



19



OFICINA ESPAÑOLA DE
PATENTES Y MARCAS

ESPAÑA

11 Número de publicación: **2 306 380**

51 Int. Cl.:

H04N 7/32 (2006.01)

H03M 7/36 (2006.01)

H04N 7/26 (2006.01)

12

TRADUCCIÓN DE PATENTE EUROPEA

T3

86 Número de solicitud europea: **06122291 .5**

86 Fecha de presentación : **16.04.2003**

87 Número de publicación de la solicitud: **1739972**

87 Fecha de publicación de la solicitud: **03.01.2007**

54

Título: **Procedimiento de codificación de imágenes en movimiento y procedimiento de decodificación de imágenes en movimiento.**

30

Prioridad: **19.04.2002 JP 2002-118484**

45

Fecha de publicación de la mención BOPI:
01.11.2008

45

Fecha de la publicación del folleto de la patente:
01.11.2008

73

Titular/es:
MATSUSHITA ELECTRIC INDUSTRIAL Co., Ltd.
1006, Oaza Kadoma
Kadoma-shi, Osaka 571-8501, JP

72

Inventor/es: **Hagai, Makoto;**
Kadono, Shinya;
Kondo, Satoshi y
Abe, Kiyofumi

74

Agente: **Roeb Díaz-Álvarez, María**

ES 2 306 380 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).

ES 2 306 380 T3

DESCRIPCIÓN

Procedimiento de codificación de imágenes en movimiento y procedimiento de decodificación de imágenes en movimiento.

Campo técnico

La presente invención se refiere a un procedimiento de codificación y decodificación de datos de imágenes en movimiento así como un medio de grabación en el que se graba un programa para ejecutar estos procedimientos como software.

Técnica anterior

En los últimos años, junto con un desarrollo de aplicaciones multimedia como imagen, audio y texto, se ha generalizado el tratamiento de toda clase de medios de manera integrada. Sin embargo, es indispensable una técnica de compresión de información para almacenamiento y transmisión de datos ya que una imagen digitalizada contiene una enorme cantidad de datos. Por otra parte, también es importante una estandarización de técnicas de compresión para interactuar con datos de imagen comprimidos. Los estándares de técnicas de compresión de imagen incluyen H.261, H.263 establecidos por la ITU (Unión Internacional de Telecomunicaciones) y MPEG (Grupo de Expertos en Imágenes en Movimiento)-1, MPEG-2 y MPEG-4 establecidos por la ISO (Organización Internacional para la Estandarización).

Una predicción entre imágenes que acompaña la compensación de movimiento puede citarse como una técnica compartida entre estos procedimientos de codificación de imágenes en movimiento. En la compensación de movimiento basada en estos procedimientos de codificación de imágenes en movimiento, una imagen de una imagen de entrada es dividida en bloques, cada uno de los cuales tiene un tamaño predeterminado, y se genera una imagen predictiva para cada bloque usando vectores de movimiento, que indican respectivamente un movimiento entre imágenes. Las siguientes predicciones se emplean para predicción entre imágenes según el MPEG: una predicción adelantada para una predicción que usa una sola imagen cuyo momento de visualización es anterior al de una imagen actual que ha de ser codificada; una predicción hacia atrás para una predicción que usa una sola imagen cuyo momento de visualización es posterior al de la imagen actual; una predicción bidireccional para una predicción que usa dos imágenes, es decir, una imagen cuyo momento de visualización es anterior al de la imagen actual y la otra imagen cuyo momento de visualización es posterior al de la imagen actual (véase, por ejemplo, la referencia ISO/IEC 14496-2:1999(E) Tecnología de información - codificación de objetos audiovisuales Parte 2: Visual (1999-12-01) págs. 150 7.6.7 Estructura de predicción temporal).

En el MPEG, una imagen de referencia que ha de usarse se determina únicamente dependiendo del tipo de predicción entre imágenes y no puede seleccionarse una imagen de referencia arbitraria. Mientras tanto, en el H.264, que actualmente está bajo proceso de estandarización por la ITU, se toma en consideración una predicción bidireccional que se expande de manera que pueden seleccionarse dos imágenes de referencia arbitrarias de una pluralidad de imágenes codificadas almacenadas en una memoria de imagen independientemente del momento de visualización de la imagen actual.

La Fig. 1 es un diagrama de bloques que muestra una estructura de un aparato de codificación de imágenes en movimiento según el estándar H.264. El aparato convencional de codificación de imágenes en movimiento mostrado en la Fig. 1 es un aparato para ejecutar un procedimiento de codificación de imágenes en movimiento que permite la selección de dos imágenes de referencia arbitrarias de imágenes codificadas plurales cuando se utiliza la predicción entre imágenes.

Este aparato de codificación de imágenes en movimiento incluye, como se muestra en la Fig. 1, una unidad de estimación de movimiento 301, una unidad de interpolación de píxeles 102, un restador 103, una unidad de codificación de imágenes 104, una unidad de decodificación de imágenes 105, un sumador 106, una unidad de codificación de longitud variable 302, una memoria intermedia de imágenes múltiples 108 y un conmutador 109.

El aparato de codificación de imágenes en movimiento divide en bloques unos datos de imagen introducidos *Img* y realiza el procesamiento para cada uno de los bloques. El restador 103 resta unos datos de imagen predictiva *Pred* de los datos de imagen *Img* introducidos en el aparato de codificación de imágenes en movimiento y los genera como datos residuales *Res*. La unidad de codificación de imágenes 104 realiza procesamiento de codificación de imágenes, como transformación ortogonal y cuantificación, sobre los datos residuales introducidos *Res* y los genera como datos residuales codificados *ERes* que incluyen coeficientes transformados ortogonales cuantificados. La unidad de decodificación de imágenes 105 realiza procesamiento de decodificación de imágenes, como cuantificación inversa y transformación ortogonal inversa, sobre los datos residuales codificados introducidos *ERes* y los genera como datos residuales decodificados *DRes*. El sumador 106 suma los datos residuales decodificados *DRes* a los datos de imagen predictiva *Pred* y los genera como datos de imagen reconstruidos *Recon*. De los datos de imagen reconstruidos *Recon*, los datos que tienen la posibilidad de ser usados para referencia en la predicción entre imágenes posterior se almacenan en la memoria intermedia de imágenes múltiples 108.

Aquí, una predicción por interpolación que usa dos imágenes de referencia realizada por el aparato convencional de codificación de imágenes en movimiento se describe con referencia a la Fig. 2. La Fig. 2 es un diagrama conceptual de la predicción por interpolación que usa imágenes de referencia plurales. Aquí, una imagen Pic es una imagen actual que ha de ser codificada. Las imágenes FwRef1~FwRef3 representan imágenes codificadas que tienen respectivamente un momento de visualización anterior al de la imagen actual mientras que las imágenes BwRef1~BwRef3 representan imágenes codificadas que tienen respectivamente un momento de visualización posterior al de la imagen actual. Se predice un bloque Blk1 usando valores de píxeles en un bloque de referencia RefBlk1 incluido en la imagen FwRef3 cuyo momento de visualización es anterior al de la imagen actual Pic y valores de píxeles en un bloque de referencia RefBlk2 incluido en la imagen BwRef1 cuyo momento de visualización es posterior al de la imagen actual Pic. Se predice un bloque Blk2 usando valores de píxeles en bloques de referencia RefBlk21 y RefBlk22 incluidos en dos imágenes FwRef1 y FwRef2 que tienen respectivamente un momento de visualización anterior al de la imagen actual. Se predice un bloque Blk3 usando valores de píxeles en bloques de referencia RefBlk31 y RefBlk32 incluidos en dos imágenes BwRef1 y BwRef2 que tienen respectivamente un momento de visualización posterior al de la imagen actual. Concretamente, un resultado de interpolar píxeles en las áreas que corresponden a los dos bloques de referencia usando un procedimiento predeterminado como el que usa un valor medio se considera que es una imagen predictiva. La característica del aparato convencional de codificación de imágenes en movimiento es realizar predicción basándose en bloque por bloque usando dos imágenes de referencia arbitrarias como se muestra en la Fig. 2. Un procedimiento para predecir con el uso de dos imágenes de referencia arbitrarias como el descrito anteriormente se denomina en lo sucesivo “predicción por interpolación de imágenes de referencia plurales”. El procedimiento de predicción incluye un procedimiento en el que un bloque incluido en una sola imagen arbitraria se usa directamente como imagen predictiva y la predicción entre imágenes aparte del procedimiento de generación de una imagen predictiva usando la interpolación de píxeles como se describió anteriormente, y es posible cambiar el procedimiento de predicción basándose en bloque por bloque.

La unidad de estimación de movimiento 301 determina un tipo de predicción para el bloque, imágenes de referencia y vectores de movimiento que han de usarse para predicción entre imágenes realizada sobre el bloque actual introducido que ha de ser codificado y genera un tipo de predicción PredType, números de imágenes de referencia RefNo1, RefNo2, y vectores de movimiento MV1, MV2. La estimación de movimiento 301 genera dos números de imágenes y dos vectores de movimiento ya que se seleccionan dos imágenes de referencia cuando se utiliza la predicción por interpolación de imágenes de referencia plurales. Aquí, la memoria intermedia de imágenes múltiples 108 genera un bloque de referencia RefBlk1 que corresponde al número de imagen de referencia RefNo1 y el vector de movimiento MV1 y el bloque de referencia RefBlk2 que corresponde al número de imagen de referencia RefNo2 y el vector de movimiento MV2. La unidad de interpolación de píxeles 102 realiza interpolación para los píxeles con respecto a los dos bloques de referencia RefBlk1 y RefBlk2 usando el valor medio y lo genera como un bloque interpolado RefPol. Por otra parte, en el caso de usar una predicción entre imágenes aparte de una predicción por interpolación de imágenes de referencia plurales, la unidad de estimación de movimiento 301 selecciona una sola imagen de referencia y, por lo tanto, genera un solo número de imagen de referencia RefNo1 y un solo vector de movimiento MV1. En este caso, la memoria intermedia de imágenes múltiples 108 genera un bloque de referencia RefBlk con respecto al número de imagen de referencia RefNo1 y el vector de movimiento MV1.

Cuando el tipo de predicción determinado por la unidad de estimación de movimiento 301 indica una predicción por interpolación de imágenes de referencia plurales, el conmutador 109 se cambia a un lado “1” y el bloque interpolado RefPol se usa como datos de imagen predictiva Pred. Cuando el tipo de predicción PredType indica una predicción entre imágenes distinta de una predicción por interpolación de imágenes de referencia plurales, el conmutador SW11 se cambia a un lado “0” y el bloque de referencia RefBlk se usa como datos de imagen predictiva Pred. La unidad de codificación longitud variable 302 realiza codificación de longitud variable sobre los datos residuales codificados ERes, el tipo de predicción PredType, los números de imágenes de referencia RefNo1, RefNo2 y los vectores de movimiento MV1, MV2 y luego los genera como datos de imágenes en movimiento codificados Str0.

La Fig. 3 es un diagrama conceptual de un formato de datos de imagen en movimiento codificada usado por el aparato convencional de codificación de imágenes en movimiento. Los datos codificados equivalentes a una sola imagen, Picture, están compuestos de datos codificados equivalentes a un solo bloque, Block, donde cada bloque compone una imagen, y similar. Aquí, los datos codificados equivalentes a un solo bloque, Block, presentan datos codificados de un bloque en el que se realiza predicción por interpolación de imágenes de referencia plurales, e incluyen en los datos codificados los números de imágenes de referencia RefNo1, RefNo2, los vectores de movimiento MV1, MV2, con respecto a las dos imágenes de referencia, el modo de predicción PredType, y similares.

La Fig. 4 es un diagrama de bloques que muestra una estructura del aparato convencional de decodificación de imágenes en movimiento. El aparato de decodificación de imágenes en movimiento incluye, como se muestra en la Fig. 4, una unidad de decodificación de longitud variable 601, una unidad de compensación de movimiento 602, una unidad de decodificación de imágenes 404, un sumador 405, una unidad de interpolación de píxeles 406, una memoria intermedia de imágenes múltiples 407 y un conmutador 408.

La unidad de decodificación de longitud variable 601 realiza decodificación de longitud variable sobre los datos de imagen codificados introducidos Str0 y genera los datos residuales codificados ERes, los vectores de movimiento MV1, MV2, los números de imágenes de referencia RefNo1, RefNo2 y el tipo de predicción PreType. La unidad de decodificación de imágenes 404 realiza procesamiento de decodificación de imágenes, como cuantificación inversa y transformación ortogonal inversa, sobre los datos residuales introducidos ERes y genera datos residuales decodificados

ES 2 306 380 T3

DRes. El sumador 405 suma los datos residuales decodificados DRes a los datos de imagen predictiva Pred y los genera como datos de imagen decodificados DImg fuera del aparato de decodificación de imágenes. La memoria intermedia de imágenes múltiples 407 almacena los datos de imagen decodificados DImg para predicción entre imágenes.

5 La unidad de compensación de movimiento 602 genera números de imágenes de referencia NRefNo1, NRefNo2 de los bloques de referencia necesarios para predicción entre imágenes según el tipo de predicción PredType así como los vectores de movimiento MV1, MV2 y ordena a la memoria intermedia de imágenes múltiples 407 que genere los bloques de referencia. Cuando el tipo de predicción PredType indica una predicción por interpolación de imágenes de referencia plurales, la memoria intermedia de imágenes múltiples 407 genera el bloque de referencia RefBlk1 que
10 corresponde al número de imagen de referencia NRefNo1 y el vector de movimiento NMV1 así como el bloque de referencia RefBlk2 que corresponde al número de imagen de referencia NRefNo2 y el vector de movimiento NMV2. La unidad de interpolación de píxeles 406 interpola los píxeles en los dos bloques de referencia RefBlk1 y RefBlk2 usando el valor medio. Por otra parte, cuando el tipo de predicción PredType indica un procedimiento de predicción entre imágenes distinto de una predicción por interpolación de imágenes de referencia plurales, la memoria intermedia de imágenes múltiples 407 genera el bloque de referencia RefBlk que corresponde al número de imagen de referencia
15 NRefNo1 y el vector de movimiento NMV1.

20 Cuando el tipo de predicción PreType indica una predicción por interpolación de imágenes de referencia plurales, el conmutador 408 se cambia a un lado "0" y se usa un bloque interpolado RefPol como datos de imagen predictiva Pred. De este modo, el aparato de decodificación de imágenes en movimiento decodifica los datos de imágenes en movimiento codificados Str0 mediante el procesamiento descrito anteriormente y los genera como datos de imagen decodificados DImg.

25 Mientras tanto, bajo el procedimiento de codificación de imágenes en movimiento basado en el MPEG-4, se define un procedimiento de predicción por interpolación de imágenes de referencia plurales denominado "modo directo" para un tipo de imagen, denominada "imagen predictiva bidireccional", que emplea una predicción por interpolación de imágenes de referencia plurales. Se define como procedimiento para abreviar los vectores de movimiento y los números de imágenes de referencia incluidos en los datos codificados del bloque calculando los vectores de movimiento con respecto a dos imágenes de referencia usadas para la generación de la imagen predictiva por medio de interpolación
30 que usa los vectores de movimiento codificados.

35 La Fig. 5 es una ilustración para un caso de usar el modo directo definido en el MPEG-4. Aquí, una imagen Pic representa una imagen actual que ha de ser codificada, una imagen Ref1 representa una imagen de referencia cuyo momento de visualización es anterior al de la imagen actual Pic y una imagen Ref2 representa una imagen de referencia cuyo momento de visualización es posterior al de la imagen actual Pic mientras que un bloque Blk representa un bloque actual que ha de ser codificado y un bloque Blk0 representa un bloque cuya posición es la misma que la del bloque actual Blk en la imagen de referencia Ref2. Un vector de movimiento MV01 representa un vector de movimiento de referencia adelantada que usa la imagen Ref1 como imagen de referencia para codificar el bloque Blk0, un vector de movimiento MV1 representa un vector de movimiento del bloque actual con respecto a la imagen de referencia Ref1, un vector de movimiento MV2 representa un vector de movimiento del bloque actual con respecto a la imagen de referencia Ref2, un bloque RefBlk1 representa un bloque de referencia al que ha de hacer referencia el vector de movimiento MV1 y un bloque RefBlk2 representa un bloque de referencia al que ha de hacer referencia el vector de movimiento MV2.

45 En cuanto a las dos imágenes que han de usarse para referencia por el bloque actual Blk, la imagen Ref2 cuyo momento de visualización es posterior y es más cercano a la imagen actual se usa como imagen de referencia atrasada, mientras que la imagen Ref1, que ha sido usada para referencia por el bloque Blk0 en el momento de codificación, se usa como imagen de referencia adelantada.

50 Para el cálculo de los vectores de movimiento, se supone que cualquier movimiento es constante o no se encuentran movimientos al comparar las imágenes. Aquí, suponiendo que un valor diferencial entre el momento de visualización de la imagen actual Pic y el de la imagen de referencia Ref1 es TRD1, un valor diferencial entre el momento de visualización de la imagen de referencia Ref1 y el de la imagen de referencia Ref2 es TRD2, y un valor diferencial entre el momento de visualización de la imagen actual Pic y el de la imagen de referencia Ref2 es TRD3, los vectores de movimiento MV1 y MV2 que han de usarse para codificar el bloque actual pueden calcularse respectivamente
55 usando las siguientes ecuaciones:

$$MV1 = MV01 \times (TRD1/TRD2) \quad (\text{Ecuación A})$$

$$60 \quad MV2 = -MV01 \times (TRD3/TRD2) \quad (\text{Ecuación B})$$

65 Usando el procedimiento anterior, pueden determinarse las imágenes de referencia y los vectores de movimiento en el caso de usar un modo directo. El procesamiento en el caso de usar un modo directo como se describió anteriormente es ejecutado por la unidad de estimación de movimiento 301 mostrada en el diagrama de bloques que ilustra el aparato convencional de codificación de imágenes en movimiento de la Fig. 1. El procesamiento para el caso de usar un modo directo descrito anteriormente, realizado por el aparato de decodificación de imágenes en movimiento, es ejecutado por

ES 2 306 380 T3

la unidad de compensación de movimiento 602 mostrada en el diagrama de bloques que ilustra el aparato convencional de decodificación de imágenes en movimiento de la Fig. 4.

