostro de usuario, mano, o similar. Más adelante se proporciona una descripción adicional sobre la unidad 152B para la selección de un bloque que ha de cortarse.

25 La unidad de lectura de bloques 156 efectúa la lectura de cada bloque de imagen y luego, transmite el bloque de imagen a la unidad de codificación 158, en el orden de pixels para un bloque que son preparados en la memoria intermedia 154. La unidad de escritura de bloques 152 y la unidad de lectura de bloques 156 se ajustan para funcionar de forma síncrona con la unidad de control 164. Dicho de otro modo, mientras se realiza la escritura de la unidad de escritura de bloques 152, cada vez que se emiten pixels desde la unidad de adquisición de imágenes 102, la unidad de interpolación cromática 104, y la unidad de filtro piramidal 170, la unidad de lectura de bloques 156 efectúa la lectura cada vez que se acumulan pixels para un bloque en la memoria intermedia 154. La temporización de sincronización se determina de conformidad con la velocidad de exposición del elemento de creación de imágenes. Es preferible un bloque único a un bloque que tiene un tamaño de ocho por ocho pixels, de conformidad con la codificación JPEG que sigue a continuación. En la explicación siguiente, un bloque de la imagen RAW se refiere como Br, y bloques de las imágenes interpoladas de 1/4, 1/16, 1/64 y 1/256 se refieren como B1, B4, B16, B64 y B256, respectivamente.

40 En la forma de realización, en lugar de transmitir pixels al terminal concentrador cuando se agotan los pixels para la imagen RAW completa o las imágenes reducidas completas, se transmiten pixels en unidades de bloques. De este modo, el tamaño de la memoria intermedia 154 es lo suficientemente grande para memorizar todos los bloques de imagen para la imagen RAW y las imágenes reducidas. Dependiendo del tipo de una aplicación, solamente es necesaria una capacidad de almacenar dos a tres bloques de imágenes. Tal como se describe, se reducen los datos que han de memorizarse, y los datos se empaquetan, de forma secuencial, cada vez que se prepara un bloque y, a continuación, se transfiere. En consecuencia, se reduce la latencia asociada con un proceso puesto en práctica dentro de la cámara. Los pixels se emiten, de forma secuencial, desde la unidad de adquisición de imágenes 102 y la unidad de filtro piramidal 170 a la unidad de escritura de bloques 152, cada vez que se completa la exposición por el elemento de creación de imágenes. De este modo, es imposible, desde el aspecto estructural, que un bloque de una trama distinta sea objeto de escritura en la memoria intermedia 154, o un bloque que ha de ser empaquetado y luego transmitido, en un orden diferente.

50 La unidad 152B para la selección de un bloque que ha de cortarse, recibe información de posición de una parte específica, transmitida desde el terminal concentrador 20, y selecciona como un bloque específico un bloque que incluye una zona que es mayor que la zona que representa la parte específica por un número predeterminado de pixels.

55 La unidad de codificación 158 realiza una codificación de compresión bien conocida, tal como JPEG o similar, en un bloque de imagen de las imágenes reducidas, distintas de la imagen RAW y luego, transmite el bloque de imagen codificada por compresión a la unidad de empaquetado 160. La unidad de empaquetado 160 empaqueta un bloque de imagen RAW y un bloque de imagen codificada, de una imagen reducida en el orden de llegada a la unidad de empaquetado 160, y escribe los bloques de imagen empaquetados en la memoria intermedia de paquetes 162. La unidad de comunicación 108 transmite un paquete en la memoria intermedia de paquetes 162 para el terminal concentrador 20 de conformidad con un protocolo de comunicación predeterminado. Se pueden utilizar también otras tecnologías de codificación bien conocidas, tales como LLVC, AVC, etc. Sin embargo, son preferidas aquellas que pueden realizar la codificación en unidades de bloques. El tamaño de un bloque leído por la unidad de lectura de bloques 156 puede cambiarse de conformidad con la codificación. A modo de ejemplo, la lectura y la codificación se pueden poner en práctica utilizando un bloque en unidades de 256 por 256 pixels.

A continuación, se proporciona una explicación de la acción de la unidad de transmisión de imágenes 150, utilizando la Figura 4. Una columna derecha S1 ilustra una imagen RAW, e imágenes interpoladas recibidas por la unidad de escritura de bloques. Un pequeño cuadrado representa un pixel. Hay que tener en cuenta que un cuadrado corresponde a un valor de pixel de cualquiera de entre R, G y B en una imagen RAW, y que un solo pixel contiene todas las señales YCbCr en una imagen interpolada. La densidad de un cuadrado muestra que los pixels se han reducido por el filtro piramidal. De conformidad con una instrucción de la unidad de control 164, la unidad de escritura de bloques escribe solamente algunas imágenes, entre estas imágenes que se han proporcionado, en la memoria intermedia. En esta realización ejemplo, se supone que se selecciona una imagen RAW e imágenes interpoladas de 1/16 y 1/64. Para la imagen RAW, solamente se selecciona un bloque que contiene una parte específica. En consecuencia, una unidad de escritura de bloques, ilustrada en la columna central S2 en la Figura, efectúa la lectura de cuatro bloques Br de imagen RAW, un bloque B16 de imagen interpolada de 1/16, y un bloque B64 de imagen interpolada de 1/64, a partir de la memoria intermedia. Conviene señalar que, en realidad, los bloques no pasan directamente a la unidad de escritura de bloques desde la unidad de escritura de bloques, y que la unidad de escritura de bloques efectúa la lectura de los bloques en unidades de bloques, después de que los bloques de un tamaño predeterminado se hayan acumulado en la memoria intermedia.

Con respecto a los bloques que han sido objeto de lectura, los bloques de imagen interpolada B16 y B64, pasan por un proceso de codificación y a continuación, se empaquetan de nuevo junto con los bloques de imagen RAW, según se ilustra en una columna izquierda S3 en la figura.

La Figura 5 ilustra la configuración de una parte de un terminal concentrador que es responsable para la generación de una imagen de baja latencia en la forma de realización. Estas configuraciones se pueden poner en práctica, además, mediante configuraciones tales como una unidad CPU, una memoria RAM, una memoria ROM, una unidad GPU y un aparato de control de entrada y salida, en términos de hardware, y mediante programas informáticos para proporcionar varias funciones tales como entrada de datos, memorización de datos, procesamiento de imagen y presentación en términos de software. La Figura 6 ilustra bloques funcionales puestos en práctica mediante la cooperación de estos componentes. En consecuencia, estos bloques funcionales se pueden poner en práctica en una diversidad de formas mediante la combinación de hardware y software.

El terminal concentrador 20 incluye una unidad de comunicación 22, una unidad de control 52, una unidad de procesamiento de imagen 50, y una unidad de control de presentación visual 54. Además de permitir que un sistema operativo realice el control del funcionamiento general de la unidad de procesamiento de imagen 50, la unidad de control 52 pone en práctica otro control necesario para el terminal concentrador 20, p.ej., realización de varias aplicaciones para un juego, conversación de chat, etc., control de driver, control de la lectura de un programa procedente de un soporte de registro o similar. La unidad de comunicación 22 recibe varios conjuntos de datos de imagen procedentes de la cámara 100 y transmite los conjuntos de datos a la unidad de procesamiento de imagen 50.

La unidad de procesamiento de imagen 50 incluye una unidad de asignación 24, una unidad de interpolación cromática de imagen de alta calidad 28, una unidad de procesamiento de alta calidad 32, una unidad de decodificación 34, una unidad de ampliación 36, una unidad de síntesis de imagen 38, una unidad de salida de imagen 40, y una unidad de detección de parte específica 42. Bloques de una parte específica de una imagen RAW e imágenes reducidas codificadas, procedentes de la cámara 100, se introducen en la unidad de asignación 24. A continuación, de conformidad con una instrucción procedente de la unidad de control 52, la unidad de asignación 24 transmite un bloque de imagen de una imagen RAW a la unidad de interpolación cromática de imagen de alta calidad 28, y otros bloques de imagen a la unidad de decodificación 34.

La unidad de interpolación cromática de imagen de alta calidad 28 pone en práctica un proceso de interpolación sobre la parte específica de la imagen RAW. En este proceso de interpolación, que es distinto de la unidad de procesamiento de interpolación cromática simple de la cámara 100, la unidad de interpolación cromática de imagen de alta calidad 28 realiza un proceso de interpolación cromática de imagen de alta calidad utilizando un recurso informático del terminal concentrador 20. Se puede utilizar un algoritmo arbitrario que ya existe, o que se desarrollará en el futuro mediante, p.ej., el uso de un valor RGB de tres por tres pixels o más, para el cálculo de una señal de imagen YCbCr de un pixel, la modificación de un coeficiente de interpolación en consideración de un grado de correlación entre pixels en una dirección horizontal y/o una dirección vertical, o similar. Dicho proceso de interpolación cromática se describe en, a modo de ejemplo, la Publicación de Solicitud de Patente Japonesa nº 7-236147. Una imagen interpolada se transmite a la unidad de procesamiento de alta calidad 32.

La unidad de procesamiento de alta calidad 32 convierte, además, la imagen interpolada en una imagen de alta precisión adecuada para su visualización. Un proceso puesto en práctica en esta etapa varía en conformidad con el tipo de una aplicación. A modo de ejemplo, cuando se ilustra un rostro en una pantalla grande para una conversación de videoconferencia, se realizan procesos adecuados tales como un ajuste del tono fácil, modificación del color facial, modificación de los ojos y la boca, o similares. Se transmite una imagen de alta calidad a la unidad de síntesis de imagen 38.

Por otro lado, la unidad de decodificación 34 recibe un bloque de imagen de una imagen reducida distinta de la

imagen RAW y luego, decodifica la imagen comprimida mediante JPEG, o similar. Si no se realiza la compresión de la imagen en la cámara 100, la unidad de decodificación 34 no es necesaria.

5 La unidad de detección de parte específica 42 recibe la imagen reducida decodificada por la unidad de decodificación 34. A continuación, de conformidad con un método de detección de imagen bien conocido, la unidad de detección de parte específica 42 identifica una parte considerada para ser particularmente importante en una imagen del usuario para la aplicación. Esta parte es, a modo de ejemplo, un rostro del usuario en el caso de conversación por videoconferencia, y un ojo, una boca, una mano, una pierna, o similar en el caso de una UI. La información de posición de la parte así especificada se retroinforma a la cámara 100 por intermedio de la unidad de comunicación 22. De conformidad con la información de posición, bloques de imagen RAW alrededor de la parte específica, se transmiten desde la cámara 100. De este modo, se realiza el proceso de interpolación cromática en solamente bloques limitados por la unidad de interpolación cromática de imagen de alta calidad.

15 La detección de una parte específica se realiza, a modo de ejemplo, mediante un método que se muestra a continuación. En el caso de detección de rostros, una imagen se divide en 20 por 20 bloques, utilizando solamente una señal de brillo (Y), y una parte que comprende una pluralidad de bloques en los que la luminosidad es relativamente más brillante se reconoce como un rostro. Como alternativa, se puede memorizar una imagen de referencia de una parte específica (un rostro, una mano, etc.) en una unidad de memorización de imagen de referencia (no ilustrada), por anticipado, y una unidad de coincidencia (no ilustrada) puede hacer coincidir la imagen de referencia y una imagen de entrada, con el fin de identificar una zona correspondiente a un rostro o mano del usuario. Después de la identificación de una zona para el rostro, se pueden identificar zonas correspondientes a un ojo, una nariz, una boca, etc., en la zona para el rostro mediante la coincidencia con imágenes de referencia para las respectivas partes.

25 La unidad de ampliación 36 recibe una imagen reducida (p.ej., 1/64) que ha sido decodificada, y amplía la imagen a un tamaño de 1/1 (mismo tamaño). Los datos de imagen ampliada se transmiten a la unidad de síntesis de imagen 38.

30 La unidad de síntesis de imagen 38 sintetiza, a la vez, una imagen de baja calidad de un tamaño de 1/1 y un bloque de imagen de alta calidad de una parte específica. Como resultado, se crea una imagen en donde una parte específica, que es apreciada en una aplicación, tiene una calidad de imagen extremadamente alta y una alta resolución, mientras que otras partes tienen una baja calidad y baja resolución. La unidad de salida de imagen 40 efectúa la escritura de una imagen sintetizada creada en una memoria intermedia de trama (no ilustrada). La unidad de control de presentación visual 54 y la unidad de salida de imagen 40 crean una señal de vídeo para la visualización, en la pantalla 4, de datos de imagen dibujados en la trama de memoria intermedia.

35 A continuación, se proporciona una explicación de la acción del terminal concentrador 20. La unidad de asignación 24, que ha recibido varios bloques de imagen procedentes de la cámara 100, envía un bloque de imagen RAW a la unidad de interpolación cromática de imagen de alta calidad 28, y otros bloques de imagen a la unidad de decodificación 34. Cuando se recogen bloques de imagen RAW correspondientes a una parte específica, la unidad de interpolación cromática de imagen de alta calidad 28 pone en práctica un proceso de interpolación cromática de imagen de alta calidad con el fin de convertir los bloques de imagen RAW en una imagen de protocolo de YCbCr. A continuación, se obtiene una imagen de parte específica de alta calidad, realizando, además, un proceso predeterminado por la unidad de procesamiento de alta calidad 32. Por otro lado, un bloque de una imagen reducida pasa por un proceso de decodificación en la unidad de decodificación 34 y luego, se amplía a un tamaño de 1/1 mediante la unidad de ampliación 36. Entonces, la unidad de síntesis de imagen 38 proporciona, a la salida, una imagen sintetizada en la que una parte específica de la imagen reducida, que ha sido ampliada, se sustituye con una imagen de parte específica de alta calidad. Mediante el uso de una señal Y de la imagen reducida después de la decodificación, la unidad de detección de parte específica 42 detecta la parte específica. Esta información se transmite a la unidad de control de la cámara y se utiliza para la selección de un bloque de imagen RAW en la unidad de escritura de bloques 152. Se repite una serie de procesos, anteriormente descritos, sobre la base de trama por trama. En consecuencia, se crea una imagen en movimiento del usuario que tiene una parte específica de una alta calidad.

55 Las Figuras 6 y 7 son diagramas de flujo que ilustran la operación completa del sistema de cámara de baja latencia 10 de conformidad con la forma de realización. En primer lugar, se proporciona una explicación del funcionamiento de la cámara. Cuando la unidad de adquisición de imágenes 102 recibe una imagen procedente del elemento de creación de imágenes, se transmite una imagen RAW a la unidad de transmisión de imágenes 150 y a la unidad de interpolación cromática 104 (S10). La unidad de interpolación cromática 104 pone en práctica un proceso de interpolación cromática de calidad de imagen relativamente baja sobre un pixel de la imagen RAW, y transmiten el pixel interpolado a la unidad de transmisión de imágenes 150 y la unidad de filtro piramidal 170 (S12). Se realiza una interpolación bilineal por cada uno de los filtros de las capas respectivas de la unidad de filtro piramidal 170, y se proporcionan, a la salida, filas de pixels de tamaños de 1/4 a 1/256, hacia la unidad de transmisión de imágenes 150 (S14).

65 De conformidad con una instrucción procedente de la unidad de control 164, la unidad de escritura de bloques 152

de la unidad de transmisión de imágenes 150, escribe una parte de una imagen reducida en la memoria intermedia 154 (S16) y, para una imagen RAW, selecciona solamente un bloque de imagen que contiene una parte específica, que ha de escribirse en el bloque de imagen en la memoria intermedia 154 (S18). Cada vez que se registran pixels en, p.ej., ocho por ocho bloques en la memoria intermedia 154, la unidad de lectura de bloques 156 efectúa la lectura de dichos bloques de imagen y luego, transmite los bloques de imagen a la unidad de codificación 158 (S20). Después de pasar por una codificación predeterminada en la unidad de codificación 158 (S22), los bloques de imagen son empaquetados y a continuación, transmitidos al terminal concentrador 20 (S24).

Ahora se proporciona una explicación del funcionamiento del terminal concentrador ilustrado en la Figura 7. La unidad de asignación 24 toma varios bloques de imagen procedentes de un paquete recibido desde la cámara, y transmite los bloques de imagen a la unidad de interpolación cromática de imagen de alta calidad 28, o la unidad de decodificación 34 (S30). Después de pasar a través de un proceso de interpolación cromática puesto en práctica en la unidad de interpolación cromática de imagen de alta calidad 28 (S32), un bloque de imagen RAW pasa por un proceso predeterminado en la unidad de procesamiento de alta calidad 32 y luego, se proporciona a la unidad de síntesis de imagen 38 (S34). Por otro lado, se decodifica un bloque de imagen reducida en la unidad de decodificación 34 (S36), y se detecta, a continuación, una parte específica por la unidad de detección de parte específica 42, utilizando una señal de brillo (S38). Esta información se transmite a la cámara 100 y se utiliza para seleccionar un bloque específico de la imagen RAW (S40). La imagen reducida después de la decodificación es objeto de ampliación a un tamaño de 1/1 por la unidad de ampliación 36, y se proporciona, a la salida, a la unidad de síntesis de imagen 38 (S42). La unidad de síntesis de imagen 38 sintetiza una imagen de alta calidad de la parte específica y otras imágenes ampliadas (S44) y la unidad de salida de imagen 40 proporciona, a la salida, una imagen sintetizada a la memoria intermedia de trama (S46).

La forma de realización está basada en el entendimiento de que una parte específica de una imagen de usuario tal como un rostro, una mano, etc., mostrada en una pantalla deseada, tiene una calidad de imagen y una resolución tan alta como sea posible y que otras imágenes de fondos o similares, pueden tener una calidad de imagen relativamente baja. Por el contrario, en una aplicación para conversación de videoconferencia o similar, es preferido, a veces, que una imagen de fondo tenga baja calidad de imagen.

Tal como se describió con anterioridad, una señal de brillo es suficiente para el reconocimiento fácil y, por lo tanto, no es necesario un proceso de interpolación cromática de imagen de alta calidad. Sin embargo, en un sistema de cámara convencional, se pone en práctica un proceso altamente derrochador de recursos, en el que se realiza una compresión de datos de una imagen, para la que la cámara realiza una interpolación cromática de imagen de alta calidad, que se pone en práctica, además, con el fin de realizar un reconocimiento facial utilizando una imagen después de la descompresión. La forma de realización requiere el uso de menos datos, ya que una imagen después de una interpolación cromática simple, se reduce todavía más y se utiliza para el reconocimiento facial. Por otro lado, para una zona frontal, se puede recibir una imagen RAW de la cámara y se puede realizar una interpolación cromática de imagen de alta calidad utilizando una capacidad informática del terminal concentrador 20. De este modo, aunque la cantidad de datos a transmitir al terminal concentrador, desde la cámara, es pequeño, y la latencia es también pequeña, la calidad de la imagen de la zona facial se mejora, además, en comparación con un sistema de cámara convencional.