5 Cuando una imagen en movimiento, en la que el movimiento entre las imágenes es pequeño, está codificada entre imágenes, un error de predicción entre las imágenes se hace muy pequeño y la mayoría de los datos residuales codificados ERes se hacen "0" realizando procesamiento de codificación de imágenes como cuantificación. Un caso en el que todos los datos residuales codificados ERes que resultan de la predicción entre imágenes usando las imágenes de referencia y los vectores de movimiento del bloque actual es "0" en la codificación en la que los vectores de movimiento y las imágenes de referencia se determinan usando un procedimiento predeterminado sin codificarlos, como
10 en el caso de usar un modo directo como se describió anteriormente, se define como uno de los tipos de predicción PredType denominado "modo de salto". Al usar un modo de salto, sólo se transmite el tipo de predicción PredType que indica el modo de salto, por lo tanto, la codificación de un bloque requiere una cantidad de código muy pequeña. La eficiencia de codificación puede mejorarse más asignando a este modo de salto código de longitud variable que sea más corto que otros tipos de predicción, o mediante codificación de la longitud de la serie el número de bloques consecutivos usados para el modo de salto.
15

En el H.264 descrito anteriormente, se define el "modo de salto" como un caso en el que se supone que son "0" todos los datos residuales codificados equivalentes a un solo bloque obtenido por la predicción entre imágenes usando un modo directo. El siguiente procesamiento se realiza cuando un bloque es codificado usando un modo de salto por el
20 aparato de codificación de imágenes en movimiento mostrado en la Fig. 1. La unidad de estimación de movimiento 301 genera los números de imágenes de referencia RefNo1, RefNo2, los vectores de movimiento MV1, MV2, así como el tipo de predicción PredType que indica un modo de salto. La unidad de codificación de longitud variable 302 realiza codificación de longitud variable sólo para el tipo de predicción PredType y la genera como datos de imágenes en movimiento codificados Str0 mediante el procesamiento explicado anteriormente, cuando el tipo de predicción PredType indica un modo de salto. El siguiente procesamiento se realiza cuando los datos codificados del bloque codificado usando un modo de salto se introducen en el aparato de decodificación de imágenes en movimiento mostrado en la Fig. 4. La unidad de decodificación de longitud variable 601 realiza decodificación de longitud variable sobre el tipo de predicción PredType. Cuando el tipo de predicción PredType indica un modo de salto, la unidad de compensación de movimiento 602 genera, mediante el procesamiento utilizado en el caso de modo directo explicado anteriormente,
30 los números de imágenes de referencia NRefNo1, NRefNo2, los vectores de movimiento NMV1, NMV2, así como el tipo de predicción PredType que indica un modo de salto.

En el H.264 como se describió anteriormente, pueden seleccionarse imágenes de referencia arbitrarias de una pluralidad de imágenes codificadas independientemente del momento de visualización de la imagen actual. Sin embargo,
35 las imágenes de referencia arbitrarias se seleccionan realizando estimación de movimiento para la pluralidad de imágenes codificadas en este caso, por lo tanto, la carga de procesamiento causada por la estimación de movimiento se hace muy grande. La predicción por interpolación de imágenes de referencia plurales también conlleva un problema de degradación de la eficiencia de codificación ya que requiere codificación de números de imágenes de referencia y vectores de movimiento para cada dos imágenes de referencia.
40

Además, cuando se realiza predicción entre imágenes para una imagen usando como imagen de referencia una imagen cuyo momento de visualización es posterior al de la imagen actual, como en el caso de predicción bidireccional descrita en la técnica convencional, la imagen tiene que ser codificada en un orden diferente de un orden de visualización, lo cual causa un retardo. En un caso de comunicación en tiempo real como videoteléfono, no pueden
45 usarse imágenes predictivas bidireccionales debido al retardo. Sin embargo, en el H.264, pueden seleccionarse dos imágenes de referencia arbitrarias independientemente de la información de orden de visualización, por lo tanto, el retardo causado por la codificación puede eliminarse realizando una predicción por interpolación de imágenes de referencia plurales con una selección de dos imágenes que tienen respectivamente un momento de visualización que es anterior al de la imagen actual. Sin embargo, la imagen cuyo momento de visualización es posterior al de la imagen actual no es almacenada en la memoria intermedia de imágenes múltiples, por lo tanto, no puede emplearse el modo directo usado convencionalmente para determinar los vectores de movimiento usando la imagen cuyo momento de visualización es posterior al de la imagen actual como se describió anteriormente.
50

El documento "Working Draft N° 2, Revision 2 (WD-2)", JVT-B118r2, del 29 de enero de 2002, págs. 1-100 describe un procedimiento de codificación de referencia que ha de usarse para el desarrollo de una nueva Recomendación ITU-T (H.26L) y Estándar ISO (MPEG-4, Parte 10) de compresión de vídeo. La configuración básica del algoritmo es similar a H-263 y MPEG-4, Parte 2. En particular, se describe una configuración en la que se proporciona una lista de imágenes de referencia para cada franja y en la que además se proporciona para cada macrobloque la información que identifica las imágenes incluidas en la lista de imágenes de referencia.
60

El documento de Fukuhara y col. "Very Low Bit-Rate Video Coding with Block Partitioning and Adaptive Selection of Two Time-Differential Frame Memories", IEEE Transactions on Circuits and Systems for Video Technology, Vol. 7, N° 1, febrero de 1997, págs. 212-220, desvela esquemas de compensación de movimiento avanzada para codificación de vídeo de tasa de bits muy baja. Las características fundamentales de los esquemas de compensación de movimiento propuestos son predicción de partición de bloques y utilización de dos fotogramas de referencia diferencial en el tiempo.
65

El documento WO01/33864A1 desvela un procedimiento de codificación de una secuencia de imágenes, estando dividida cada imagen en bloques de píxeles no superpuestos.

5 El documento de Karczewicz y col. "A Proposal for SP-frames", VCEG-L27, ITU-Telecommunications Standardization Sector VCEG-L27, 9 de enero de 2001, págs. 1-9, desvela un nuevo tipo de imagen que hace uso de codificación predictiva compensada en movimiento para aprovecharse de la redundancia temporal en la secuencia permitiendo además reconstrucción idéntica del fotograma aun cuando se estén usando fotogramas de referencia diferentes.

Exposición de la invención

10 La presente invención está concebida en vista de las circunstancias anteriores, y se propone proporcionar un procedimiento de codificación de imágenes en movimiento y un *aparato de codificación* de imágenes en movimiento para llevar a cabo una codificación efectiva así como una reducción de la carga de procesamiento cuando se realiza una predicción por interpolación de imágenes de referencia plurales.

15 Para lograr los objetos anteriores, el procedimiento de codificación de imágenes en movimiento *se propone como se define en la reivindicación 1*.

20 *En la reivindicación 3 se define un aparato de codificación correspondiente.*

25 Cuando la imagen predictiva se genera usando una imagen de referencia, puede reducirse la carga de procesamiento ya que no se requiere el procesamiento para seleccionar una imagen que ha de usarse como imagen de referencia, basándose en bloque por bloque, de entre una pluralidad de imágenes codificadas. No es necesaria la codificación de esta imagen de referencia basándose en bloque por bloque. Por lo tanto, puede reducirse la cantidad de bits. En general, es muy posible que la mayoría de los bloques en los datos de imagen seleccionen la misma imagen como imagen de referencia óptima. Por lo tanto, es posible reducir la carga de procesamiento manteniendo mientras tanto una alta eficiencia de codificación compartiendo una imagen de referencia basándose en bloque por bloque, por ejemplo.

30 Cuando la imagen predictiva se genera con referencia a dos imágenes de referencia, puede reducirse la carga de procesamiento ya que no es necesario el procesamiento para seleccionar una sola imagen basándose en bloque por bloque como la única imagen de referencia de entre una pluralidad de imágenes codificadas. También puede reducirse la cantidad de bits ya que no es necesaria la codificación de esta imagen de referencia basándose en bloque por bloque. Por lo general, es muy posible que la mayoría de los bloques en los datos de imagen seleccionen la misma imagen como imagen de referencia óptima. Por lo tanto, es posible reducir la carga de procesamiento manteniendo mientras tanto una alta eficiencia de codificación compartiendo una de las imágenes de referencia basándose en bloque por bloque, por ejemplo.

35 El procedimiento de codificación de imágenes en movimiento comprende una etapa de descripción de información de describir información para especificar la imagen de referencia común en un área de información común asignada para la pluralidad de bloques en los datos de imágenes en movimiento codificados que han de generarse.

40 De ese modo, la información para especificar la imagen de referencia común puede describirse en los datos de imágenes en movimiento codificados y luego generarse, por lo tanto, la imagen de referencia puede especificarse con certeza cuando se decodifican los datos de imágenes en movimiento codificados.

45 La presente invención puede llevarse a cabo no sólo como el procedimiento de codificación de imágenes en movimiento y el procedimiento de decodificación de imágenes en movimiento tal como se describió anteriormente, sino también como un aparato de codificación de imágenes en movimiento y un aparato de decodificación de imágenes en movimiento que tiene etapas características incluidas en el procedimiento de codificación de imágenes en movimiento y el procedimiento de decodificación de imágenes en movimiento como unidades. También puede llevarse a cabo como un programa que tiene un ordenador que ejecuta estas etapas o como datos de imágenes en movimiento codificados que son codificados con el uso del procedimiento de codificación de imágenes en movimiento. Huelga decir que tal programa y tales datos de imágenes en movimiento codificados pueden ser distribuidos por medio de un medio de grabación como un CD-ROM y un medio de transmisión como una Internet.

Breve descripción de los dibujos

50 La Fig. 1 es un diagrama de bloques que muestra una estructura de un aparato convencional de codificación de imágenes en movimiento.

55 La Fig. 2 es un diagrama conceptual que muestra una predicción por interpolación que usa una pluralidad de imágenes de referencia.

60 La Fig. 3 es un diagrama conceptual que muestra un formato de datos de una imagen en movimiento codificada empleado por el aparato convencional de codificación de imágenes en movimiento.

65 La Fig. 4 es un diagrama de bloques que muestra una estructura de un aparato convencional de decodificación de imágenes en movimiento.

ES 2 306 380 T3

La Fig. 5 es una ilustración para un modo directo convencional.

La Fig. 6 es un diagrama de bloques que muestra una estructura de un aparato de codificación de imágenes en movimiento según una primera realización.

5

La Fig. 7 es un diagrama conceptual que muestra un formato de datos de una imagen en movimiento codificada según la primera realización.

La Fig. 8 es un diagrama de bloques que muestra una estructura de un aparato de decodificación de imágenes en movimiento según una segunda realización.

10

La Fig. 9 es un diagrama de bloques que muestra una estructura de un aparato de codificación de imágenes en movimiento según una tercera realización.

La Fig. 10 es un diagrama conceptual que muestra un formato de datos de una imagen en movimiento codificada según la tercera realización.

15

La Fig. 11 es un diagrama de bloques que muestra una estructura de una variación del aparato de codificación de imágenes en movimiento según la tercera realización.

20

La Fig. 12 es un diagrama conceptual que muestra un formato de datos de una imagen en movimiento codificada según la variación de la tercera realización.

La Fig. 13 es un diagrama de bloques que muestra una variación del aparato de codificación de imágenes en movimiento según la tercera realización.

25

La Fig. 14 es un diagrama de bloques que muestra una estructura de un aparato de decodificación de imágenes en movimiento según la cuarta realización.

La Fig. 15 es un diagrama de bloques que muestra una estructura de una variación del aparato de decodificación de imágenes en movimiento según la cuarta realización.

30

La Fig. 16 es una ilustración de un modo directo según una quinta realización, con el uso de imágenes de referencia plurales que tienen respectivamente información sobre un momento de visualización que es anterior al de una imagen actual.

35

La Fig. 17 es una ilustración de un modo directo según la quinta realización, con el uso de imágenes de referencia plurales que tienen respectivamente información sobre un momento de visualización que es posterior al de la imagen actual.

40

La Fig. 18 es una ilustración de una predicción entre imágenes que usa un modo de salto según una sexta realización.

Las Figs. 19A, 19B y 19C son ilustraciones de un medio de grabación para almacenar un programa para llevar a cabo el procedimiento de codificación de imágenes en movimiento o el procedimiento de decodificación de imágenes en movimiento según cada una de las realizaciones en un sistema informático. La Fig. 19A es una ilustración que muestra un ejemplo de un formato físico de un disco flexible que es un cuerpo principal de un medio de almacenamiento. La Fig. 19B es una ilustración que muestra un aspecto completo del disco flexible, una estructura en sección transversal y el propio disco flexible. La Fig. 19C es una ilustración que muestra una configuración para grabar y reproducir el programa en el disco flexible FD.

50

La Fig. 20 es un diagrama de bloques que muestra una configuración íntegra de un sistema de distribución de contenido para llevar a cabo un servicio de distribución de contenido.

La Fig. 21 es un croquis que muestra un ejemplo de un teléfono celular.

55

La Fig. 22 es un diagrama de bloques que muestra una estructura interna del teléfono celular.

La Fig. 23 es un diagrama de bloques que muestra una configuración íntegra de un sistema de radiodifusión digital.

60

Mejor modo de llevar a cabo la invención

Primera realización

65

La Fig. 6 es un diagrama de bloques que muestra una estructura de un aparato de codificación de imágenes en movimiento según la primera realización. Se ponen las mismas marcas para las unidades y los datos que funcionan de la misma manera que se describe en el diagrama de bloques que muestra una estructura de un aparato convencional de

ES 2 306 380 T3

codificación de imágenes en movimiento de la Fig. 1 y se abreviará la descripción. Es posible que el aparato de codificación de imágenes en movimiento y el aparato de decodificación de imágenes en movimiento según cada realización descrita más adelante cambie, basándose en bloque por bloque, entre los siguientes procedimientos de predicción: un procedimiento de generación de una imagen predictiva por interpolación de píxeles usando dos imágenes de referencia (una predicción por interpolación de imágenes de referencia plurales); un procedimiento de uso de un bloque incluido en una sola imagen arbitraria directamente como imagen predictiva; un procedimiento de generación de una imagen predictiva usando una predicción entre imágenes.

El aparato de codificación de imágenes en movimiento es un aparato para dividir en bloques unos datos de imagen introducidos *Img* y realizar codificación sobre cada uno de los bloques, e incluye una unidad de estimación de movimiento 101, la unidad de interpolación de píxeles 102, el restador 103, la unidad de codificación de imágenes 104, la unidad de decodificación de imágenes 105, el sumador 106, una unidad de codificación de longitud variable 107, la memoria intermedia de imágenes múltiples 108, y el conmutador 109.

Se introduce en el aparato de codificación de imágenes en movimiento un número de imagen de referencia por omisión *DefRefNo* que indica la una de las imágenes de referencia que ha de usarse para un bloque que es codificado usando una predicción por interpolación de imágenes de referencia plurales. La unidad de estimación de movimiento 101 realiza estimación de movimiento fijando una de dos imágenes de referencia como la imagen de referencia indicada por el número de imagen de referencia por omisión *DefRefNo* introducido, cuando se realiza una predicción por interpolación de imágenes de referencia plurales. El número de imagen de referencia *RefNo1* generado por la unidad de estimación de movimiento 101 indica, por lo tanto, el mismo valor que el indicado por el número de imagen de referencia por omisión *DefRefNo*. La unidad de codificación de longitud variable 107 realiza codificación de longitud variable para los datos residuales codificados *ERes*, el tipo de predicción *PredType*, el número de imagen de referencia *RefNo2*, los vectores de movimiento *MV1*, *MV2*, el número de imagen de referencia por omisión *DefRefNo* y los genera como datos de imágenes en movimiento codificados *Str*.

Lo siguiente describe una operación realizada por el aparato de codificación de imágenes en movimiento construido como el anterior, cuando el tipo de predicción del bloque actual es una predicción por interpolación de imágenes de referencia plurales.

Los datos de imagen introducidos *Img* se introducen en la unidad de estimación de movimiento 101 y el restador 103 basándose en bloque por bloque.

La unidad de estimación de movimiento 101 determina un tipo de predicción del bloque actual introducido y genera el tipo de predicción hacia el conmutador 109 y la unidad de codificación de longitud variable 107. Cuando el tipo de predicción *PredType* determinado es una predicción por interpolación de imágenes de referencia plurales, la unidad de estimación de movimiento 101 determina la una de las dos imágenes de referencia como la imagen de referencia indicada por el número de imagen de referencia por omisión *DefRefNo* introducido y determina respectivamente la otra imagen de referencia y los vectores de movimiento *MV1* y *MV2* con respecto a estas dos imágenes de referencia. La unidad de estimación de movimiento 101 genera luego un número de imagen de referencia *RefNo2* y los vectores de movimiento *MV1* y *MV2* hacia la memoria intermedia de imágenes múltiples 108 y la unidad de codificación de longitud variable 107 así como el número de imagen de referencia *RefNo1* para la memoria intermedia de imágenes múltiples 108. El número de imagen de referencia por omisión *DefRefNo* puede ser generado desde la estimación de movimiento 101 hacia la unidad de codificación de longitud variable 107.