En un sistema de cámara convencional, solamente se puede transmitir la imagen RAW directamente, o transmitir la imagen RAW después de la interpolación seguida por una compresión de JPEG. En el primer caso, puesto que aumenta la cantidad de datos, una banda de comunicación se convierte en un denominado 'cuello de botella' lo que causa una latencia. En el último caso, puesto que se realiza una compresión con pérdida, la mejora en la calidad de la imagen en una etapa posterior está limitada. En general, la interpolación cromática determina la calidad de imagen de una imagen. Sin embargo, la capacidad informática se utiliza, en gran medida, para ello. Por lo tanto, es ventajoso realizar la interpolación cromática en el terminal concentrador que tiene suficientes recursos operativos. En la forma de realización, puesto que solamente se transmite una parte específica, tal como un rostro, una mano, o elemento similar, en una imagen RAW, al terminal concentrador, la cantidad de datos a transmitirse es pequeña y, por lo tanto, se reduce la latencia. Al mismo tiempo, el grado de libertad para la mejora de la calidad de imagen, utilizando la imagen RAW, en el terminal concentrador, puede aumentarse en gran medida.

La Figura 8 es un diagrama que ilustra la descripción de un proceso de imagen cuando se aplica el sistema de cámara de baja latencia 10, de conformidad con la forma de realización, a una aplicación de videoconferencia. En este ejemplo, una cámara y un terminal concentrador están conectados por intermedio de un USB, y se supone que se está ejecutando un programa de aplicación en el que una parte del rostro de un usuario, que está en la videoconferencia, se detecta y a continuación, se transmite la imagen, teniendo una imagen de alta calidad y una imagen de alta resolución para la parte facial, para un terminal concentrador de un destino de chat.

En esta aplicación, se supone que el terminal concentrador 20 proporciona instrucciones a la cámara 100 para transmitir solamente una imagen RAW y una imagen reducida de 1/16. Un bloque B16 de imagen reducida a 1/16 se selecciona por la unidad de selección de tamaño 152A (S50), y se codifica por la unidad de codificación 158 (S52). Para la imagen RAW, la unidad de escritura de bloques 152 recibe información de posición de una parte específica procedente del terminal concentrador (S54), y la unidad 152B, para la selección de un bloque que ha de cortarse,

- 5 selecciona un margen E como un bloque de parte específica y lo proporciona, a la salida, al bloque de parte específica (S56), obteniéndose el margen E mediante la expansión, de forma vertical y horizontal, de una zona D especificada por la información de posición de un resultado de reconocimiento facial por una cantidad de pixels predeterminados. Lo que antecede se debe al hecho de que, aunque solamente se reconoce una parte del color de la piel en el reconocimiento facial, se necesita una imagen de la parte de la cabeza completa para el chat. Este margen se establece, preferentemente, de modo que el margen coincida con el límite de los bloques. El bloque de parte específica de la imagen RAW y el bloque de imagen reducida, así obtenidos, se empaquetan y transmiten al terminal concentrador 20 (S58).
- 10 En el terminal concentrador 20, el bloque de parte específica de la imagen RAW pasa por un proceso de interpolación cromática y un proceso de alta calidad (S62) y se proporciona, a la salida, como una imagen de alta calidad (S64). La imagen reducida pasa por un proceso de decodificación (S66) y se amplía, además, a un tamaño de 1/1 (S68). Se pone en práctica una síntesis de imagen para sustituir una parte que corresponde a la cabeza del usuario, en la imagen ampliada, con una imagen de alta calidad (S70) y se proporciona una imagen sintetizada. Al mismo tiempo, se utiliza una señal Y de la imagen reducida para la detección facial (S72), y la información de posición de un rostro detectado se retroinforma a la cámara 100 (S74) y se utiliza para seleccionar un bloque de parte específica.
- 15 Se puede poner en práctica un proceso para el reconocimiento facial en cualquier momento durante una conversación de videoconferencia. Sin embargo, en la práctica, es suficiente realizar el proceso para cada pluralidad de tramas dentro de un intervalo temporal predeterminado (p.ej., cada segundo). Esto es así puesto que normalmente se generan 60 tramas o más por segundo para una imagen en movimiento, de modo que la cabeza del usuario no se mueva mucho entre tramas.
- 20 Como en esta realización, a modo de ejemplo, una imagen facial es más importante en el caos de una aplicación para videoconferencia, y el grado de importancia es bajo para otras partes del cuerpo, y fondo, y similar. Por lo tanto, se puede obtener una imagen facial de imagen de alta calidad y alta definición mediante el corte de solamente un bloque que corresponde a una parte facial, a partir de la imagen RAW y a continuación, realizar un proceso de interpolación cromática y un proceso de alta definición, que utiliza el bloque objeto de corte, usando suficientes recursos informáticos del terminal concentrador. Al mismo tiempo, utilizando una imagen comprimida para partes diferentes de un rostro, se puede reducir el tráfico entre la cámara y el terminal concentrador y, además, se puede reducir la latencia.
- 25 Tal como se explicó con anterioridad, de conformidad con la presente forma de realización, se recibe una imagen RAW desde una cámara, y se pone en práctica un proceso de interpolación cromática de imagen de alta calidad en un terminal concentrador, para una parte necesaria para tener una imagen de alta calidad y una alta resolución. Para otras partes que pueden tener una baja calidad de imagen y una baja resolución, se recibe una imagen reducida desde la cámara y se amplía para su utilización. En consecuencia, se puede suprimir la latencia debida a los procesos realizados en el interior de la cámara y en la comunicación entre la cámara y el terminal concentrador, mientras que se asegura una imagen de imagen de alta calidad parcial. Por lo tanto, se puede realizar una respuesta, más rápida que nunca antes, en una aplicación de una interfaz de usuario. Por lo tanto, existe una posibilidad de que, a modo de ejemplo, el movimiento del cursor, la operación de marcación, la operación del jugador o similar, se puede obtener en una pantalla sobre la base de un reconocimiento de imagen del gesto del usuario con alta capacidad de respuesta. Además, se puede proporcionar una imagen con un movimiento suave y menos latencia en una aplicación para videoconferencia. Tal como se describe, solamente se pueden seleccionar y transmitir datos de imagen que cumplen las necesidades para cada aplicación.
- 30 Tal como se explicó con anterioridad, de conformidad con la presente forma de realización, se recibe una imagen RAW desde una cámara, y se pone en práctica un proceso de interpolación cromática de imagen de alta calidad en un terminal concentrador, para una parte necesaria para tener una imagen de alta calidad y una alta resolución. Para otras partes que pueden tener una baja calidad de imagen y una baja resolución, se recibe una imagen reducida desde la cámara y se amplía para su utilización. En consecuencia, se puede suprimir la latencia debida a los procesos realizados en el interior de la cámara y en la comunicación entre la cámara y el terminal concentrador, mientras que se asegura una imagen de imagen de alta calidad parcial. Por lo tanto, se puede realizar una respuesta, más rápida que nunca antes, en una aplicación de una interfaz de usuario. Por lo tanto, existe una posibilidad de que, a modo de ejemplo, el movimiento del cursor, la operación de marcación, la operación del jugador o similar, se puede obtener en una pantalla sobre la base de un reconocimiento de imagen del gesto del usuario con alta capacidad de respuesta. Además, se puede proporcionar una imagen con un movimiento suave y menos latencia en una aplicación para videoconferencia. Tal como se describe, solamente se pueden seleccionar y transmitir datos de imagen que cumplen las necesidades para cada aplicación.
- 35 Recientemente, el número de pixels ha aumentado incluso en cámaras web económicas, y la velocidad de escaneo ha aumentado también mediante el uso del sensor CMOS como un elemento de creación de imágenes. Con el fin de transmitir una imagen de gran tamaño, debe realizarse una compresión alta en una cámara. Sin embargo, lo anterior aumentaría el tiempo de procesamiento de la cámara. Por lo tanto, la latencia asociada con los procesos puestos en práctica dentro de la cámara, y la comunicación, se ha vuelto más evidente que con anterioridad. Dicho de otro modo, aunque la capacidad del elemento de creación de imágenes y la capacidad del terminal concentrador se han mejorado, entre ellos se genera un 'cuello de botella' y, por lo tanto, su capacidad no se puede aprovechar plenamente. Sin embargo, si se puede preparar un terminal concentrador que tenga un alto rendimiento operativo, resulta más ventajoso recibir una imagen RAW que no ha sido procesada por una cámara puesto que no existe deterioro en la calidad de imagen, permitiendo, de este modo, un procesamiento posterior de equilibrio de blancos y exposición.
- 40 En la forma de realización, la interpolación cromática en la cámara se simplifica con el fin de reducir la latencia y, para una parte que necesita tener una imagen de alta calidad y alta resolución, para un proceso de reconocimiento o similar, se recibe directamente una imagen RAW, procedente de la cámara, de modo que se pueda realizar la interpolación cromática en un terminal concentrador que tiene un más alto rendimiento informático. De este modo, el rendimiento del elemento de creación de imágenes se puede aprovechar al máximo mientras que se reduce la latencia.
- 45
- 50
- 55
- 60
- 65

Además, las imágenes reducidas jerarquizadas por un filtro piramidal, se preparan en todo momento en la cámara. De este modo, el terminal concentrador puede demandar una imagen reducida de un tamaño necesario para una aplicación que se está ejecutando en cualquier momento. A modo de ejemplo, si el usuario permanece estacionario, solamente necesita prepararse una parte específica de una imagen RAW, desde el principio, para ser utilizada en todo momento. Sin embargo, dicha situación es menos probable que suceda realmente. En la forma de realización, una parte específica de una imagen RAW se puede recibir, y se puede rastrear el movimiento del usuario utilizando imágenes reducidas en cualquier momento, con el fin de detectar el movimiento de una parte para el rostro o la mano. De este modo, un bloque de la imagen RAW que ha de demandarse se puede cambiar, de forma apropiada, de conformidad con el movimiento. Con respecto al tamaño que se requerirá de una imagen reducida, se puede seleccionar una imagen reducida de un tamaño óptimo de conformidad con el tamaño de un rostro o mano del usuario que ha de detectarse, el número de usuarios, la magnitud del movimiento del usuario, etc.

Además, un tamaño pequeño de una memoria intermedia necesaria para un proceso de imagen se permite en la unidad de escritura de bloques y la unidad de escritura de bloques. Esto se debe al hecho de que mientras la unidad de escritura de bloques realiza la escritura en los pixels de la memoria intermedia desde la unidad de interpolación cromática simple, o la unidad de filtro piramidal, cada vez que se realiza un escaneo de imagen para una fila, por el elemento de creación de imágenes, la unidad de escritura de bloques efectúa la lectura, para un bloque, desde la memoria intermedia cada vez que los pixels para un bloque se memorizan y transmiten en un paquete.

Además, en un dispositivo de procesamiento de imagen de conformidad con la forma de realización, no es necesario proporcionar un dispositivo de proceso de reconocimiento, o similar, en la cámara, como en un dispositivo convencional, lo que permite que el hardware sea pequeño. Un sistema de cámara con baja latencia se puede establecer solamente añadiendo una unidad de filtro piramidal y una unidad de transmisión de imágenes a una videocámara digital existente.

Lo anteriormente descrito es una explicación de la presente invención sobre la base de la forma de realización. La forma de realización está prevista para ser ilustrativa solamente, y será evidente para los expertos en esta técnica que se pueden desarrollar varias modificaciones a los elementos y procesos constituyentes y que dichas modificaciones están también dentro del alcance de la presente invención.

Combinaciones opcionales de los elementos constituyentes descritos en las formas de realización, y las puestas en práctica de la invención en la forma de métodos, aparatos, sistema, programas informáticos, y soportes de registro, se pueden poner en práctica, además, como modos operativos adicionales de la presente invención. Además de los procesos realizados en su orden anterior y en orden cronológico, el método descrito como un diagrama de flujo en la presente especificación, incluye procesos puestos en práctica en paralelo o de forma individual.

En la forma de realización, se selecciona un bloque específico solamente para una imagen RAW. Sin embargo, se puede seleccionar un bloque específico para otras imágenes interpoladas. A modo de ejemplo, cuando se detecta un gesto de una mano, una parte de la mano se determina para ser una parte específica con referencia a un rostro de usuario, y se transmite una imagen interpolada de una zona de una mano a un terminal concentrador. Cuando el movimiento del usuario es amplio, se puede rastrear el movimiento utilizando una imagen gruesa, es decir, una imagen interpolada de un pequeño tamaño. Cuando el movimiento del usuario es pequeño, se utiliza una imagen fina, es decir, una imagen interpolada de gran tamaño. Según se describe, dado que los datos de imagen jerarquizados están preparados por un filtro piramidal en todo momento en la cámara, se puede cambiar, de forma adecuada, un tamaño de imagen requerida de conformidad con la velocidad y la magnitud del movimiento del usuario. De modo similar, incluso cuando la cantidad de usuarios aumenta a dos, se puede seleccionar, de nuevo, rápidamente un bloque de imagen RAW puesto que se está supervisando una imagen completa. Además, el procesamiento de una parte específica utilizando una imagen RAW permite la detección de un pequeño movimiento, tal como un movimiento de la yema del dedo, lo que resulta extremadamente complicado en el reconocimiento utilizando una cámara convencional.

En la forma de realización, se describe que un proceso de interpolación cromática en la cámara se simplifica en comparación con el del terminal concentrador. Sin embargo, si un efecto de latencia asociado con un proceso de interpolación cromática es relativamente pequeño, se puede poner en práctica un proceso de interpolación cromática que tenga un rendimiento equivalente al del terminal concentrador. Dicho de otro modo, siempre que se realice un proceso de interpolación cromática de conformidad con un algoritmo que sea distinto en la cámara y el terminal concentrador, el tipo de algoritmo para un proceso de interpolación cromática no está limitado en la presente invención. De este modo, se pueden tratar dos tipos de imágenes decodificadas en el terminal concentrador y se aumenta el grado de libertad de composición de una imagen sintetizada que ha de enviarse a una pantalla.

En la forma de realización, se describe que se transmite una imagen reducida completa desde la cámara al terminal concentrador. Como alternativa, solamente se puede transmitir una parte de una imagen reducida. A modo de ejemplo, sólo se puede transmitir una señal Y de una imagen reducida de un tamaño de 1/16, y se puede transmitir una señal CbCr de una imagen reducida de 1/64 en función de las propiedades de JPEG.

En la forma de realización, un ordenador personal o consola de juego se ilustra a modo de terminal concentrador.

Como alternativa, un terminal concentrador puede ser un ordenador portátil, un dispositivo de juegos portátil, o similar. En este caso, una cámara se instala, de forma preferentemente sobre, o se incorpora en, a modo de ejemplo, la parte superior de una pantalla.

5 En la forma de realización, se describe que la detección del rostro, mano, etc. de un usuario, se pone en práctica en el terminal concentrador. Como alternativa, la detección se puede realizar en la cámara. A modo de ejemplo, se considera que la latencia que ha de aumentarse es relativamente pequeña para un proceso de reconocimiento que utiliza una señal de brillo, tal como un reconocimiento facial. En este caso, se puede seleccionar un bloque de una imagen RAW que corresponde a un rostro, en la cámara y luego, transmitirse al terminal concentrador sin esperar una instrucción procedente del terminal concentrador.

15 Se da a conocer una explicación como una aplicación que se utiliza cuando un gesto del usuario se utiliza en una interfaz y para cuando se realiza una videoconferencia. La presente invención se puede poner en práctica, con las modificaciones necesarias, para una aplicación arbitraria en donde se recibe una parte de un cuerpo o gesto de un usuario con el reconocimiento que sirve como una entrada, y una aplicación arbitraria en donde se utiliza la imagen del usuario.

DESCRIPCIÓN DE LOS NÚMEROS DE REFERENCIA

- 20 4 pantalla
- 6 usuario
- 10 sistema de cámara de baja latencia
- 25 20 terminal concentrador
- 22 unidad de comunicación
- 30 24 unidad de asignación
- 28 unidad de interpolación cromática de imagen de alta calidad
- 32 unidad de procesamiento de alta calidad
- 35 34 unidad de decodificación
- 36 unidad de ampliación
- 40 38 unidad de síntesis de imagen
- 40 unidad de salida de imagen
- 42 unidad de detección de parte específica
- 45 50 unidad de procesamiento de imagen
- 52 unidad de control
- 50 54 unidad de control de visualización
- 100 cámara
- 102 unidad de adquisición de imágenes
- 55 104 unidad de interpolación cromática
- 106 unidad de procesamiento de interpolación cromática simple
- 60 108 unidad de comunicación
- 110 primer filtro
- 120 segundo filtro
- 65 130 tercer filtro

	140	cuarto filtro
5	150	unidad de transmisión de imágenes
	152	unidad de escritura de bloques
	154	memoria intermedia
10	156	unidad de escritura de bloques
	158	unidad de codificación
15	160	unidad de empaquetado
	162	memoria intermedia de paquete
	164	unidad de control
20	170	unidad de filtro piramidal

#### APLICABILIDAD INDUSTRIAL

25 De conformidad con la presente invención, se puede obtener una imagen de alta resolución parcial utilizando una imagen bruta para una parte específica al mismo tiempo que se reduce la latencia asociada con una comunicación de imagen desde un dispositivo de captura de imágenes a un terminal concentrador.