A continuación, la memoria intermedia de imágenes múltiples 108 genera un bloque de referencia *RefBlk1* que corresponde al número de imagen de referencia *RefNo1* y el vector de movimiento *MV1* así como un bloque de referencia *RefBlk2* que corresponde al número de imagen de referencia *RefNo2* y el vector de movimiento *MV2* hacia la unidad de interpolación de píxeles 102. La interpolación de píxeles 102 interpola píxeles con respecto a los dos bloques de referencia *RefBlk1* y *RefBlk2* con el uso de un valor medio y lo genera como un bloque interpolado *RefPol*. Aquí, el tipo de predicción determinado por la unidad de estimación de movimiento 101 es una predicción por interpolación de imágenes de referencia plurales, el conmutador 109 se cambia a un lado "1" y el bloque interpolado *RefPol* se genera como datos de imagen predictiva *Pred* hacia el restador 103 y el sumador 106.

El restador 103 resta los datos de imagen predictiva *Pred* de los datos de imagen introducidos *Img* y los genera como datos residuales *Res* hacia la unidad de codificación de imágenes 104. La unidad de codificación de imágenes 104 realiza procesamiento de codificación de imágenes, como transformación ortogonal y cuantificación, sobre los datos residuales introducidos *Res* y los genera como datos residuales codificados *ERes* hacia la unidad de decodificación de imágenes 105 y la unidad de codificación de longitud variable 107. La unidad de decodificación de imágenes 105 realiza procesamiento de decodificación de imágenes, como cuantificación inversa y transformación ortogonal inversa, sobre los datos residuales codificados *ERes* introducidos y los genera como datos residuales decodificados *DRes* hacia el sumador 106. El sumador 106 suma los datos residuales decodificados *DRes* a los datos de imagen predictiva *Pred* y los genera como datos de imagen reconstruidos *Recon*. Los datos que tienen la posibilidad de ser usados para referencia en la predicción entre imágenes posterior de los datos reconstruidos *Recon* se almacenan en la memoria intermedia de imágenes múltiples 108.

La unidad de codificación de longitud variable 107 realiza codificación de longitud variable para los datos residuales codificados *ERes* introducidos, el tipo de predicción *PredType*, el número de imagen de referencia *RefNo2*

ES 2 306 380 T3

y los vectores de movimiento MV1, MV2 para cada bloque y los genera como datos de imágenes en movimiento codificados Str.

Para la imagen indicada por el número de imagen de referencia por omisión DefRefNo, puede seleccionarse una imagen arbitraria de las imágenes almacenadas en la memoria intermedia de imágenes múltiples 108. Por ejemplo, son imaginables una imagen codificada que tiene información de orden de visualización que es el más cercano al de la imagen actual, una imagen codificada que tiene información de orden de visualización que es anterior y el más cercano al de la imagen actual, una imagen codificada que tiene información de orden de visualización que es posterior y el más cercano al de la imagen actual, y similares, en la memoria intermedia de imágenes múltiples 108.

Igualmente, también son imaginables una imagen que es la más cercana a la imagen actual en un orden de codificación, una imagen que tiene información de orden de visualización que es anterior al de la imagen actual y un orden de codificación que es el más cercano al de la imagen actual, una imagen que tiene información de orden de visualización que es anterior al de la imagen actual y un orden de codificación que es el más cercano al de la imagen actual, una imagen que tiene información de orden de visualización que es posterior al de la imagen actual y un orden de codificación que es el más cercano al de la imagen actual, y similares.

La Fig. 7 es un diagrama conceptual que muestra un formato de datos de una imagen en movimiento codificada según la primera realización. Se ponen las mismas marcas para los mismos datos que los descritos en el diagrama conceptual que muestra el formato de datos de la imagen en movimiento codificada empleado por el aparato convencional de codificación de imágenes en movimiento mostrado en la Fig. 3, y se abreviará la descripción. La diferencia entre el formato de datos de la imagen en movimiento codificada según la presente realización y el empleado por el aparato convencional de codificación de imágenes en movimiento es que para cada imagen está incluido un número de imagen de referencia por omisión DefRefNo y que sólo se incluye un único dato para el número de imagen de referencia en los datos codificados del bloque que es codificado usando una predicción por interpolación de imágenes de referencia plurales.

Según la presente realización tal como se describió anteriormente, puede mejorarse la eficiencia de codificación ya que el número de imagen de referencia fijo no tiene que ser codificado basándose en bloque por bloque. Esto es porque se selecciona una imagen arbitraria para la una imagen de referencia basándose en bloque por bloque de entre imágenes codificadas plurales, y la otra imagen de referencia puede fijarse como una imagen entre imágenes codificadas plurales basándose en imagen por imagen.

En la presente realización, como procedimiento de especificación de una imagen de referencia por omisión, se asignan los números de imagen a las imágenes, sin embargo, la presente invención no estará limitada a esto. Por ejemplo, es posible especificar una imagen de referencia por omisión usando un valor diferencial relativo entre el número de imagen poseído por la imagen actual y el número de imagen poseído por la imagen seleccionada como imagen de referencia por omisión, o usando información como un comando que indica una imagen de referencia por omisión.

En la presente realización, sólo se especifica una imagen de referencia como imagen de referencia por omisión, sin embargo, los dos números de imágenes de referencia en los datos codificados del bloque pueden abreviarse codificando dos números de imágenes de referencia por omisión.

Además, en la presente realización, se proporciona la descripción para la predicción por interpolación de imágenes de referencia plurales para generar una imagen predictiva por interpolación de píxeles usando dos imágenes de referencia. Sin embargo, puede tratarse de la misma manera un caso de predicción por interpolación de imagen de referencia única usando un bloque incluido en una sola imagen de referencia arbitraria como imagen predictiva. En este caso, no hay necesidad de describir la información de imagen de referencia para cada bloque y, por lo tanto, la información de imagen de referencia se describe sólo en un área de información común.

Los números de imágenes de referencia por omisión son codificados basándose en imagen por imagen en la presente realización, sin embargo, pueden ser codificados usando una estructura de sintaxis que almacena un solo número de imagen de referencia por omisión para todas las imágenes plurales o pueden ser codificados usando una estructura de sintaxis que almacena un solo número de imagen de referencia por omisión para una estructura de sintaxis inferior a una imagen como un macrobloque que está compuesto de bloques plurales o una franja que está formada por macrobloques plurales, o similar.

60 Segunda realización

La Fig. 8 es un diagrama de bloques que muestra un aparato de decodificación de imágenes en movimiento según la segunda realización de la presente invención. Se ponen las mismas marcas para las unidades y los datos que funcionan de la misma manera que se ilustra en el diagrama de bloques que muestra una estructura del aparato convencional de decodificación de imágenes en movimiento de la Fig. 4, y se abreviará la descripción. La diferencia entre el aparato de decodificación de imágenes en movimiento de la presente realización y el convencional mostrado en la Fig. 4 es que al anterior se añade una memoria intermedia de números de imágenes de referencia por omisión 402.

ES 2 306 380 T3

El aparato de decodificación de imágenes en movimiento, como se muestra en la Fig. 8, incluye una unidad de decodificación de longitud variable 401, una memoria intermedia de números de imágenes de referencia por omisión 402, una unidad de compensación de movimiento 403, una unidad de decodificación de imágenes 404, un sumador 405, una unidad de interpolación de píxeles 406, una memoria intermedia de imágenes múltiples 407 y un conmutador 408.

La unidad de decodificación de longitud variable 401 realiza decodificación de longitud variable sobre los datos de imágenes en movimiento codificados Str introducidos y genera datos residuales codificados ERes, un tipo de predicción PredType, un número de imagen de referencia RefNo2, vectores de movimiento MV1 y MV2, un número de imagen de referencia por omisión DefRefNo. El número de imagen de referencia por omisión DefRefNo decodificado tiene que ser compartido entre bloques plurales de manera que se almacena en la memoria intermedia de números de imágenes de referencia por omisión 402. El número de imagen de referencia por omisión DefRefNo almacenado en la memoria intermedia de números de imágenes de referencia por omisión 402 se introduce como número de imagen de referencia RefNo1 en la unidad de compensación de movimiento 403.

Lo siguiente describe una operación del aparato de decodificación de imágenes en movimiento construido como antes cuando un tipo de predicción de un bloque actual que ha de ser decodificado es una predicción por interpolación de imágenes de referencia plurales.

Los datos de imágenes en movimiento codificados Str se introducen en la unidad de decodificación de longitud variable 401. La unidad de decodificación de longitud variable 401 realiza decodificación de longitud variable sobre los datos de imágenes en movimiento codificados Str introducidos y genera respectivamente lo siguiente: los datos residuales codificados ERes hacia la unidad de decodificación de imágenes 404; el número de imagen de referencia RefNo2 y los vectores de movimiento MV1, MV2 hacia la unidad de compensación de movimiento 403; el tipo de predicción PredType hacia la unidad de compensación de movimiento 403 y el conmutador 408; y el número de imagen de referencia por omisión DefRefNo hacia la memoria intermedia de números de imágenes de referencia por omisión 402. La memoria intermedia de números de imágenes de referencia por omisión 402 genera el número de imagen de referencia por omisión DefRefNo almacenado como número de imagen de referencia RefNo1 hacia la unidad de compensación de movimiento 403.

Como el tipo de predicción PredType es una predicción por interpolación de imágenes de referencia plurales, la unidad de compensación de movimiento 403 genera, hacia la memoria intermedia de imágenes múltiples 407, el número de imagen de referencia NRefNo1 introducido por la memoria intermedia de números de imágenes de referencia por omisión 402 y el número de imagen de referencia RefNo2 y los vectores de movimiento MV1, MV2 introducidos por la unidad de decodificación de longitud variable 401 y ordena una salida de los bloques de referencia. La memoria intermedia de imágenes múltiples 407 genera, hacia la unidad de interpolación de píxeles 406, el bloque de referencia RefBlk1 que corresponde al número de imagen de referencia NRefNo1 y el vector de movimiento NMV1 y el bloque de referencia RefBlk2 que corresponde al número de imagen de referencia NRefNo2 y el vector de movimiento NMV2. La unidad de interpolación de píxeles 406 interpola los valores de píxeles con respecto a los dos bloques de referencia RefBlk1 y RefBlk2 usando el valor medio y lo genera como un bloque interpolado RefPol. Aquí, como el tipo de predicción es una predicción por interpolación de imágenes de referencia plurales, el conmutador 408 se cambia a un lado "0" y el bloque interpolado RefPol se genera así como datos de imagen predictiva Pred hacia el sumador 405.

Por otra parte, la unidad de decodificación de imágenes 404 en la que se introducen los datos residuales codificados ERes realiza procesamiento de decodificación de imágenes como cuantificación inversa y transformación ortogonal inversa y genera datos residuales decodificados DRes hacia el sumador 405. El sumador 405 suma los datos residuales decodificados DRes a los datos de imagen predictiva Pred y los genera como datos de imagen decodificados DImg fuera del aparato de decodificación de imágenes en movimiento. La memoria intermedia de imágenes múltiples 407 almacena los datos de imagen decodificados DImg para predicción entre imágenes. El aparato de decodificación de imágenes en movimiento decodifica los datos de imágenes en movimiento codificados Str mediante tal procesamiento y los genera como los datos de imagen decodificados DImg.

Según la presente realización tal como se describió anteriormente, es posible decodificar correctamente los datos de imágenes en movimiento codificados Str que son codificados por el aparato de codificación de imágenes en movimiento que usa el procedimiento de codificación de imágenes en movimiento de la presente invención descrito en la primera realización.

60 Tercera realización

La Fig. 9 es un diagrama de bloques que muestra un aparato de codificación de imágenes en movimiento según la tercera realización de la presente invención. Se ponen las mismas marcas para las unidades y los datos que funcionan de la misma manera que se muestra en el diagrama de bloques que ilustra el aparato de codificación de imágenes en movimiento según la primera realización en la Fig. 6, y se abreviará la descripción.

El aparato de codificación de imágenes en movimiento de la presente realización incluye una unidad de generación de números de imágenes de referencia por omisión 201 además de la estructura mostrada en la primera realización.

ES 2 306 380 T3

La unidad de generación de números de imágenes de referencia por omisión 201 genera un número de imagen de referencia por omisión DefRefNo usando un procedimiento predeterminado y lo genera hacia la unidad de estimación de movimiento 101. La unidad de estimación de movimiento 101 realiza estimación de movimiento, fijando una de dos imágenes de referencia como la imagen de referencia indicada por el número de imagen de referencia por omisión DefRefNo introducido, cuando se realiza la predicción por interpolación de imágenes de referencia plurales como en el caso del aparato de codificación de imágenes en movimiento según la primera realización. La unidad de codificación de longitud variable 202 realiza codificación de longitud variable sobre los datos residuales codificados ERes, el tipo de predicción PredType, el número de imagen de referencia RefNo2 y los vectores de movimiento MV1, MV2 y los genera como datos de imágenes en movimiento codificados Str2.

Por ejemplo, se dispone de los siguientes procedimientos como procedimiento de generación de un número de imagen de referencia por omisión DefRefNo empleado por la unidad de generación de números de imágenes de referencia por omisión 201. El primer procedimiento es determinar, como número de imagen de referencia por omisión DefRefNo, un número de imagen que indica una imagen que tiene información de orden de visualización que es el más cercano al de la imagen actual de las imágenes codificadas almacenadas en la memoria intermedia de imágenes múltiples 108. El segundo procedimiento es determinar, como número de imagen de referencia por omisión DefRefNo, un número de imagen que indica una imagen que tiene información de orden de visualización que es anterior y el más cercano al de la imagen actual de las imágenes codificadas almacenadas en la memoria intermedia de fotografías múltiples 108. El tercer procedimiento es determinar, como número de imagen de referencia por omisión DefRefNo, un número de imagen que indica una imagen que tiene información de orden de visualización que es posterior y el más cercano al de la imagen actual de las imágenes codificadas almacenadas en la memoria intermedia de imágenes múltiples 108. El cuarto procedimiento es determinar, como número de imagen de referencia por omisión DefRefNo, un número de imagen que indica una imagen cuyo orden de codificación es el más cercano al de la imagen actual de las imágenes codificadas almacenadas en la memoria intermedia de imágenes múltiples 108. El quinto procedimiento es determinar, como número de imagen de referencia por omisión DefRefNo, un número de imagen que indica una imagen que tiene información de orden de visualización anterior al de la imagen actual y cuyo orden de codificación es el más cercano al de la imagen actual de las imágenes codificadas almacenadas en la memoria intermedia de imágenes múltiples 108. El sexto procedimiento es determinar, como número de imagen de referencia por omisión DefRefNo, un número de imagen que indica una imagen que tiene información de orden de visualización posterior al de la imagen actual y cuyo orden de codificación es el más cercano al de la imagen actual de las imágenes codificadas almacenadas en la memoria intermedia de imágenes múltiples 108.

El formato de datos de la imagen en movimiento codificada usado por el aparato de codificación de imágenes en movimiento según la presente realización es como se muestra en la Fig. 10, de la cual se omite el número de imagen de referencia por omisión DefRefNo mostrado en el formato de datos de la imagen en movimiento codificada mostrada en la Fig. 7. Por lo tanto, no tiene que codificarse un número de imagen de referencia por omisión DefRefNo, lo cual mejora la eficiencia de codificación.

En la realización mencionada anteriormente, se explica un procedimiento para llevar a cabo la codificación sin describir en absoluto información sobre la imagen de referencia por omisión en un formato de datos fijando un procedimiento a uno arbitrario para determinar una imagen de referencia por omisión. Sin embargo, es posible cambiar entre los procedimientos para determinar una imagen de referencia por omisión basándose en imagen por imagen. Por ejemplo, esto puede llevarse a cabo codificando cualquiera de los siguientes identificadores: un identificador que indica un procedimiento de selección, como imagen de referencia por omisión, de una imagen que tiene información de momento de visualización que es el más cercano al de la imagen actual de las imágenes codificadas almacenadas en la memoria intermedia de imágenes múltiples; un identificador que indica un procedimiento de selección, como imagen de referencia por omisión, de una imagen que tiene información de momento de visualización que es anterior y el más cercano al de la imagen actual de las imágenes codificadas almacenadas en la memoria intermedia de imágenes múltiples; y un identificador que indica un procedimiento de selección, como imagen de referencia por omisión, de una imagen que tiene información sobre un momento de visualización que es posterior y el más cercano al de la imagen actual de las imágenes codificadas almacenadas en la memoria intermedia de imágenes múltiples.

La Fig. 11 es un diagrama de bloques que muestra el aparato de codificación de imágenes en movimiento usado para este caso. La unidad de generación de números de imágenes de referencia por omisión 203 genera un identificador Ident que indica un procedimiento de selección de una imagen de referencia por omisión hacia la unidad de codificación de longitud variable 204, como se muestra en la Fig. 11. La unidad de codificación de longitud variable 204 realiza codificación de longitud variable sobre los datos residuales codificados ERes, el tipo de predicción PredType, la imagen de referencia RefNo2, los vectores de movimiento MV1 y MV2 así como el identificador Ident y los genera como datos de imágenes en movimiento codificados Str3. El formato de datos para este caso incluye un identificador Ident para indicar un procedimiento de selección de una imagen de referencia por omisión como se muestra en la Fig. 12 en lugar del número de imagen de referencia por omisión DefRefNo, es decir información que especifica directamente una imagen de referencia por omisión como se muestra en la Fig. 7.