30

**REIVINDICACIONES**

- 5 **1.** Un programa de procesamiento de imagen en movimiento incorporado en un soporte de grabación legible por ordenador no transitorio, que es operativo en un dispositivo de procesamiento de imagen en movimiento, que permite al dispositivo de procesamiento de imagen en movimiento obtener:
- un módulo de adquisición de imagen, configurado para adquirir una imagen bruta que se captura utilizando un elemento de creación de imágenes;
- 10 un módulo de interpolación cromática (*demosaicing*), configurado para realizar un proceso de interpolación cromática en la imagen bruta con el fin de obtener una imagen interpolada;
- un módulo de filtrado configurado para convertir la imagen interpolada en una pluralidad de imágenes reducidas cuyas resoluciones varían en etapas;
- 15 un módulo de selección, configurado para seleccionar una parte de la imagen bruta como una parte específica, siendo la parte específica una parte de la imagen bruta para la cual es necesaria una imagen de alta calidad y, además, para seleccionar una o más de entre la pluralidad de imágenes reducidas como una imagen reducida especificada; y
- 20 un módulo de transmisión configurado para transmitir a un terminal concentrador, para un procesamiento de imagen adicional, la parte específica de la imagen bruta y la imagen reducida especificada que han sido seleccionadas.
- 25 **2.** El programa de procesamiento de imagen en movimiento según la reivindicación 1, en donde el módulo de selección selecciona, de conformidad con una instrucción proporcionada por el terminal concentrador, la imagen reducida especificada como una imagen de fondo que ha de combinarse con la parte específica.
- 3.** El programa de procesamiento de imagen en movimiento según la reivindicación 1 o 2, en donde
- 30 el módulo de selección incluye:
- un módulo de escritura configurado para seleccionar y escribir en una memoria intermedia, un valor de píxel correspondiente a la parte específica cada vez que se emite un valor de píxel de la imagen bruta desde el elemento de creación de imágenes, y para seleccionar y escribir, en la memoria intermedia, un valor de píxel correspondiente
- 35 a la imagen reducida especificada, cada vez que se recibe un valor de píxel de una imagen reducida que se convierte por el módulo de filtrado; y
- un módulo de lectura configurado para leer un valor de píxel desde la memoria intermedia cada vez que el valor de píxel almacenado en la memoria intermedia alcanza una magnitud predeterminada de un bloque de imagen.
- 40 **4.** Un programa de procesamiento de imagen en movimiento incorporado en un soporte de grabación legible por ordenador no transitorio, que opera en un terminal concentrador que procesa una imagen en movimiento que es capturada por un dispositivo de captura de imágenes y luego proporciona, a la salida, la imagen en movimiento procesada a una pantalla, permitiendo al terminal concentrador obtener:
- 45 un módulo configurado para recibir, a la vez, una parte específica que es una parte de una salida de imagen bruta procedente de un elemento de creación de imágenes, y una imagen reducida que se obtiene al realizar un proceso de interpolación cromática y un proceso de reducción en la imagen bruta, en donde la parte específica es una parte de la imagen bruta para la cual es necesaria una imagen de alta calidad;
- 50 un módulo configurado para realizar el proceso de interpolación cromática en la parte específica de la imagen bruta;
- un módulo configurado para convertir la imagen reducida en una imagen de tamaño completo, ampliando la imagen reducida a un tamaño igual al de la imagen bruta; y
- 55 un módulo configurado para combinar la parte específica, sobre la cual se ha realizado el proceso de interpolación cromática, y la imagen de tamaño completo, con el fin de proporcionar, a la salida, una imagen sintetizada con resoluciones parcialmente diferentes a la pantalla.
- 60 **5.** Un programa de procesamiento de imagen en movimiento según la reivindicación 4 que comprende, además, un módulo de detección de parte específica, configurado para detectar, utilizando la imagen reducida, una parte específica que se establece para tener una imagen de alta calidad, en una imagen.
- 65 **6.** El programa de procesamiento de imagen en movimiento según la reivindicación 4 o 5, en donde un usuario, que utiliza una aplicación que usa el terminal concentrador, debe ser objeto de captura por el dispositivo de captura de imágenes, y en donde la parte específica es una parte del cuerpo del usuario que proporciona una instrucción de

funcionamiento a la aplicación.

7. Un soporte de grabación legible por ordenador, no transitorio, al que se incorpora el programa de procesamiento de imagen en movimiento de conformidad con una cualquiera de las reivindicaciones 1 a 6.

8. Un dispositivo de procesamiento de imagen en movimiento (102, 104, 170, 150, 108) que comprende:

una unidad de adquisición de imágenes (102), configurada para adquirir una imagen bruta que se captura utilizando un elemento de creación de imágenes;

una unidad de procesamiento de interpolación cromática (104), configurada para realizar un proceso de interpolación cromática sobre la imagen bruta, con el fin de generar una imagen interpolada;

una unidad de filtrado (170), configurada para convertir la imagen interpolada en una pluralidad de imágenes reducidas cuyas resoluciones varían en etapas;

una unidad de selección (150), configurada para seleccionar una parte de la imagen bruta como una parte específica, siendo la parte específica una parte de la imagen bruta para la que es necesaria una imagen de alta calidad y, además, para seleccionar una o más de entre la pluralidad de imágenes reducidas como una imagen reducida especificada;

una unidad de transmisión (108), configurada para transmitir, a un terminal concentrador, para un procesamiento adicional de imagen, la parte específica de la imagen bruta y la imagen reducida especificada que han sido seleccionadas.

9. Un dispositivo de captura de imágenes (100) que se proporciona con el dispositivo de procesamiento de imagen en movimiento según la reivindicación 8.

10. Un terminal concentrador (20) para la visualización de una imagen en movimiento capturada por un dispositivo de captura de imágenes (100) en una pantalla, que comprende:

una unidad de recepción (22) configurada para recibir, del dispositivo de captura de imágenes (100), a la vez: una parte específica que es una parte de una imagen bruta capturada usando un elemento de creación de imágenes, y una imagen reducida obtenida mediante la realización de un proceso de interpolación cromática, y un proceso de reducción en la imagen bruta, en donde la parte específica es una parte de la imagen bruta para la cual es necesaria una imagen de alta calidad;

una unidad de interpolación cromática (28) configurada para realizar el proceso de interpolación cromática en la parte específica de la imagen bruta;

una unidad de ampliación (36) configurada para proporcionar, a la salida, la imagen reducida como una imagen de tamaño completo mediante la ampliación de la imagen reducida a un tamaño igual al de la imagen bruta; y

una unidad de síntesis de imagen (38) configurada para combinar la parte específica sobre la cual se ha realizado el proceso de interpolación cromática, y la imagen de tamaño completo con el fin de proporcionar, a la salida, una imagen sintetizada con resoluciones parcialmente diferentes a la pantalla.

11. Un sistema de captura de imágenes (10) para procesamiento en el terminal concentrador (20) según la reivindicación 10, de una imagen en movimiento capturada por el dispositivo de captura de imágenes (100) según la reivindicación 10, y generando, a la salida, la imagen en movimiento procesada para una pantalla (4), en donde

la unidad de procesamiento de interpolación cromática (104) del dispositivo de captura de imágenes (100) es una primera unidad de procesamiento de interpolación cromática (104) y

la unidad de interpolación cromática (28) del terminal concentrador (20) es una segunda unidad de procesamiento de interpolación cromática (28).

12. El sistema de captura de imágenes según la reivindicación 11, donde la capacidad de procesamiento de la primera unidad de procesamiento de interpolación cromática (104) del dispositivo de captura de imágenes (100), es inferior a la de la segunda unidad de procesamiento de interpolación cromática (28) del terminal concentrador (20).

13. Un método de procesamiento de imagen en movimiento, que comprende:

la adquisición y memorización, en una memoria intermedia, de una imagen bruta que se captura utilizando un elemento de creación de imágenes;

la realización de un proceso de interpolación cromática sobre la imagen bruta, con el fin de generar una imagen interpolada;

5 la conversión, utilizando un filtro piramidal, de la imagen interpolada en una pluralidad de imágenes reducidas cuyas resoluciones varían en etapas;

10 la selección de una parte de la imagen bruta como una parte específica, siendo la parte específica una parte de la imagen bruta para la que es necesaria una imagen de alta calidad, y la selección, además, de una o más de entre la pluralidad de imágenes reducidas como una imagen reducida especificada;

la transmisión, a un terminal concentrador, para el procesamiento adicional de imagen, de la parte específica de la imagen bruta y la imagen reducida especificada, que han sido seleccionadas.

15 **14.** Un método de procesamiento de imagen en movimiento para el procesamiento, en un terminal concentrador, de una imagen en movimiento que se captura por un dispositivo de captura de imágenes y a continuación, la generación, a la salida, de la imagen en movimiento procesada, para una pantalla, que comprende:

20 la recepción, a la vez, de una parte específica, que es una parte de una imagen bruta, procedente de un elemento de creación de imágenes, y una imagen reducida que se obtiene mediante la realización de un proceso de interpolación cromática, y un proceso de reducción sobre la imagen bruta, en donde la parte específica es una parte de la imagen bruta para la que es necesaria una imagen de alta calidad;

la realización del proceso de interpolación cromática sobre la parte específica de la imagen bruta;

25 la conversión de la imagen reducida en una imagen de tamaño completo mediante la ampliación de la imagen reducida a un tamaño igual al de la imagen bruta; y

30 la combinación de la parte específica, en la que se ha realizado el proceso de interpolación cromática y la imagen en tamaño completo, con el fin de proporcionar a la salida, una imagen sintetizada con resoluciones parcialmente diferentes, a la pantalla.

FIG.1

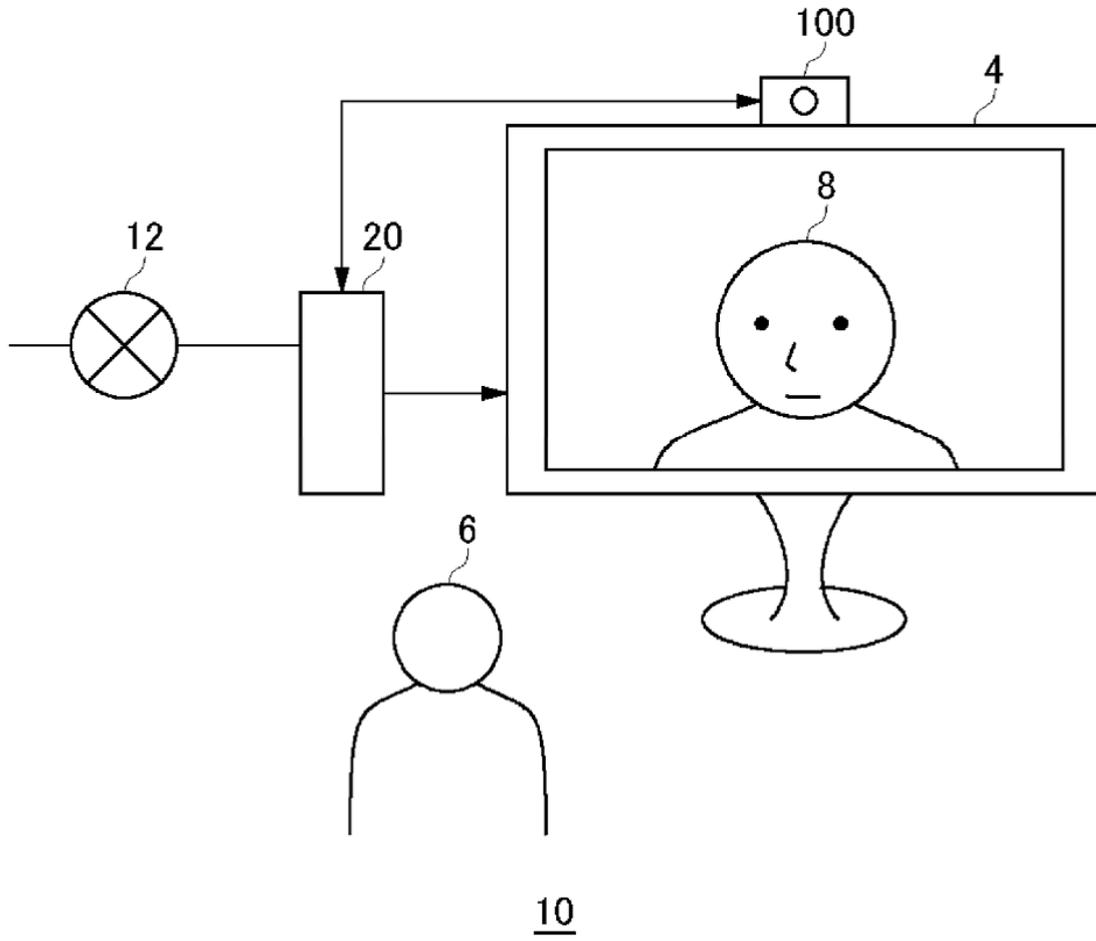
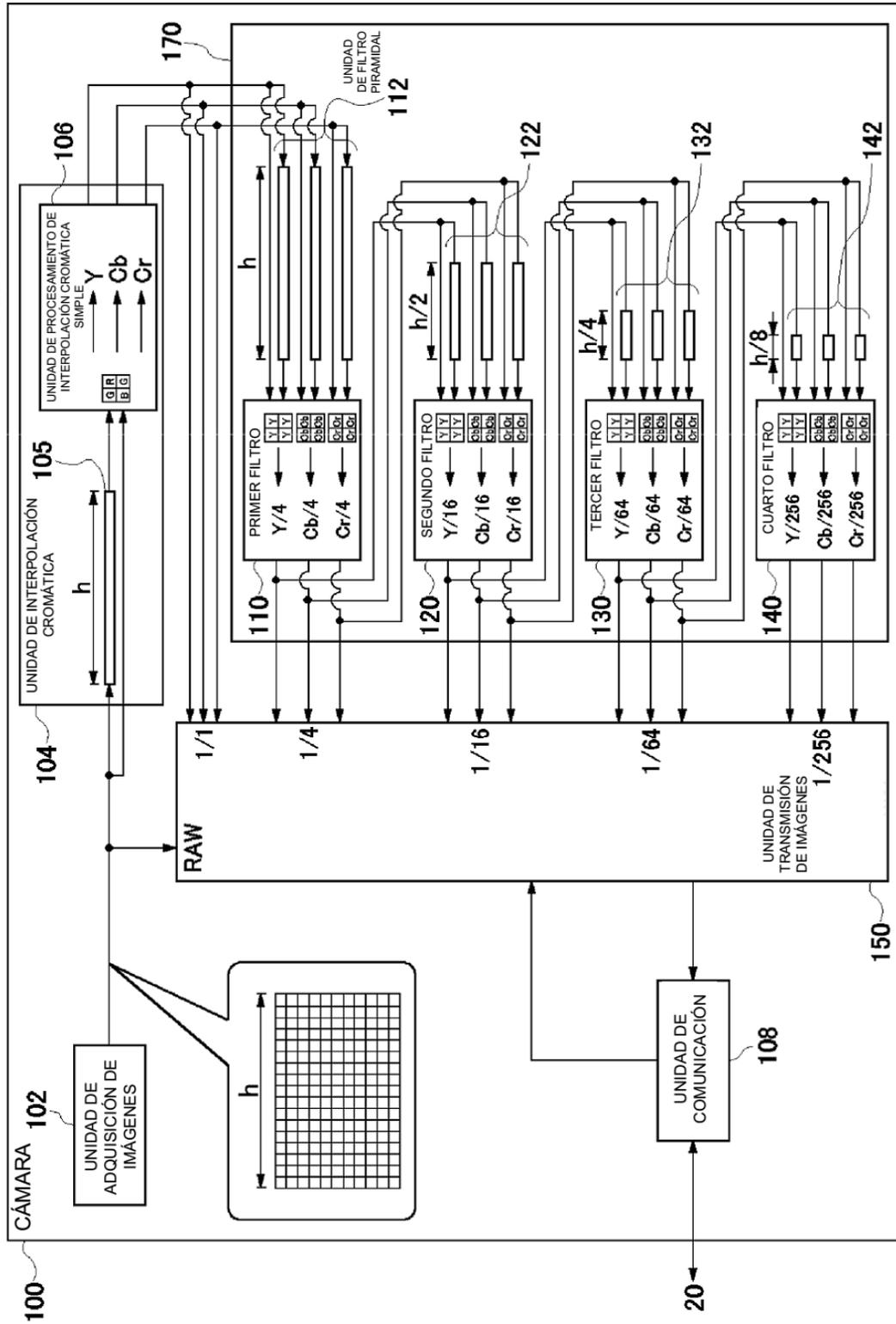