Igualmente, es posible codificar un identificador que indica un procedimiento de selección, como imagen de referencia por omisión, una imagen cuyo orden de codificación es el más cercano al de la imagen actual de las imágenes codificadas almacenadas en la memoria intermedia de imágenes múltiples, un identificador que indica un procedimiento de selección, como imagen de referencia por omisión, de una imagen que tiene información de momento de visualización anterior al de la imagen actual y cuyo orden de codificación es el más cercano al de la imagen actual

ES 2 306 380 T3

de las imágenes codificadas almacenadas en la memoria intermedia de imágenes múltiples o un identificador que indica un procedimiento de selección, como imagen de referencia por omisión, de una imagen que tiene información de momento de visualización posterior al de la imagen actual y cuyo orden de codificación es el más cercano a la imagen actual. Los datos de imágenes en movimiento codificados que se generan usando este procedimiento pueden ser decodificados usando el procedimiento de decodificación, que tiene una estructura según la cuarta realización, que se explicará más adelante.

También es posible codificar el propio número de imagen DefRefNo que indica una imagen de referencia por omisión, como en la Fig. 7, para codificar un valor diferencial entre el número de imagen de la imagen actual y el número de imagen de la imagen seleccionada como imagen de referencia por omisión, o codificar información como un comando para indicar una imagen de referencia por omisión.

La Fig. 13 es un diagrama de bloques que muestra un aparato de codificación de imágenes en movimiento usado para tal caso. La unidad de generación de números de imágenes de referencia por omisión 205 genera el número de imagen de referencia por omisión DefRefNo hacia la unidad de codificación de longitud variable 206, como se muestra en la Fig. 13. La unidad de codificación de longitud variable 206 realiza codificación de longitud variable sobre los datos residuales codificados ERes, el tipo de predicción PredType, el número de imagen de referencia RefNo2, los vectores de movimiento MV1, MV2, así como el número de imagen de referencia por omisión DefRefNo y los genera como datos de imágenes en movimiento codificados Str4. El formato de datos para este caso es el mismo que el mostrado en la Fig. 7. Los datos de imágenes en movimiento codificados generados usando este procedimiento pueden ser decodificados con el uso del procedimiento de decodificación que tiene la estructura descrita en la segunda realización.

25 Cuarta realización

La Fig. 14 es un diagrama de bloques que muestra un aparato de decodificación de imágenes en movimiento según la cuarta realización de la presente invención. Se ponen las mismas marcas para las unidades y los datos que funcionan de la misma manera que se muestra en el diagrama de bloques para el aparato de decodificación de imágenes en movimiento según la segunda realización en la Fig. 8, y se abreviará la descripción.

El aparato de decodificación de imágenes en movimiento según la presente realización incluye una unidad de generación de números de imágenes de referencia por omisión 502 en lugar de la memoria intermedia de números de imágenes de referencia por omisión 402 mostrada en la estructura de la segunda realización. La unidad de decodificación de longitud variable 501 realiza decodificación de longitud variable sobre los datos de imágenes en movimiento codificados Str2 introducidos y genera los datos residuales codificados ERes, el tipo de predicción PredType, el número de imagen de referencia RefNo2, y los vectores de movimiento MV1, MV2. La unidad de generación de números de imágenes de referencia por omisión 502 genera un número de imagen de referencia por omisión DefRefNo de la misma manera que la unidad de generación de números de imágenes de referencia por omisión 201 descrita en la tercera realización y genera, hacia la unidad de compensación de movimiento 403, el número de imagen de referencia por omisión DefRefNo como número de imagen de referencia RefNo1.

Según la presente realización tal como se describió anteriormente, es posible decodificar correctamente los datos de imágenes en movimiento codificados Str2 que son codificados por el aparato de codificación de imágenes en movimiento que usa el procedimiento de codificación de imágenes en movimiento según la presente invención descrito en la tercera realización.

El aparato de decodificación de imágenes en movimiento está construido como se indica más adelante cuando se decodifican los datos de imágenes en movimiento codificados Str3 en los que está incluido el identificador Ident para indicar un procedimiento de selección de una imagen de referencia por omisión, como se ilustra en la variación de la tercera realización descrita anteriormente.

La Fig. 15 es un diagrama de bloques que muestra el aparato de decodificación de imágenes en movimiento usado para este caso. La unidad de decodificación de longitud variable 503 realiza decodificación de longitud variable sobre los datos de imágenes en movimiento codificados Str3 introducidos y genera los datos residuales codificados ERes, el tipo de predicción PredType, el número de imagen de referencia RefNo2, los vectores de movimiento MV1, MV2, así como el identificador Ident para indicar un procedimiento de selección de una imagen de referencia por omisión, como se muestra en la Fig. 15. La unidad de generación de números de imágenes de referencia por omisión 504 genera un número de imagen de referencia por omisión DefRefNo usando el procedimiento de selección de la imagen de referencia por omisión, indicado por el identificador introducido desde la unidad de decodificación de longitud variable 503, y genera, hacia la unidad de compensación de movimiento 403, el número de imagen de referencia por omisión DefRefNo como número de imagen de referencia RefNo1.

De este modo, es posible decodificar correctamente los datos de imágenes en movimiento codificados Str3, en los que se incluye el identificador Ident para identificar un procedimiento de selección de una imagen de referencia por omisión, como se describió anteriormente en la tercera realización.

Quinta realización

La presente realización describe codificación usando un modo directo cuando se realiza codificación sólo con referencia a las imágenes, cada una de las cuales tiene información de orden de visualización que es anterior al de la imagen actual.

La Fig. 16 es un diagrama que ilustra un modo directo que usa imágenes de referencia plurales, cada una de las cuales tiene información de orden de visualización anterior al de la imagen actual, según la quinta realización de la presente invención. Aquí, una imagen Pic representa una imagen actual que ha de ser codificada, las imágenes de referencia Ref1 y Ref2 representan imágenes de referencia, un bloque Blk representa un bloque actual que ha de ser codificado y un bloque Blk0 representa un bloque en la imagen de referencia Ref1, que es contiguo con el bloque actual Blk. Un vector de movimiento MV01 representa una imagen de referencia adelantada que se usa para codificar el bloque Blk0, una imagen Ref3 representa una imagen de referencia usada por el vector de movimiento MV01, un vector de movimiento MV1 representa un vector de movimiento con respecto a la imagen de referencia Ref1, un vector de movimiento MV2 representa un vector de movimiento con respecto a la imagen de referencia Ref2, un bloque RefBlk1 representa un bloque de referencia al que hace referencia el vector de movimiento MV1, y un bloque RefBlk2 representa un bloque de referencia al que hace referencia el vector de movimiento MV2.

Para las imágenes de referencia, por ejemplo, se seleccionan de las imágenes codificadas almacenadas en la memoria intermedia de imágenes múltiples imágenes que tienen respectivamente información de orden de visualización que es anterior y es el más cercano y el segundo más cercano al de la imagen actual. En este caso, suponiendo que TRD1 representa un valor diferencial entre la información de orden de visualización de la imagen actual Pic y la de la imagen de referencia Ref1, TRD2 representa un valor diferencial entre la información de orden de visualización de la imagen de referencia Ref1 y la de la imagen de referencia Ref3 y TRD3 representa un valor diferencial entre la información de orden de visualización de la imagen actual Pic y la de la imagen de referencia Ref2, los vectores de movimiento MV1 y MV2 que han de usarse para codificar el bloque actual pueden calcularse usando las siguientes ecuaciones:

$$MV1 = MV01 \times (TRD1/TRD2) \quad \dots(\text{Ecuación A})$$

$$MV2 = MV01 \times (TRD3/TRD2) \quad \dots(\text{Ecuación B})$$

Usando el procedimiento descrito anteriormente, pueden determinarse las imágenes de referencia y los vectores de movimiento en el caso de usar un modo directo.

En el H.264 descrito anteriormente, se trata de un procedimiento para controlar explícitamente imágenes que han de ser almacenadas en la memoria intermedia de imágenes múltiples incluyendo información de control para almacenar y quitar las imágenes codificadas en y de la memoria intermedia de imágenes múltiples en los datos de imágenes en movimiento codificados. Bajo tal control, podría haber un caso en el que sólo se almacenan en la memoria intermedia de imágenes múltiples las imágenes que tienen información de orden de visualización posterior al de la imagen actual. Lo siguiente describe un procedimiento para llevar a cabo un modo directo para una imagen a la que se aplica una predicción por interpolación de imágenes de referencia plurales, cuando sólo se almacenan en la memoria intermedia de imágenes múltiples las imágenes que tienen información de orden de visualización que es posterior al de la imagen actual.

La Fig. 17 es una ilustración que muestra un modo directo que usa imágenes de referencia plurales que tienen respectivamente información de orden de visualización que es posterior al de la imagen actual, según la quinta realización de la presente invención. Aquí, una imagen Pic representa una imagen actual que ha de ser codificada, las imágenes Ref1 y Ref2 representan imágenes de referencia, un bloque Blk representa un bloque actual que ha de ser codificado, un bloque Blk0 representa un bloque en la imagen de referencia Ref1, que es contiguo con el bloque actual Blk. Un vector de movimiento MV01 representa un vector de movimiento de referencia adelantada usado para codificar el bloque Blk0, un vector de movimiento MV1 representa un vector de movimiento con respecto a la imagen de referencia Ref1 y un vector de movimiento MV2 representa un vector de movimiento con respecto a la imagen de referencia Ref2, mientras que un bloque RefBlk1 representa un bloque de referencia al que hace referencia el vector de movimiento MV1 y un bloque RefBlk2 representa un bloque de referencia al que hace referencia el vector de movimiento MV2.

Para las imágenes de referencia, por ejemplo, una imagen que tiene información de orden de visualización que es posterior y es el más cercano y el segundo más cercano al de la imagen actual, se seleccionan de las imágenes codificadas almacenadas en la memoria intermedia de imágenes múltiples. En este caso, suponiendo que TRD1 representa un valor diferencial entre la información de orden de visualización de la imagen actual Pic y la de la imagen de referencia Ref1, TRD2 representa un valor diferencial entre la información de orden de visualización de la imagen de referencia Ref1 y la de la imagen de referencia Ref3 y TRD3 representa un valor diferencial entre la información de orden de visualización de la imagen actual Pic y la de la imagen de referencia Ref2, los vectores de movimiento MV1 y MV2 que han de usarse para codificar el bloque actual pueden calcularse usando las siguientes ecuaciones (Ecuación C) y (Ecuación D):

ES 2 306 380 T3

$$MV1 = -MV01 \times (TRD1/TRD2) \quad \dots(\text{Ecuación C})$$

$$MV2 = -MV01 \times (TRD3/TRD2) \quad \dots(\text{Ecuación D})$$

5

Usando el procedimiento descrito anteriormente, pueden determinarse las imágenes de referencia y los vectores de movimiento en el caso de usar un modo directo.

10 El procesamiento del modo directo como el descrito anteriormente realizado por el aparato de codificación de imágenes en movimiento mostrado en la Fig. 6 se ejecuta por la unidad de estimación de movimiento 101. Igualmente, el realizado por el aparato de decodificación de imágenes en movimiento mostrado en la Fig. 8 se ejecuta por la unidad de compensación de movimiento 403.

15 De este modo, el aparato de codificación de imágenes en movimiento utilizable para el modo directo como se describe en la presente realización permite el uso del modo directo aun cuando la memoria intermedia de imágenes múltiples sólo almacene las imágenes codificadas que tienen información de orden de visualización que es anterior o posterior al de la imagen actual y, por lo tanto, puede mejorar la eficiencia de codificación ya que pueden omitirse las imágenes de referencia y los vectores de movimiento. El aparato de decodificación de imágenes en movimiento
20 utilizable para el modo directo descrito en la presente realización puede decodificar los datos de imágenes en movimiento codificados generados por el aparato de codificación de imágenes en movimiento utilizable para el modo directo descrito en la presente realización.

25 Puede definirse un modo de salto como un caso en el que los datos residuales codificados ERes obtenidos por la predicción entre imágenes usando las imágenes de referencia y los vectores de movimiento calculados usando un modo directo según la presente realización son "0". El modo directo según la presente realización permite el uso del modo directo aun cuando la memoria intermedia de imágenes múltiples sólo tenga las imágenes codificadas que tienen información de orden de visualización que es anterior o posterior al de la imagen actual, por lo tanto, puede seleccionarse un modo de salto para tal caso. El aparato de decodificación de imágenes en movimiento utilizable para
30 el modo de salto descrito anteriormente permite el uso del modo de salto y por lo tanto puede mejorar la eficiencia de codificación. El aparato de decodificación de imágenes en movimiento utilizable para el modo directo descrito en la presente realización puede decodificar los datos de imágenes en movimiento codificados generados por el aparato de codificación de imágenes en movimiento utilizable para el modo directo descrito en la presente realización.

35 En la descripción anterior para las Figs. 16 y 17, puede seleccionarse libremente un vector de movimiento con respecto a la imagen de referencia Ref1 y también puede codificarse un vector diferencial entre el vector de movimiento y el vector de movimiento MV2 descritos anteriormente. Igualmente, puede seleccionarse libremente un vector de movimiento con respecto a la imagen de referencia Ref2 y también puede codificarse un vector diferencial entre el vector de movimiento y el vector de movimiento MV2 descritos anteriormente.

40

En la presente realización, el modo de salto descrito en la presente realización se usa cuando la memoria intermedia de imágenes múltiples sólo tiene las imágenes codificadas que tienen información de orden de visualización que es anterior o posterior al de la imagen actual. Sin embargo, una imagen que tiene información de orden de visualización que es el más cercano y el segundo más cercano al de la imagen actual puede seleccionarse de las imágenes almacenadas en la memoria intermedia de imágenes múltiples. El procedimiento puede modificarse de manera que el modo de salto descrito en la presente realización se adapte al caso en el que las dos imágenes seleccionadas son las imágenes que tienen información de orden de visualización que es anterior o posterior al de la imagen actual.

Sexta realización

50

En el H.264 como se describió anteriormente, un modo de salto para una imagen a la que se aplica una predicción por interpolación de imágenes de referencia plurales indica que los datos residuales codificados que resultan de la predicción entre imágenes usando un modo directo es "0". En contraste, el aparato de codificación de imágenes en movimiento y el aparato de decodificación de imágenes en movimiento de la presente invención emplean, como
55 procedimiento de predicción que ha de usarse para un modo de salto, una predicción entre imágenes que usa una imagen de referencia que tiene información de orden de visualización que es el más cercano al de la imagen actual de las imágenes codificadas de la memoria intermedia de imágenes múltiples.

60 La Fig. 18 es una ilustración que muestra la predicción entre imágenes en el caso de usar el modo de salto según la sexta realización de la presente invención. Aquí, un imagen Pic representa una imagen actual que ha de ser codificada, una imagen Ref1 representa una imagen codificada que tiene información de orden de visualización inmediatamente anterior al de la imagen actual, una imagen Ref2 es una imagen codificada que tiene información de orden de visualización inmediatamente posterior al de la imagen actual, un bloque Blk es un bloque actual que ha de ser codificado, un vector de movimiento MV1 representa un vector de movimiento que indica valor "0" con respecto a la imagen ref1 y un bloque RefBlk1 representa un bloque de referencia al que hace referencia el vector de movimiento MV1. También,
65 un valor diferencial entre la información de orden de visualización de la imagen actual Pic y la de la imagen Ref1, será menor que TRD2, un valor diferencial entre la información de orden de visualización de la imagen actual Pic y la de la imagen Ref2.

ES 2 306 380 T3

En la presente realización, se usa como imagen de referencia una imagen que tiene información de orden de visualización que es el más cercano al de la imagen actual. En la Fig. 18, una imagen que tiene información de orden de visualización que es el más cercano al de la imagen actual es una imagen Ref1. El vector de movimiento MV1 con respecto a la imagen Ref1 indica "0" tanto en las componentes vertical como horizontal dentro de la imagen y usa el bloque de referencia RefBlk1, al que hace referencia el vector de movimiento MV1, como imagen predictiva. Usando tal procedimiento de predicción, las imágenes de referencia y los vectores de movimiento se determinan de manera unívoca por el aparato de codificación de imágenes en movimiento y el aparato de decodificación de imágenes en movimiento, por lo tanto, no hay necesidad de incluir la información que indica imágenes de referencia así como vectores de movimiento en los datos de imágenes en movimiento codificados. Con la definición del modo de salto como el caso en el que los datos residuales codificados obtenidos como resultado de la predicción entre imágenes descrita anteriormente son "0", sólo se puede incluir en los datos codificados el tipo de predicción que indica un modo de salto para el bloque al que se aplica un modo de salto y después se transmite.

En la presente realización, se determina como imagen de referencia, de las imágenes codificadas almacenadas en la memoria intermedia de imágenes múltiples, una imagen que tiene información de orden de visualización que es el más cercano al de la imagen actual. Sin embargo, puede determinarse como imagen de referencia, de las imágenes codificadas de la memoria intermedia de imágenes múltiples, una imagen que tiene información de orden de visualización que es anterior y el más cercano al de la imagen actual.

Además, en la presente realización, se determina como imagen de referencia, de las imágenes codificadas almacenadas en la memoria intermedia de imágenes múltiples, una imagen que tiene información de orden de visualización que es el más cercano al de la imagen actual. Sin embargo, también puede determinarse como imagen de referencia, de las imágenes codificadas de la memoria intermedia de imágenes múltiples, una imagen que tiene información de orden de visualización que es posterior y el más cercano al de la imagen actual.