FIG.2



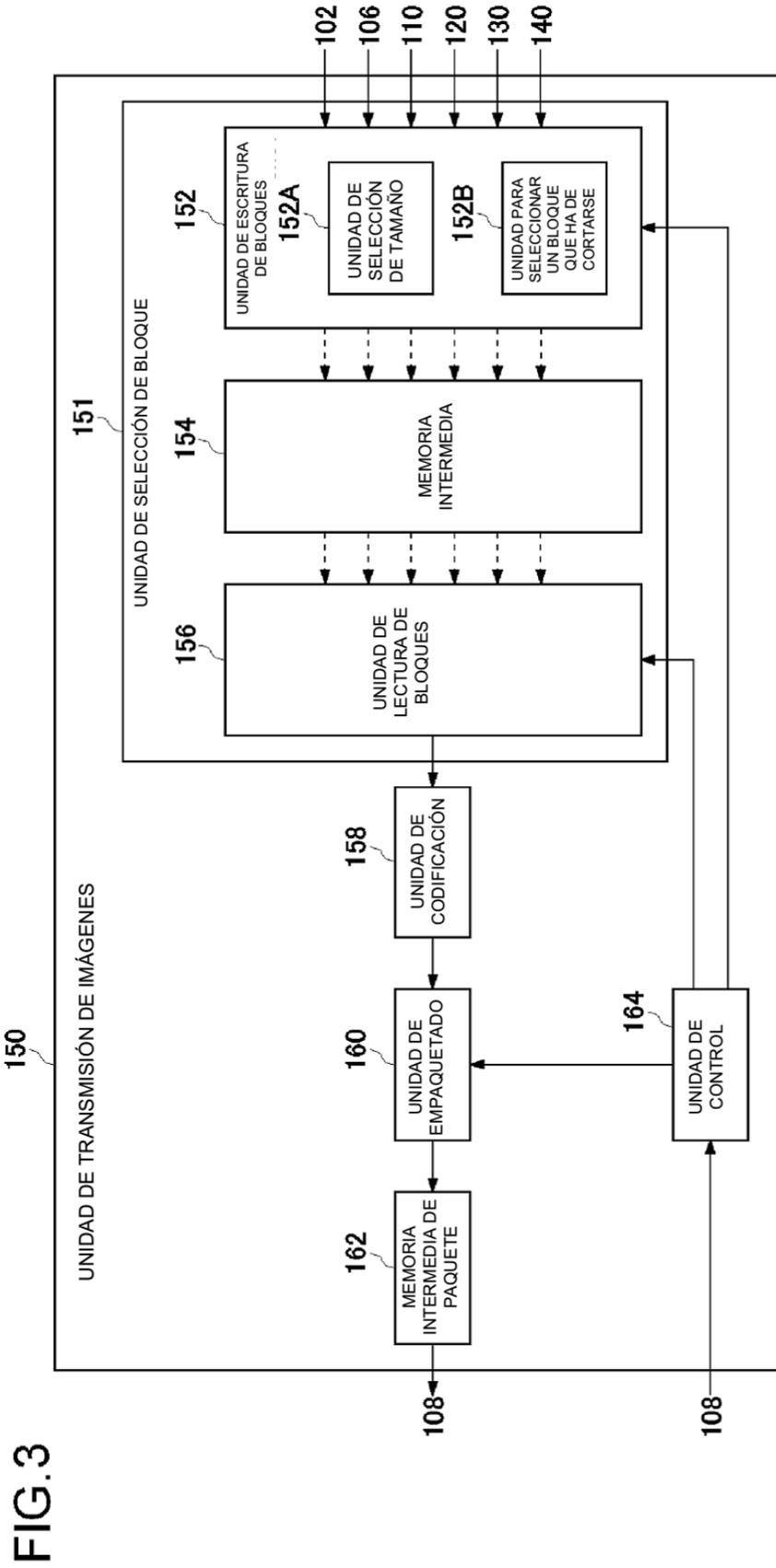


FIG.3

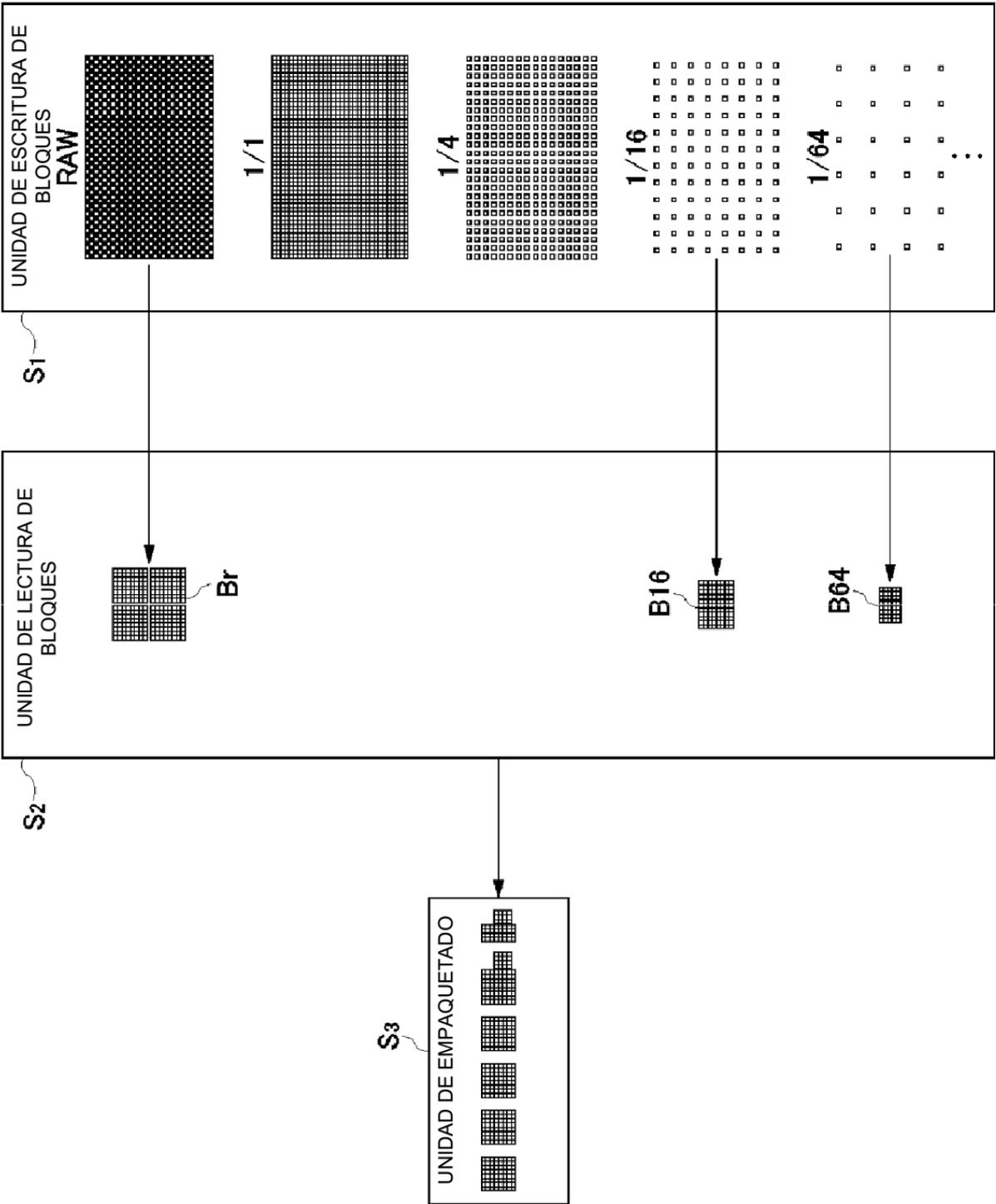


FIG.4

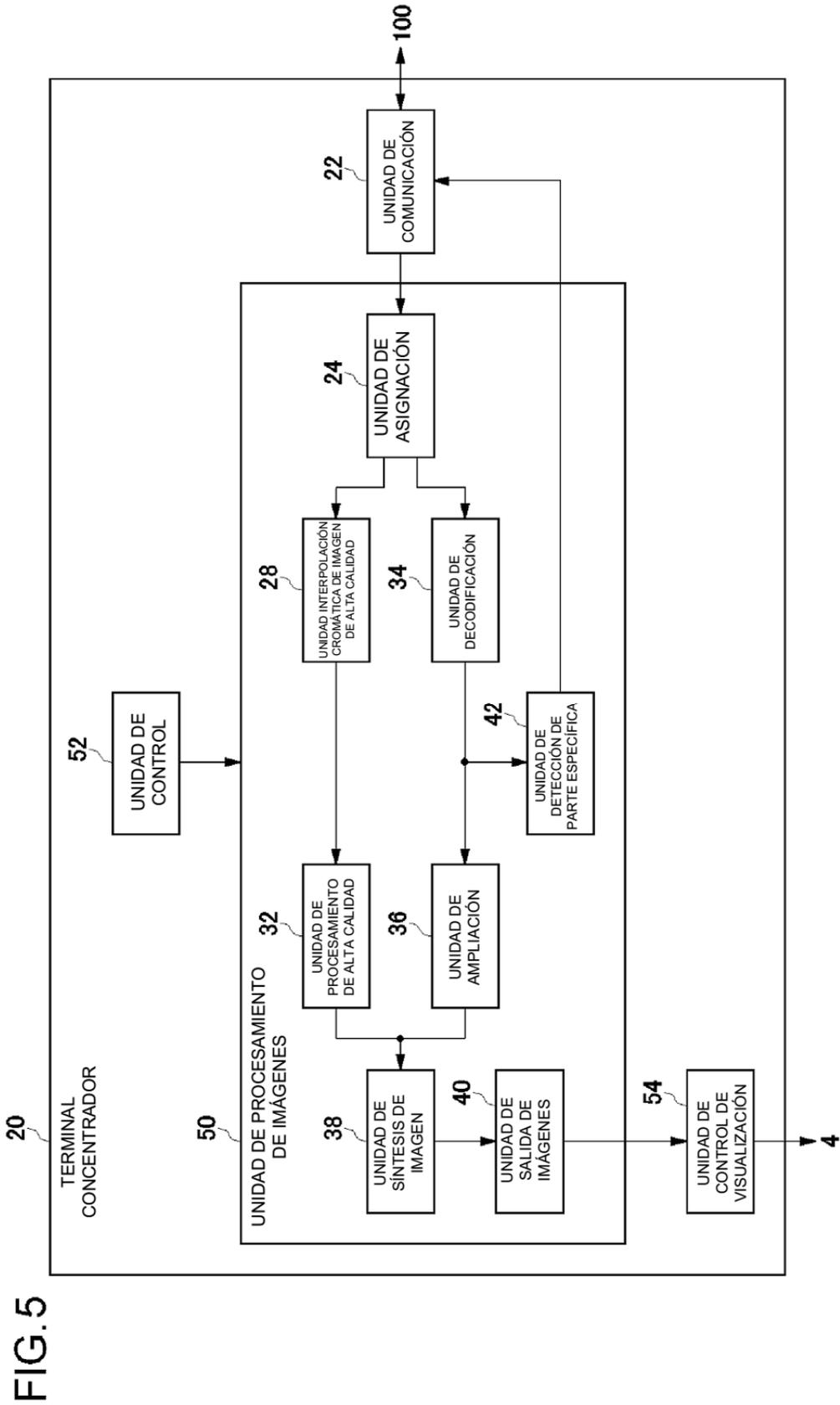


FIG. 5

FIG.6

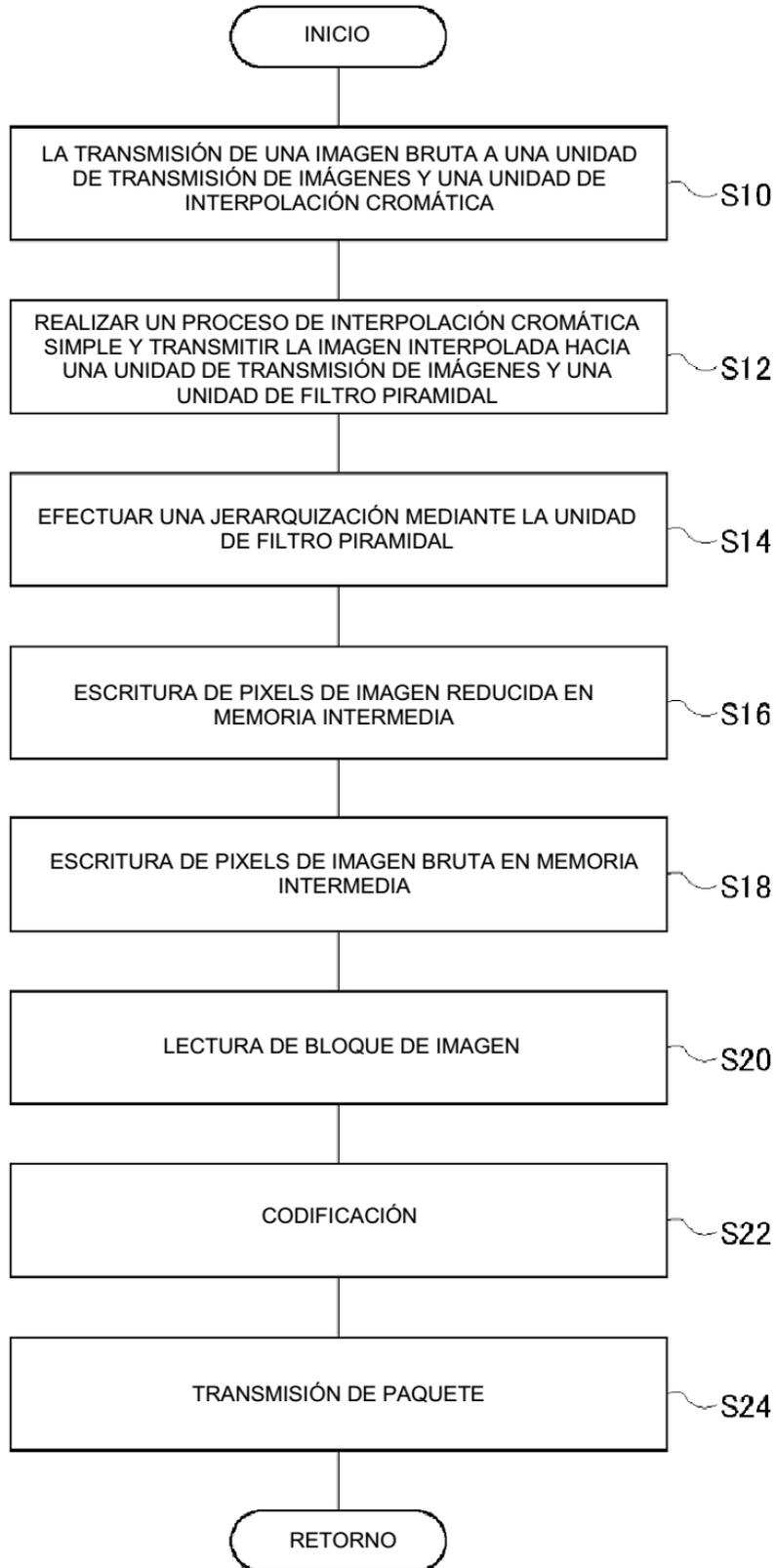
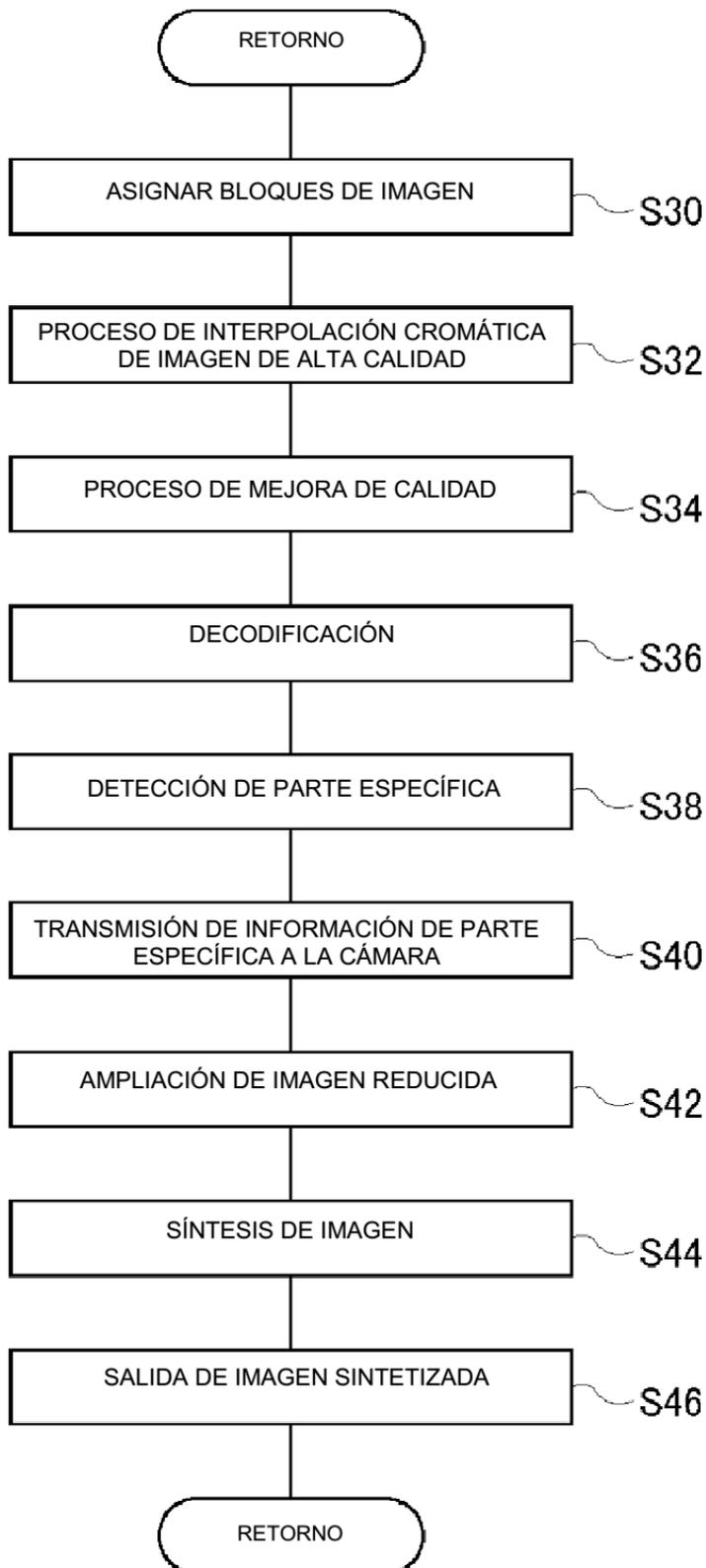


FIG.7



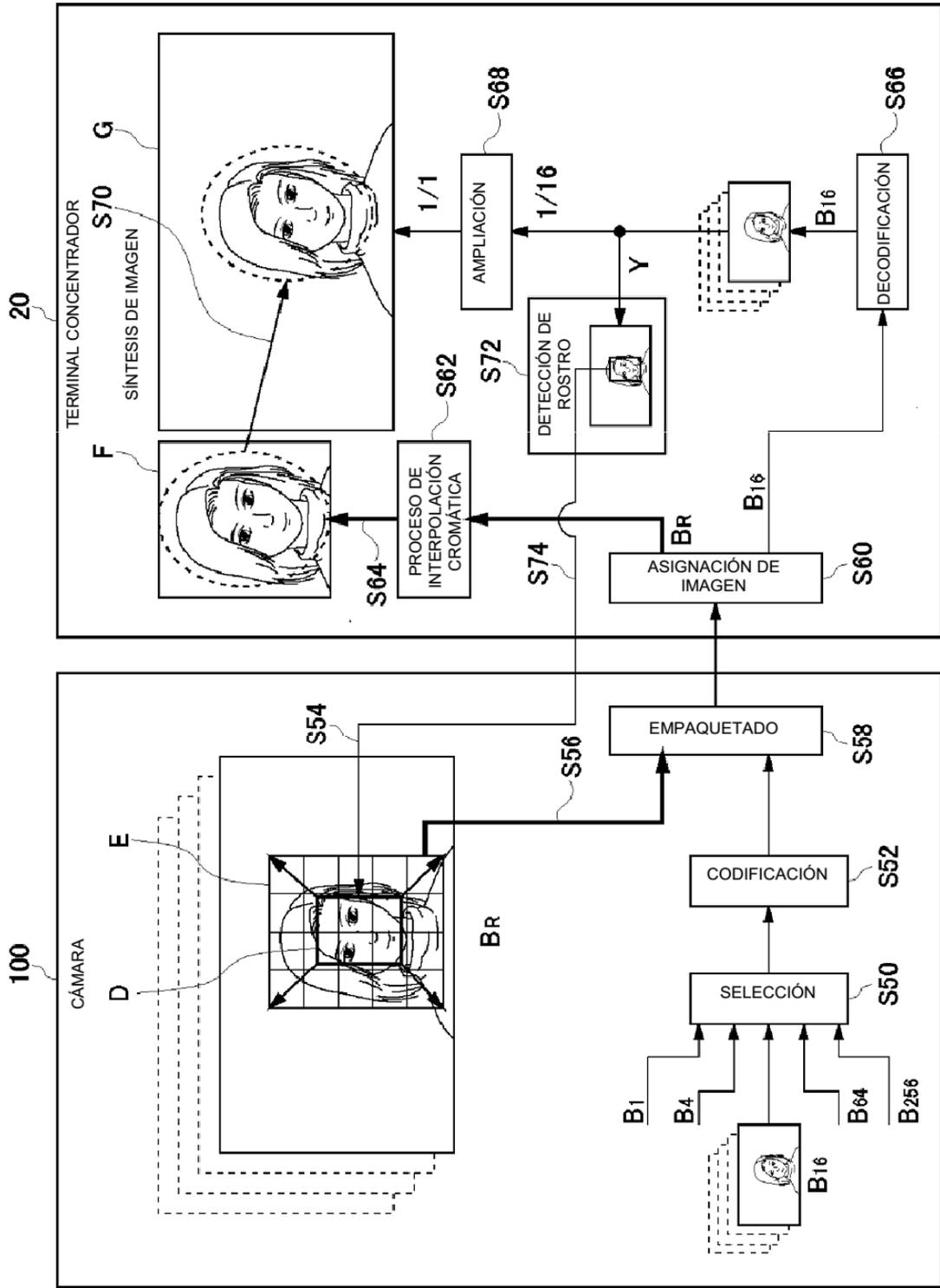


FIG.8