La información de orden de visualización de las imágenes usadas en cada una de las realizaciones anteriores puede ser un valor que indica el momento para visualizar las imágenes o información que indica una relación relativa en el orden de visualización de las imágenes.

La imagen mencionada anteriormente significa tanto un fotograma como un campo: se usa un fotograma para codificación de fotograma mientras que se usa un campo para codificación entrelazada (codificación de campo).

En cada una de las realizaciones anteriores, puede realizarse el mismo procesamiento incluso en el caso de codificación entrelazada para codificar una imagen como dos campos, un campo superior y un campo inferior. En la codificación entrelazada, puede lograrse mejor la eficiencia de codificación, ya que se duplica el número de imágenes de referencia. En este caso, puede usarse como prioridad una imagen que tiene el mismo atributo que la imagen actual. Concretamente, cuando una imagen actual está en un campo superior, se prioriza un campo superior para usarse como imagen indicada por el número de imagen de referencia por omisión DefRefNo. Por otra parte, cuando una imagen actual está en un campo inferior, se prioriza un campo inferior para usarse como imagen indicada por el número de imagen de referencia por omisión DefRefNo.

Séptima realización

Además, el procesamiento mostrado en cada una de las realizaciones anteriores puede llevarse a cabo fácilmente en un sistema informático independiente grabando un programa para realizar el procedimiento de codificación/decodificación de imágenes descrito en cada una de las realizaciones anteriores sobre un medio de grabación como un disco flexible o similar.

Las Figs. 19A, 19B y 19C son ilustraciones de un medio de grabación para grabar un programa para realizar el procedimiento de codificación/decodificación descrito en las realizaciones anteriores en el sistema informático.

La Fig. 19B muestra un aspecto completo de un disco flexible, su estructura en sección transversal y el propio disco flexible mientras que la Fig. 19A muestra un ejemplo de un formato físico del disco flexible como cuerpo principal de un medio de grabación. Un disco flexible FD está contenido en una caja F con una pluralidad de pistas Tr formadas concéntricamente desde la periferia hacia el interior sobre la superficie del disco, y cada pista está dividida en 16 sectores Se en la dirección angular. De este modo, el programa se almacena en un área asignada para ella sobre el disco flexible FD.

La Fig. 19C muestra una configuración para grabar y reproducir el programa en el disco flexible FD. Cuando el programa se graba en el disco flexible FD, el sistema informático Cs escribe el programa por medio de una unidad de disco flexible FDD. Cuando el aparato de codificación y el aparato de decodificación se construyen en el sistema informático usando el programa del disco flexible, el programa se lee del disco flexible y luego se transfiere al sistema informático mediante la unidad de disco flexible FDD.

La explicación anterior se hace suponiendo que un medio de grabación es un disco flexible, pero también puede realizarse el mismo procesamiento usando un disco óptico. Además, el medio de grabación no está limitado a un disco flexible y un disco óptico, sino que puede usarse cualquier otro medio como una tarjeta IC y una casete ROM capaz de grabar un programa.

ES 2 306 380 T3

Lo siguiente es una descripción para las aplicaciones del procedimiento de codificación/decodificación de imágenes ilustrado en las realizaciones anteriormente mencionadas y un sistema que las usa.

La Fig. 20 es un diagrama de bloques que muestra una configuración general de un sistema de suministro de contenido ex100 para realizar servicio de distribución de contenido. El área para proporcionar servicio de comunicación está dividida en celdas de tamaño deseado, y estaciones base ex107 ~ ex110, que son estaciones inalámbricas fijas, están situadas en celdas respectivas.

Este sistema de suministro de contenido ex100 está conectado a aparatos como un ordenador ex111, un asistente personal digital (PDA) ex112, una cámara ex113, un teléfono celular ex114, y un teléfono celular con una cámara ex115 a través, por ejemplo, de la Internet ex101, un proveedor de servicios de Internet ex102, una red telefónica ex104, así como las estaciones base ex107 ~ ex110.

Sin embargo, el sistema de suministro de contenido ex100 no está limitado a la configuración mostrada en la Fig. 20, y puede estar conectado a una combinación de cualquiera de ellos. Además, cada aparato puede estar conectado directamente a la red telefónica ex104, no a través de las estaciones base ex107 ~ ex110.

La cámara ex113 es un aparato capaz de filmar vídeo, como una cámara de vídeo digital. El teléfono celular ex114 puede ser un teléfono celular de cualquiera de los siguientes sistemas: un sistema de comunicaciones digitales personales (PDC), un sistema de acceso múltiple por división de código (CDMA), un sistema de acceso múltiple por división de código-banda ancha (W-CDMA) o un sistema global para comunicaciones móviles (GSM), un sistema de teléfono personal (PHS) o similar.

Un servidor de descarga y reproducción en tiempo real ex103 está conectado a la cámara ex113 a través de la red telefónica ex104 y también la estación base ex109, que lleva a cabo una distribución en directo o similar usando la cámara ex113 basándose en los datos codificados transmitidos desde el usuario. Puede codificar los datos la cámara ex113, el servidor que transmite los datos y similares. Los datos de imágenes en movimiento filmados por una cámara ex116 pueden ser transmitidos al servidor de descarga y reproducción en tiempo real ex103 a través del ordenador ex111. En este caso, la cámara ex116 o el ordenador ex111 pueden codificar los datos de imágenes en movimiento. Un LSI ex117 incluido en el ordenador ex111 y la cámara ex116 realiza el procesamiento de codificación. El software para codificar y decodificar imágenes puede estar integrado en cualquier tipo de medio de grabación (como un CD-ROM, un disco flexible y un disco duro) es decir, un medio de grabación que sea legible por el ordenador ex111 o similar. Además, un teléfono celular con una cámara ex115 puede transmitir los datos de imágenes en movimiento. Estos datos de imágenes en movimiento son los datos codificados por el LSI incluido en el teléfono celular ex115.

El sistema de suministro de contenido ex100 codifica el contenido (como un vídeo de música en directo) filmado por un usuario usando la cámara ex113, la cámara ex116 o similar de la misma manera que se muestra en las realizaciones anteriormente mencionadas y lo transmite al servidor de descarga y reproducción en tiempo real ex103, y el servidor de descarga y reproducción en tiempo real ex103 hace distribución en tiempo real de los datos de contenido a los clientes a petición suya. Los clientes incluyen el ordenador ex111, la PDA ex112, la cámara ex113, el teléfono celular ex114, etcétera, capaces de decodificar los datos codificados anteriormente mencionados. En el sistema de suministro de contenido ex100, los clientes pueden recibir y reproducir así los datos codificados, y además pueden recibir, decodificar y reproducir los datos en tiempo real para llevar a cabo radiodifusión personal.

Cuando cada aparato en este sistema realiza codificación o decodificación, puede usarse el aparato de codificación de imágenes o el aparato de decodificación de imágenes mostrados en la realización anteriormente mencionada.

Como ejemplo de tal aparato se explicará un teléfono celular.

La Fig. 21 es un diagrama que muestra el teléfono celular ex115 que usa el procedimiento de codificación/decodificación de imágenes explicado en las realizaciones anteriormente mencionadas. El teléfono celular ex115 tiene una antena ex201 para comunicación con la estación base ex110 por medio de ondas de radio, una unidad de cámara ex203 como una cámara CCD capaz de filmar imágenes en movimiento y fijas, una unidad de visualización ex202 como una pantalla de cristal líquido para visualizar los datos como imágenes decodificadas y similares filmados por la unidad de cámara ex203 o recibidos por la antena ex201, una unidad de cuerpo que incluye un conjunto de teclas de operación ex204, una unidad de salida de voz ex208 como un altavoz para generar voz, una unidad de entrada de voz ex205 como un micrófono para introducir voz, un medio de grabación ex207 para grabar datos codificados o decodificados como datos de imágenes en movimiento o fijas filmadas por la cámara, datos de correos electrónicos recibidos y datos de imágenes en movimiento o fijas, y una unidad de ranura ex206 para acoplar el medio de grabación ex207 al teléfono celular ex115. El medio de grabación ex207 guarda en sí mismo un elemento de memoria flash, una clase de memoria de sólo lectura borrrable y programable electrónicamente (EEPROM) que es una memoria no volátil borrrable y regrabable eléctricamente almacenada en una caja de plástico como una tarjeta SD.

A continuación se explicará el teléfono celular ex115 con referencia a la Fig. 22. En el teléfono celular ex115, una unidad de control principal ex311, diseñada para controlar de manera centralizada cada unidad del cuerpo principal que contiene la unidad de visualización ex202 así como las teclas de operación ex204, está conectada mutuamente a una unidad de circuito de alimentación eléctrica ex310, una unidad de control de entrada de operación ex304,

ES 2 306 380 T3

una unidad de codificación de imágenes ex312, una unidad de interfaz de cámara ex303, una unidad de control de pantalla de cristal líquido (LCD) ex302, una unidad de decodificación de imágenes ex309, una unidad de multiplexación/demultiplexación ex308, una unidad de lectura/escritura ex307, una unidad de circuito de módem ex306, y una unidad de procesamiento de voz ex305 a través de un bus síncrono ex313.

5

Cuando una tecla de fin de llamada o una tecla de encendido es activada por la utilización de un usuario, la unidad de circuito de alimentación eléctrica ex310 suministra a las unidades respectivas energía procedente de un paquete de baterías para activar el teléfono celular digital con una cámara ex115 para dejarlo dispuesto para funcionar.

10

En el teléfono celular ex115, la unidad de procesamiento de voz ex305 convierte las señales de voz recibidas por la unidad de entrada de voz ex205 en modo de conversación en datos de voz digital bajo el control de la unidad de control principal ex311 que incluye una CPU, una ROM y una RAM, la unidad de circuito de módem ex306 realiza procesamiento de espectro ensanchado para los datos de voz digital, y la unidad de comunicación ex301 realiza procesamiento de conversión digital a analógica y conversión de frecuencia para los datos, para transmitirlos a través de la antena ex201. También, en el teléfono celular ex115, la unidad de circuito de comunicación ex301 amplifica los datos recibidos por la antena ex201 en modo de conversación y realiza conversión de frecuencia y la conversión analógica a digital a los datos, la unidad de circuito de módem ex306 realiza procesamiento inverso de espectro ensanchado de los datos, y la unidad de procesamiento de voz ex305 los convierte en datos de voz analógica para generarla a través de la unidad de salida de voz ex208.

15

20

Además, al transmitir un correo electrónico en modo de comunicación de datos, los datos de texto del correo electrónico introducido accionando las teclas de operación ex204 en el cuerpo principal son enviados a la unidad de control principal ex311 a través de la unidad de control de entrada de operación ex304. En la unidad de control principal ex311, después de que la unidad de circuito de módem ex306 realiza procesamiento de espectro ensanchado de los datos de texto y la unidad de comunicación ex301 realiza la conversión digital a analógica y la conversión de frecuencia para los datos de texto, los datos se transmiten a la estación base ex110 a través de la antena ex201.

25

30

Cuando se transmiten datos de imágenes en modo de comunicación de datos, los datos de imágenes filmados por la unidad de cámara ex203 son suministrados a la unidad de codificación de imágenes ex312 a través de la unidad de interfaz de cámara ex303. Cuando no se transmiten, también es posible visualizar los datos de imágenes filmados por la unidad de cámara ex203 directamente en la unidad de visualización ex202 a través de la unidad de interfaz de cámara ex303 y la unidad de control de LCD ex302.

35

40

La unidad de codificación de imágenes ex312, que incluye el aparato de codificación de imágenes según se describe para la presente invención, comprime y codifica los datos de imágenes suministrados desde la unidad de cámara ex203 usando el procedimiento de codificación empleado por el aparato de codificación de imágenes como se muestra en las realizaciones anteriormente mencionadas para transformarlos en datos de imágenes codificados, y los envía a la unidad de multiplexación/demultiplexación ex308. En este momento, el teléfono celular ex115 envía la voz recibida por la unidad de entrada de voz ex205 durante la filmación con la unidad de cámara ex203 a la unidad de multiplexación/demultiplexación ex308 como datos de voz digital a través de la unidad de procesamiento de voz ex305.

45

La unidad de multiplexación/demultiplexación ex308 multiplexa los datos de imágenes codificados suministrados desde la unidad de codificación de imágenes ex312 y los datos de voz suministrados desde la unidad de procesamiento de voz ex305 usando un procedimiento predeterminado, luego la unidad de circuito de módem ex306 realiza procesamiento de espectro ensanchado de los datos multiplexados obtenidos como resultado de la multiplexación, y por último, la unidad de comunicación ex301 realiza conversión digital a analógica y transformación de frecuencia de los datos para la transmisión a través de la antena ex201.

50

En cuanto a recepción de datos de un archivo de imágenes en movimiento que está enlazado a una página Web o similar en modo de comunicación de datos, la unidad de circuito de módem ex306 realiza procesamiento inverso de espectro ensanchado para los datos recibidos desde la estación base ex110 a través de la antena ex201, y envía los datos multiplexados obtenidos como resultado del procesamiento inverso de espectro ensanchado.

55

Para decodificar los datos multiplexados recibidos a través de la antena ex201, la unidad de multiplexación/demultiplexación ex308 demultiplexa los datos multiplexados en un flujo codificado de datos de imágenes y un flujo codificado de datos de voz, y suministra los datos de imágenes codificados a la unidad de decodificación de imágenes ex309 y los datos de voz a la unidad de procesamiento de voz ex305, respectivamente a través del bus síncrono ex313.

60

65

A continuación, la unidad de decodificación de imágenes ex309, que incluye el aparato de decodificación de imágenes según se describe en la presente invención, decodifica el flujo codificado de los datos de imágenes usando el procedimiento de decodificación que corresponde al procedimiento de codificación como se muestra en las realizaciones anteriormente mencionadas para generar datos de imágenes en movimiento reproducidos, y suministra estos datos a la unidad de visualización ex202 a través de la unidad de control de LCD ex302, y de este modo se visualizan los datos de imágenes incluidos en el archivo de imágenes en movimiento enlazado a una página Web, por ejemplo. Al mismo tiempo, la unidad de procesamiento de voz ex305 convierte los datos de voz en datos de voz analógica, y suministra estos datos a la unidad de salida de voz ex208, y de este modo se reproducen los datos de voz incluidos en el archivo de imágenes en movimiento enlazado a una página Web, por ejemplo.

ES 2 306 380 T3

La presente invención no está limitada al sistema anteriormente mencionado ya que últimamente se ha hablado mucho de la radiodifusión digital basada en tierra o por satélite y al menos el aparato de codificación de imágenes o el aparato de decodificación de imágenes descritos en las realizaciones anteriormente mencionadas pueden estar incorporados dentro de un sistema de radiodifusión digital como se muestra en la Fig. 23. Más concretamente, un flujo codificado de información de vídeo es transmitido o comunicado desde una estación de radiodifusión ex409 a un satélite de radiodifusión ex410 a través de ondas de radio. Al recibirlo, el satélite de radiodifusión ex410 transmite ondas de radio para radiodifusión. Después, una antena de uso doméstico ex406 con una función de recepción de radiodifusión por satélite recibe las ondas de radio, y una televisión (receptor) ex401 o un decodificador (STB) ex407 decodifica un flujo de bits codificados para la reproducción. El aparato de decodificación de imágenes como el mostrado en las realizaciones anteriormente mencionadas puede implementarse en el aparato de reproducción ex403 para leer y decodificar el flujo codificado grabado en un medio de grabación ex402 como un CD y un DVD. En este caso, las señales de imágenes en movimiento reproducidas se visualizan en un monitor ex404. También se concibe implementar el aparato de decodificación de imágenes en el STB ex407 conectado a un cable ex405 para una televisión por cable o la antena ex406 para radiodifusión por satélite y/o basada en tierra para reproducirlas en un monitor ex408 de la televisión ex401. El aparato de decodificación de imágenes puede estar incorporado dentro de la televisión, no en el decodificador. También, un automóvil ex412 que tiene una antena ex411 puede recibir señales procedentes del satélite ex410 o la estación base ex107 para reproducir imagen en movimiento en un dispositivo de visualización como un sistema de navegación para automóviles ex413 montado en el automóvil ex412.

Además, el aparato de codificación de imágenes como el mostrado en las realizaciones anteriormente mencionadas puede codificar señales de imágenes y grabarlas en el medio de grabación. Como ejemplo concreto, pueden citarse un grabador ex420 como un grabador de DVD para grabar señales de imagen en un disco DVD ex421, un grabador de disco para grabarlas en un disco duro. Pueden grabarse en una tarjeta SD ex422. Cuando el grabador ex420 incluye el aparato de decodificación de imágenes como el mostrado en la realización anteriormente mencionada, las señales de imagen grabadas en el disco DVD ex421 o la tarjeta SD ex422 pueden reproducirse para visualización en el monitor ex408.

En cuanto a la estructura del sistema de navegación para automóvil ex413, es concebible la estructura sin la unidad de cámara ex203, la unidad de interfaz de cámara ex303 y la unidad de codificación de imágenes ex312, de los componentes mostrados en la Fig. 22. Se aplica lo mismo para el ordenador ex111, la televisión (receptor) ex401 y otros.

Además, pueden concebirse tres tipos de implementaciones para un terminal como el teléfono celular ex114: un terminal de envío/recepción implementado tanto con un codificador como con un decodificador, un terminal de envío implementado sólo con un codificador, y un terminal de recepción implementado sólo con un decodificador.

Tal como se describió anteriormente, es posible usar el procedimiento de codificación de imágenes y el procedimiento de decodificación de imágenes descritos en las realizaciones anteriormente mencionadas para cualquiera de los aparatos y sistemas anteriormente mencionados, y usando estos procedimientos, pueden obtenerse los efectos descritos en las realizaciones anteriormente mencionadas.

De este modo, como se describió en detalle anteriormente, con el procedimiento de codificación de imágenes en movimiento según la presente invención, no hay necesidad de seleccionar, basándose en bloque por bloque, una sola imagen de imágenes codificadas plurales para una imagen de referencia, y tampoco hay necesidad de codificar esta imagen de referencia basándose en bloque por bloque, por lo tanto, puede llevarse a cabo la codificación eficiente y puede reducirse la carga de procesamiento.

Usar el procedimiento de decodificación de imágenes en movimiento según la presente invención también permite que los datos de imágenes en movimiento codificados, que son codificados usando una imagen de referencia común y una imagen de referencia según cada bloque y luego son generados, sean decodificados correctamente.

Aplicabilidad industrial

De este modo, el procedimiento de codificación de imágenes en movimiento y el procedimiento de decodificación de imágenes en movimiento según la presente invención son prácticos como procedimiento de codificación de cada imagen que compone una imagen en movimiento de entrada, con el uso, por ejemplo, de un teléfono celular, un aparato de DVD y un ordenador personal o similar, generándolos como datos de imágenes en movimiento codificados y decodificando los datos de imágenes en movimiento codificados.

REIVINDICACIONES

5 1. Un procedimiento de codificación de imágenes para dividir, en bloques, una imagen actual que ha de ser codificada, seleccionar una imagen de referencia de entre imágenes de referencia basándose en bloques, describir información que identifica la imagen de referencia seleccionada, y realizar codificación predictiva sobre el bloque, estando **caracterizado** dicho procedimiento por las etapas de:

10 seleccionar, para codificar una unidad de imagen de bloques plurales formada por una pluralidad de bloques, una imagen de referencia común a la que ha de hacerse referencia comúnmente, de entre imágenes de referencia plurales;

describir información común que identifica la imagen de referencia común seleccionada, en un área de información común para la unidad de imagen de bloques plurales, en lugar de describir, por bloque, información que identifica la imagen de referencia común seleccionada;

15 generar una imagen predictiva de un bloque actual que ha de ser codificado incluido en la unidad de imagen de bloques plurales, usando la imagen de referencia común seleccionada; y

codificar el bloque actual usando la imagen predictiva.

20 2. El procedimiento de codificación de imágenes según la reivindicación 1,

en el que la unidad de imagen de bloques plurales es una de una unidad de imágenes plurales, una unidad de imagen, y una unidad de franja.

25 3. Un aparato de codificación de imágenes para dividir, en bloques, una imagen actual que ha de ser codificada, seleccionar una imagen de referencia de entre imágenes de referencia basándose en bloques, describir información que identifica la imagen de referencia seleccionada, y realizar codificación predictiva sobre el bloque, estando **caracterizado** dicho aparato por:

30 una unidad (101) utilizable para seleccionar, para codificar una unidad de imagen de bloques plurales formada por una pluralidad de bloques, una imagen de referencia común a la que ha de hacerse referencia comúnmente, de entre imágenes de referencia plurales;

35 una unidad (102) utilizable para describir información común que identifica la imagen de referencia común seleccionada, en un área de información común para la unidad de imagen de bloques plurales, en lugar de describir, por bloque, información que identifica la imagen de referencia común seleccionada;

40 una unidad (102) utilizable para generar una imagen predictiva de un bloque actual que ha de ser codificado incluido en la unidad de imagen de bloques plurales, usando la imagen de referencia común seleccionada; y

una unidad (104) utilizable para codificar el bloque actual usando la imagen predictiva.

45 4. El aparato de codificación de imágenes según la reivindicación 3,

en el que la unidad de imagen de bloques plurales es una de una unidad de imágenes plurales, una unidad de imagen, y una unidad de franja.

50

55

60

65

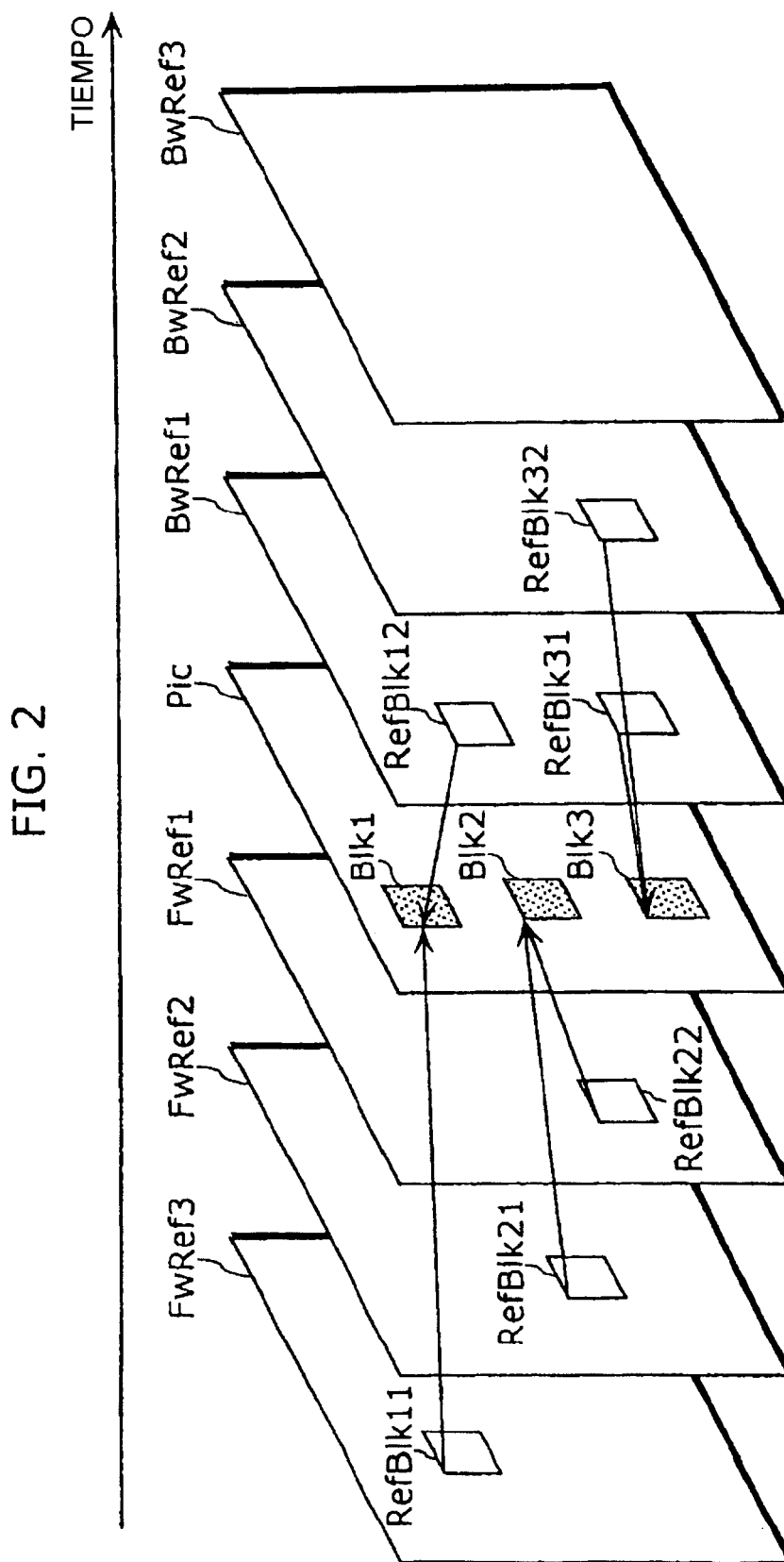


FIG. 3

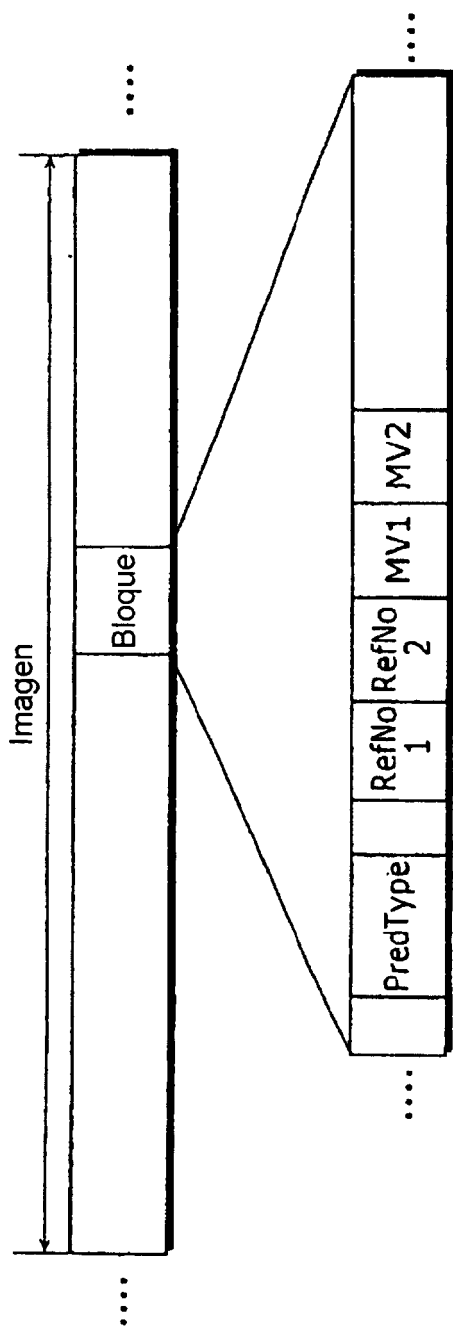


FIG. 4

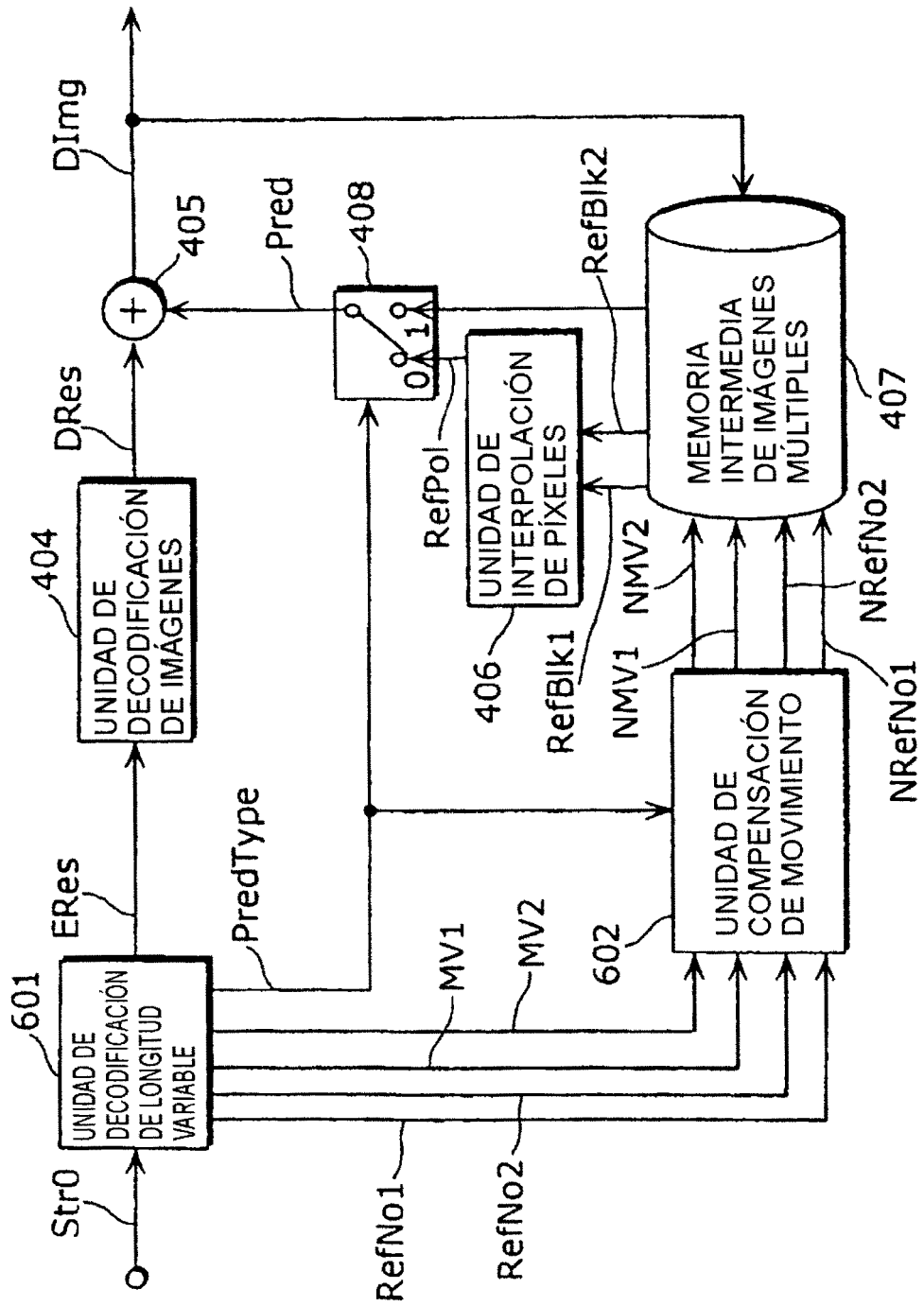
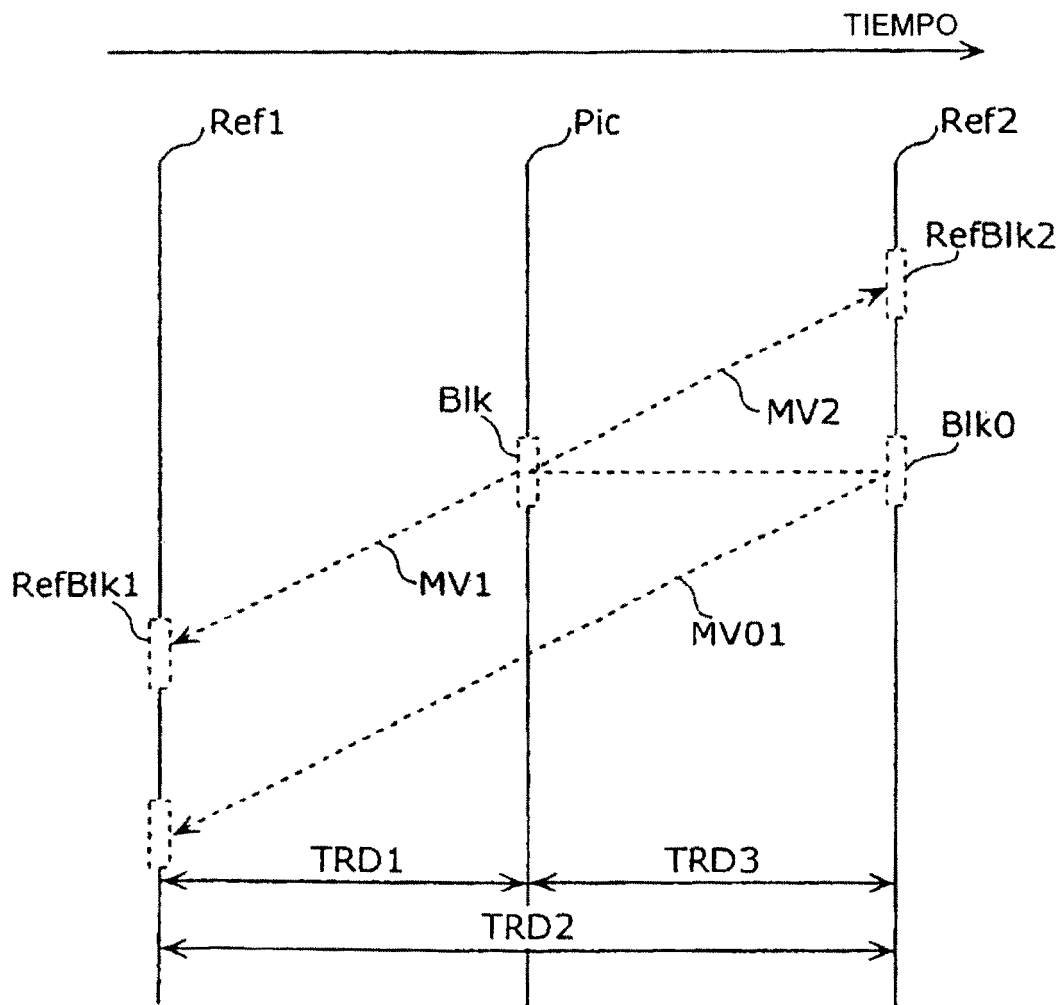


FIG. 5



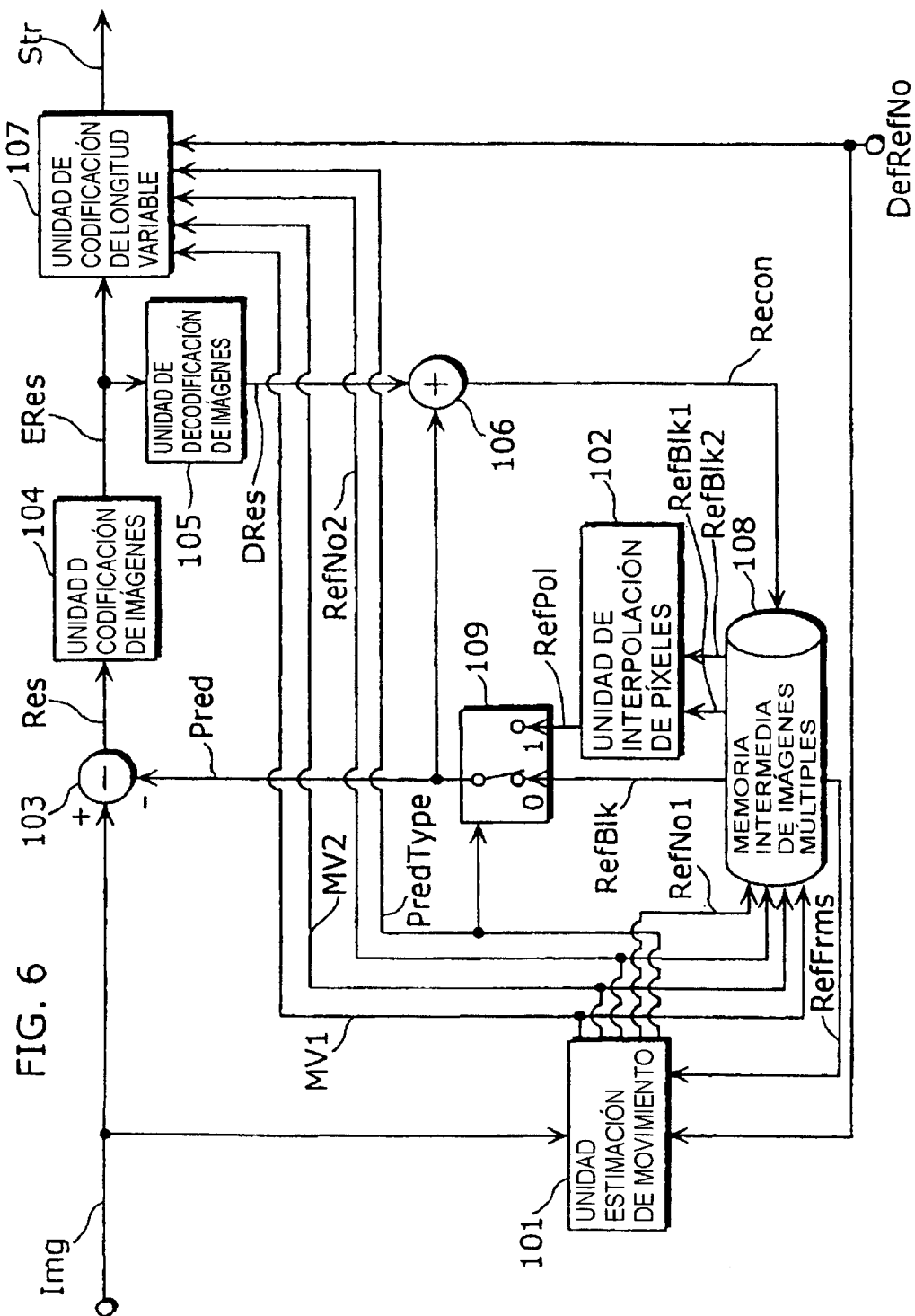


FIG. 6

FIG. 7

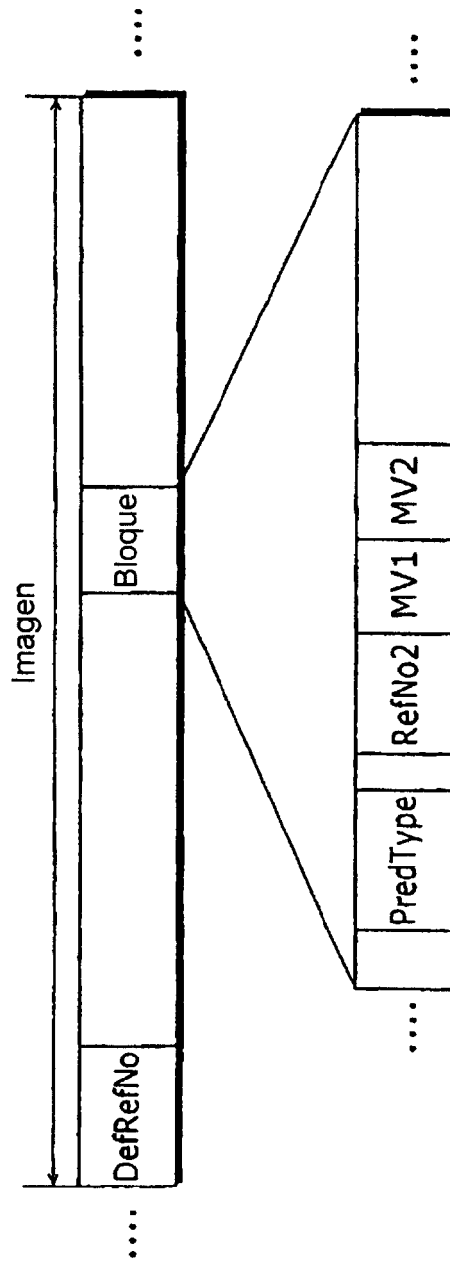


FIG. 8

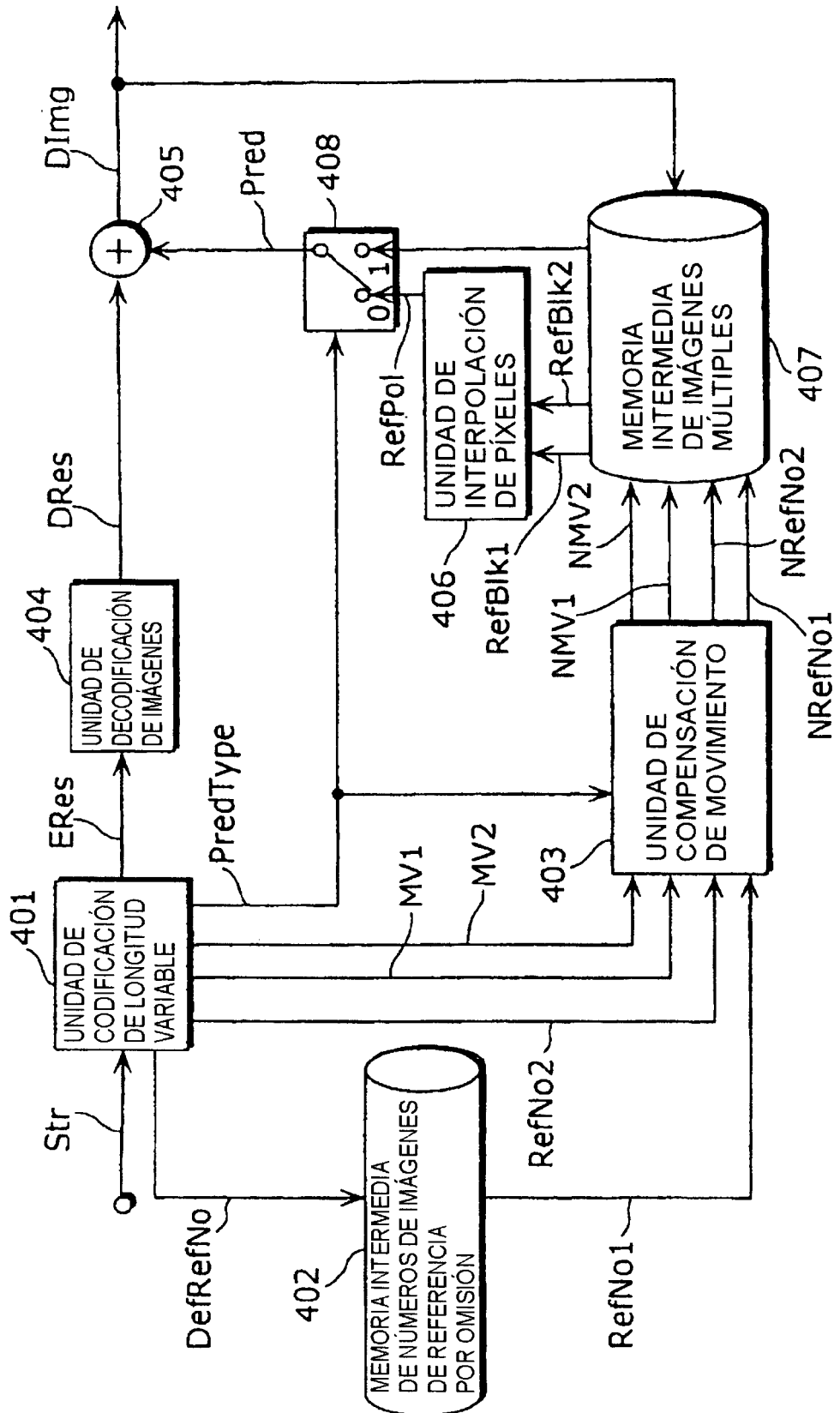
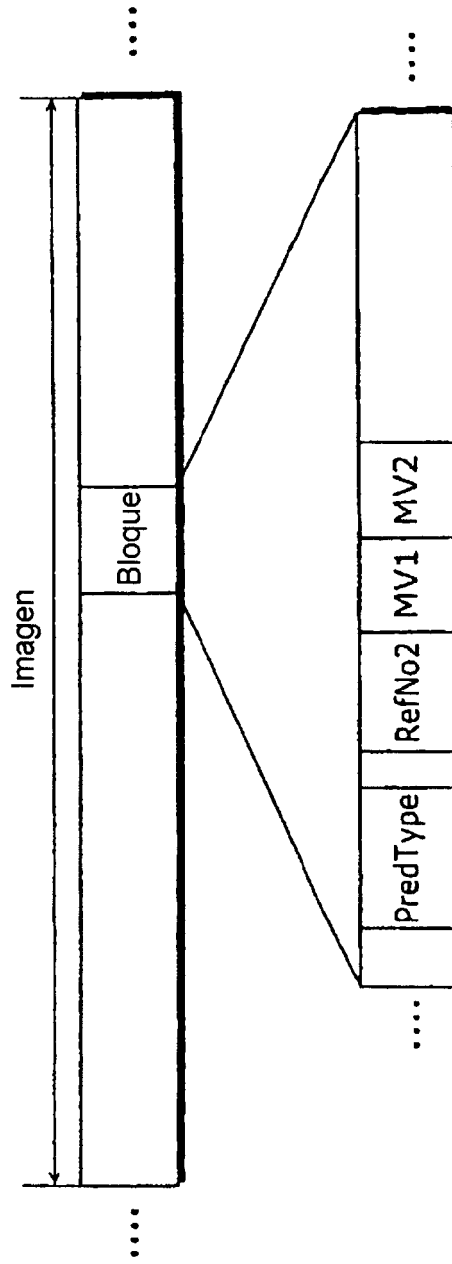


FIG. 10



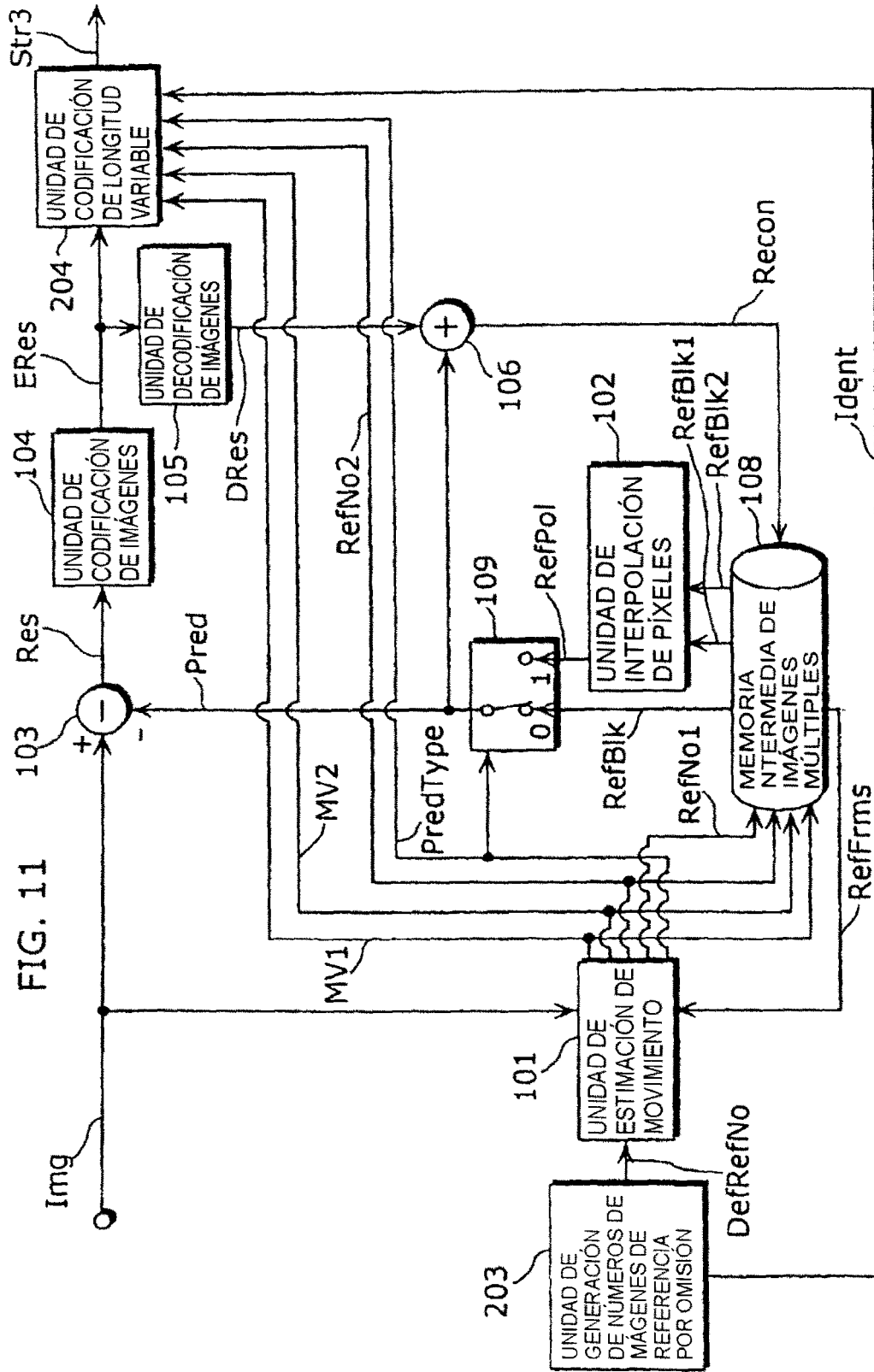


FIG. 12

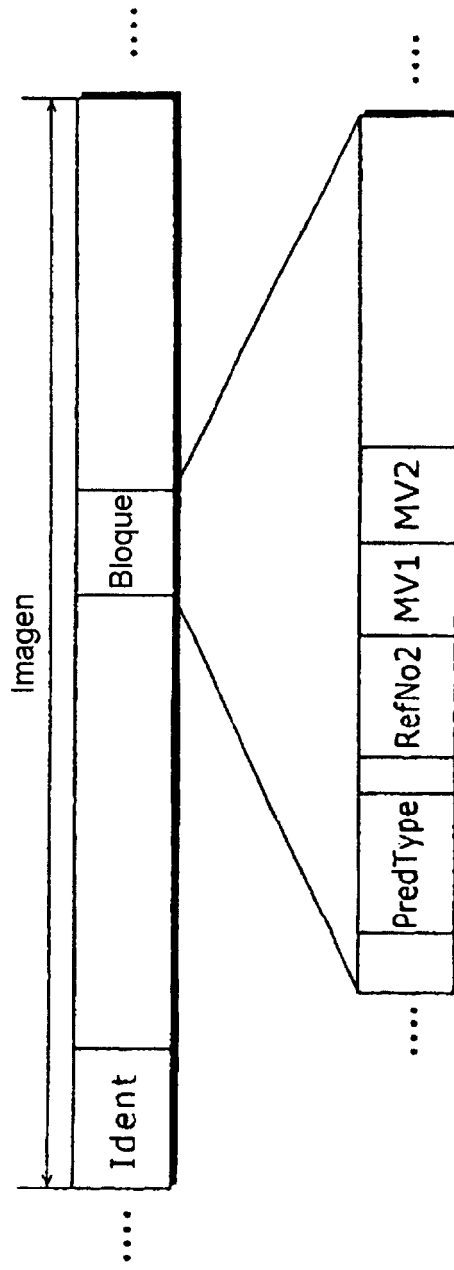


FIG. 14

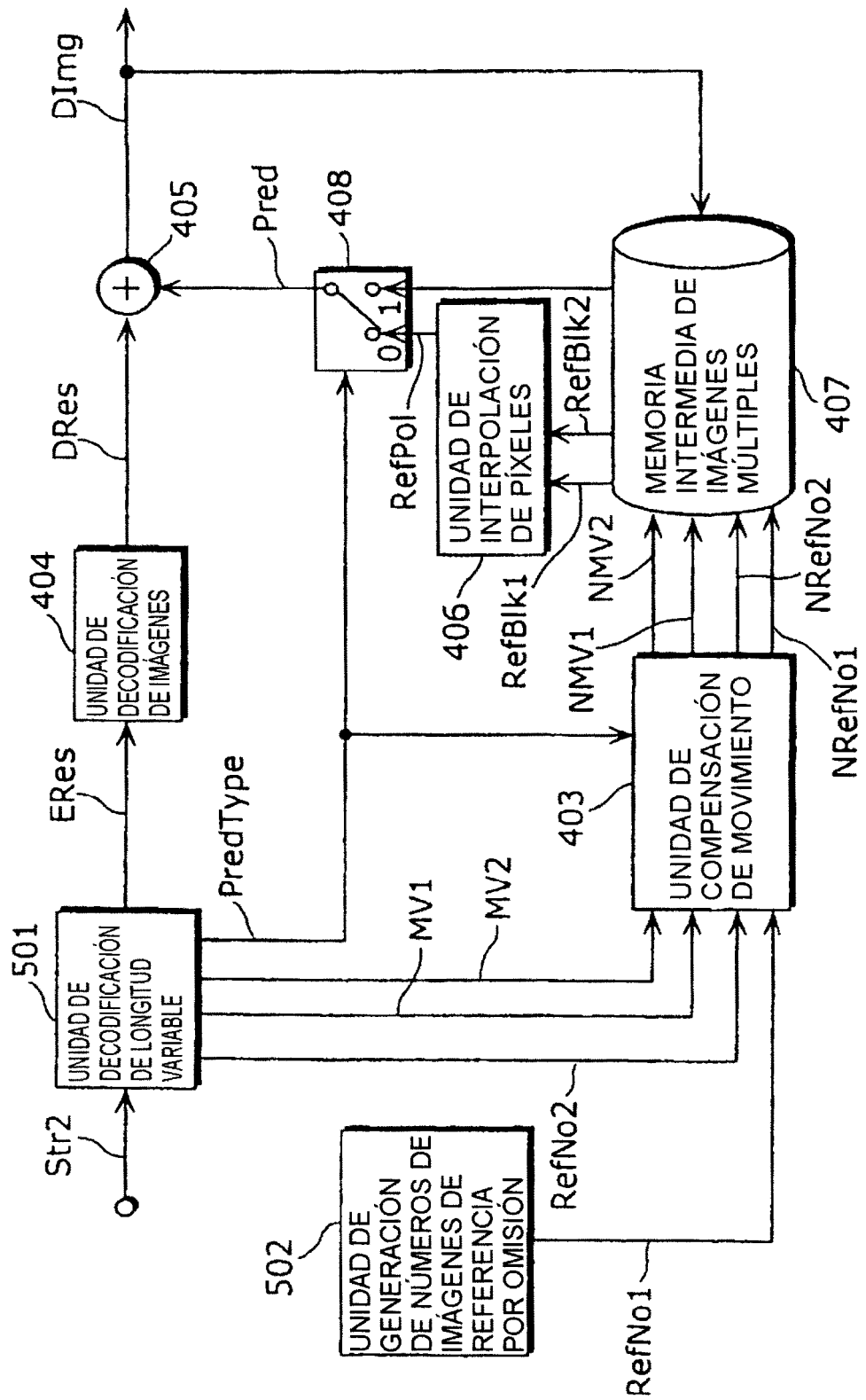
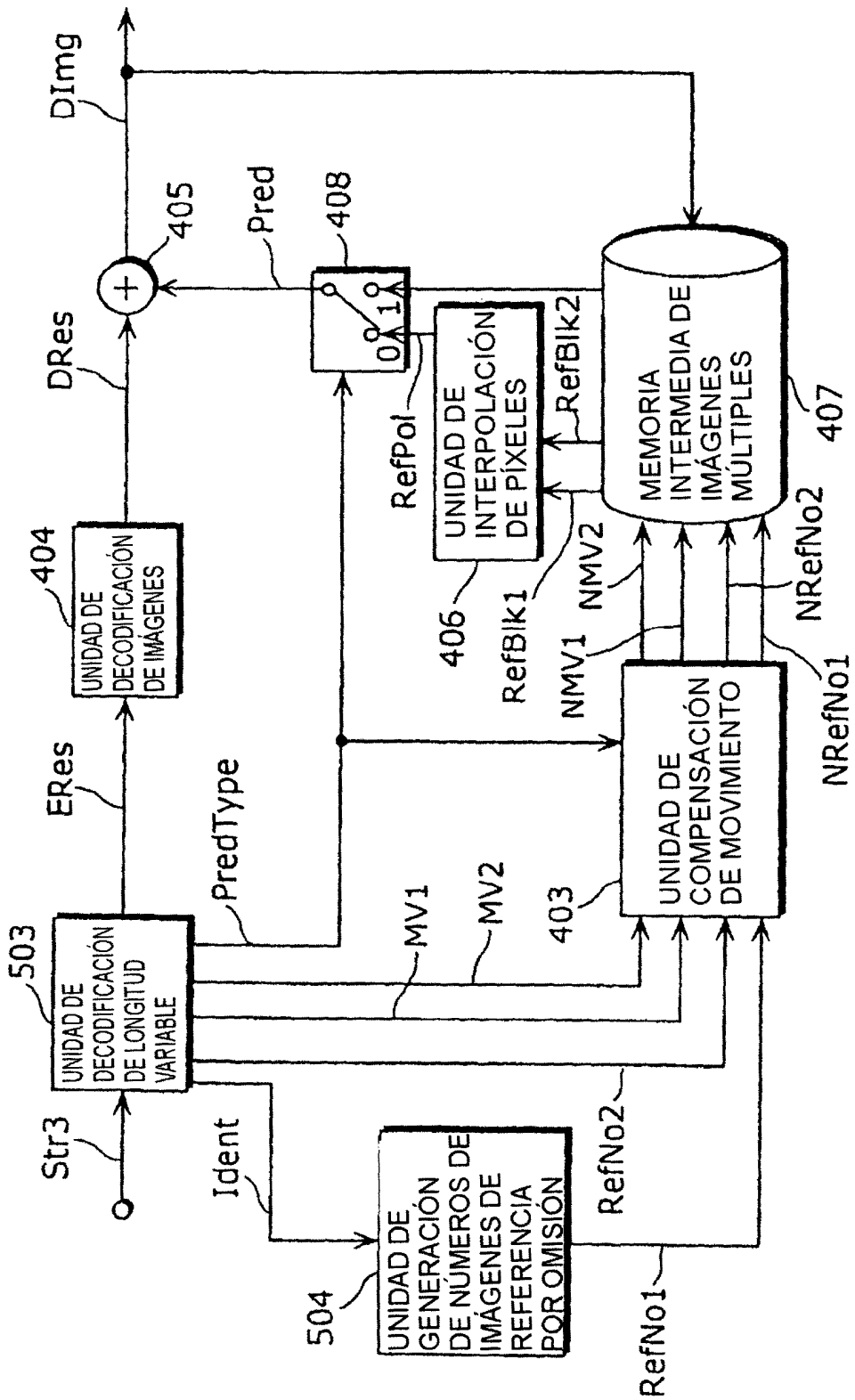


FIG. 15



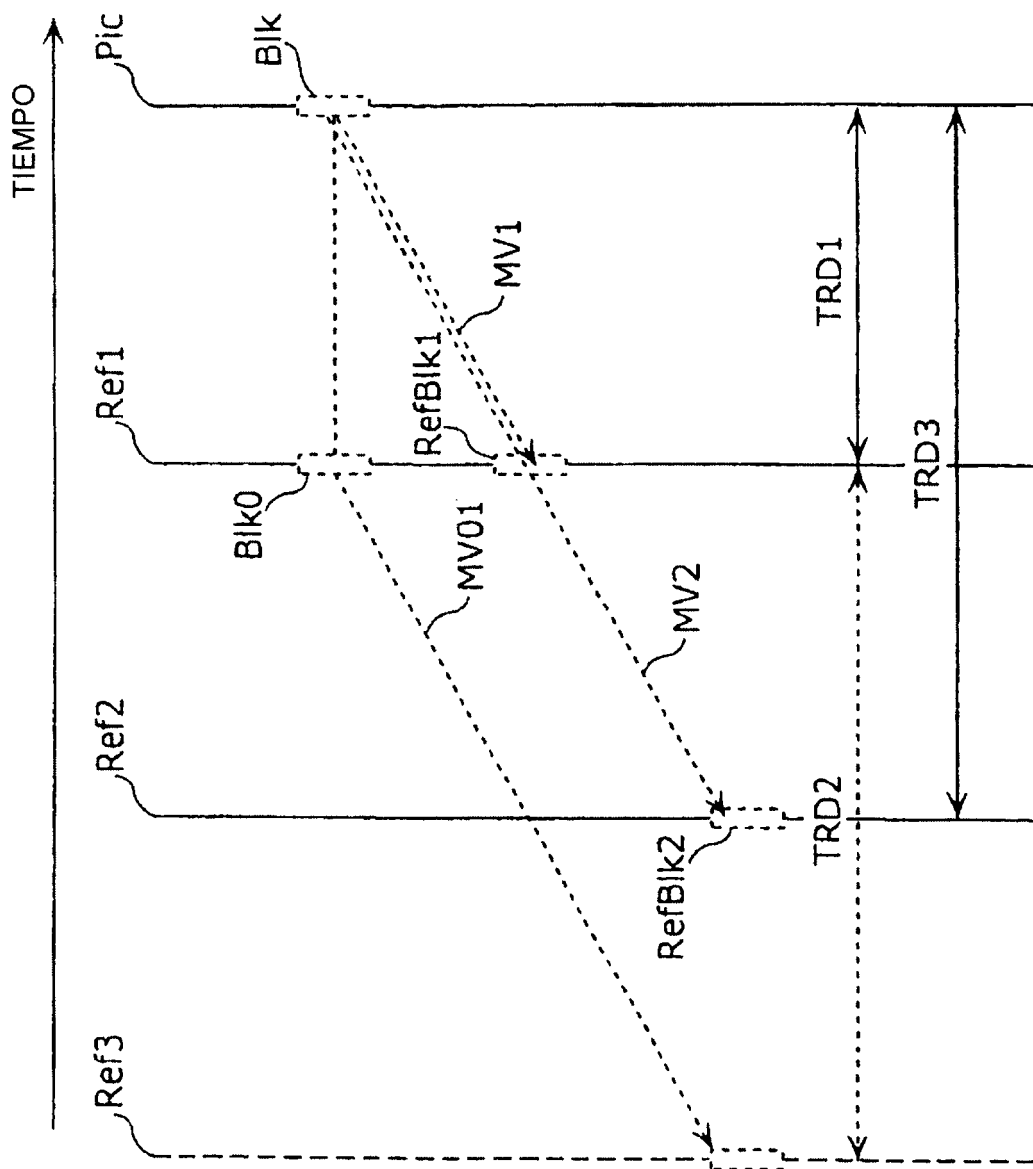


FIG. 16

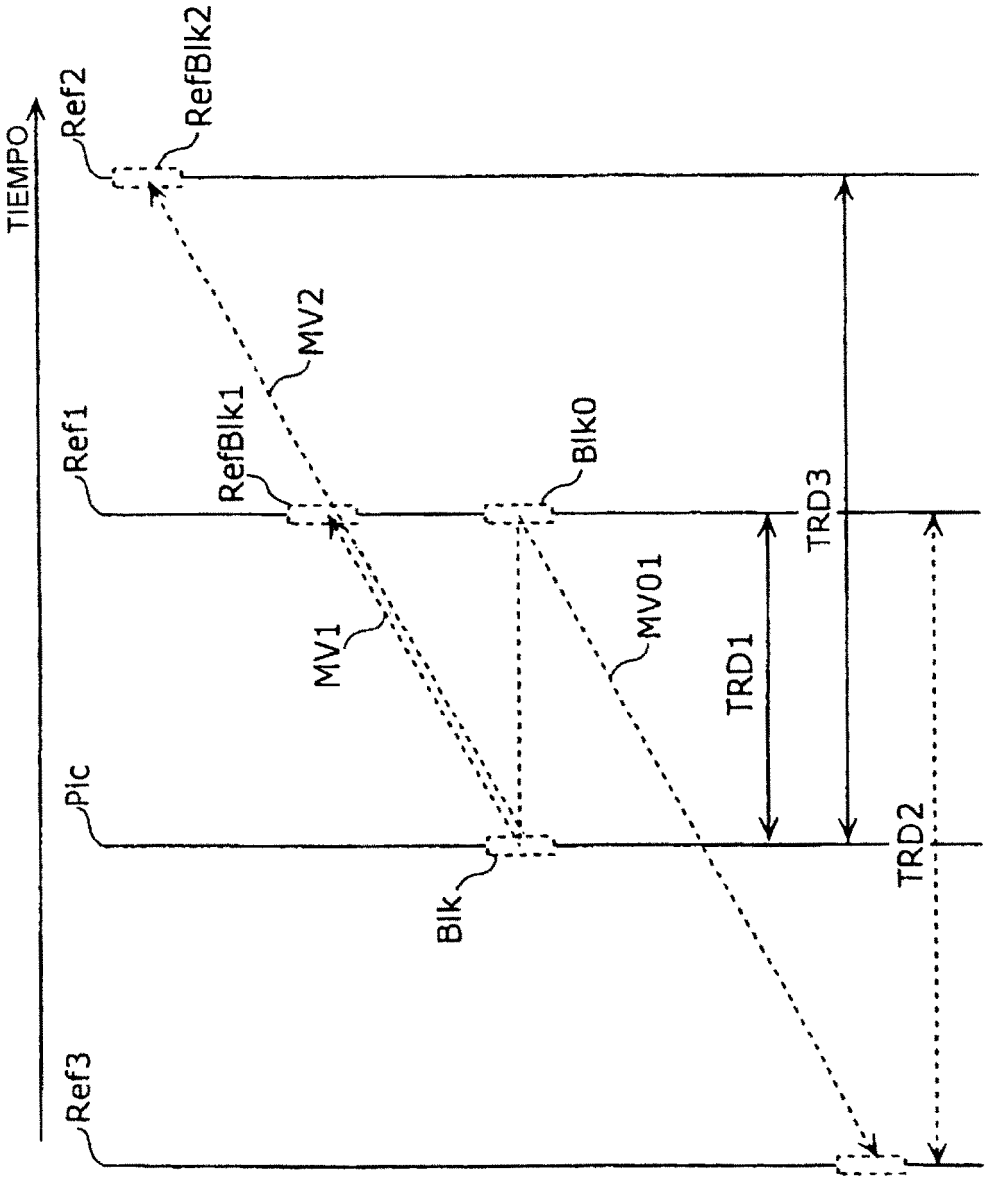


FIG. 17

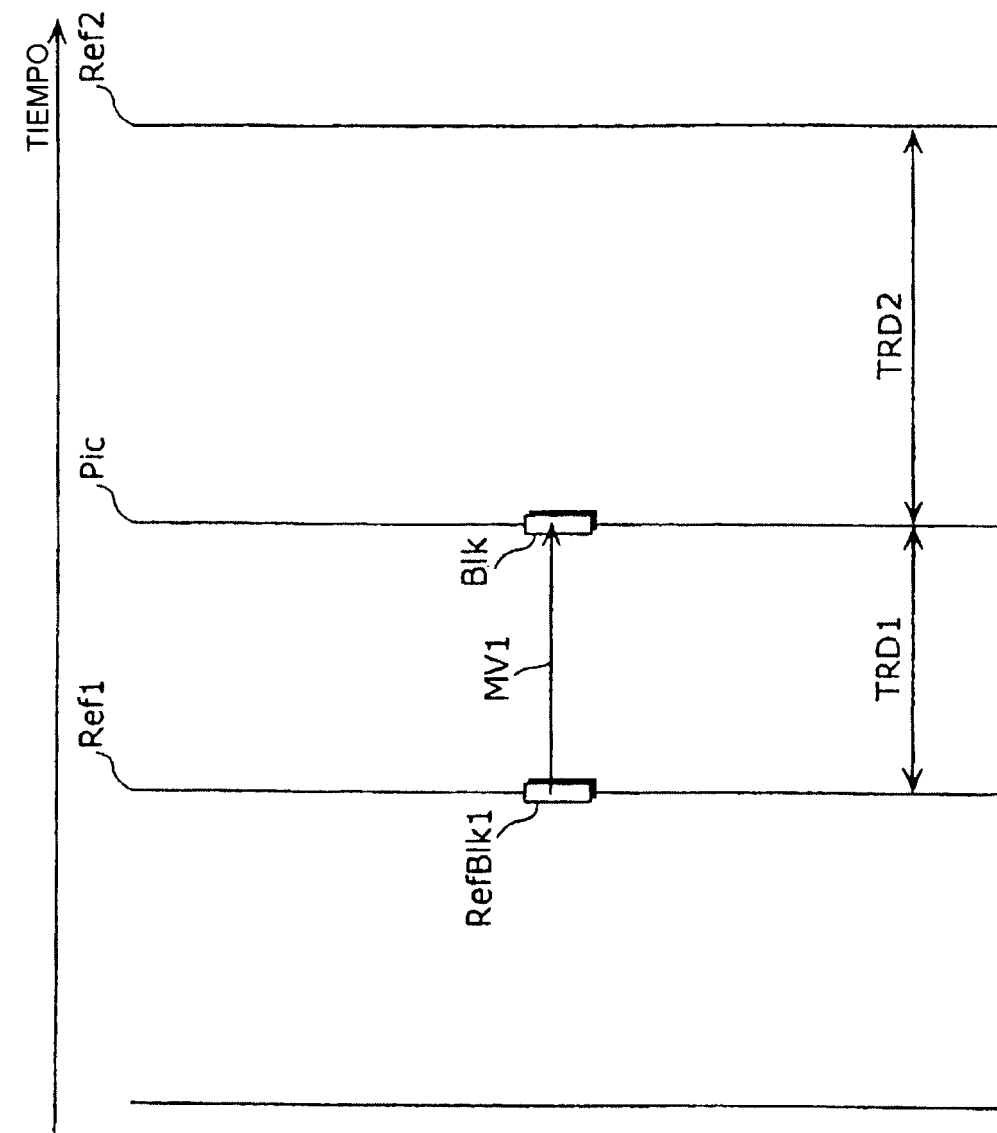
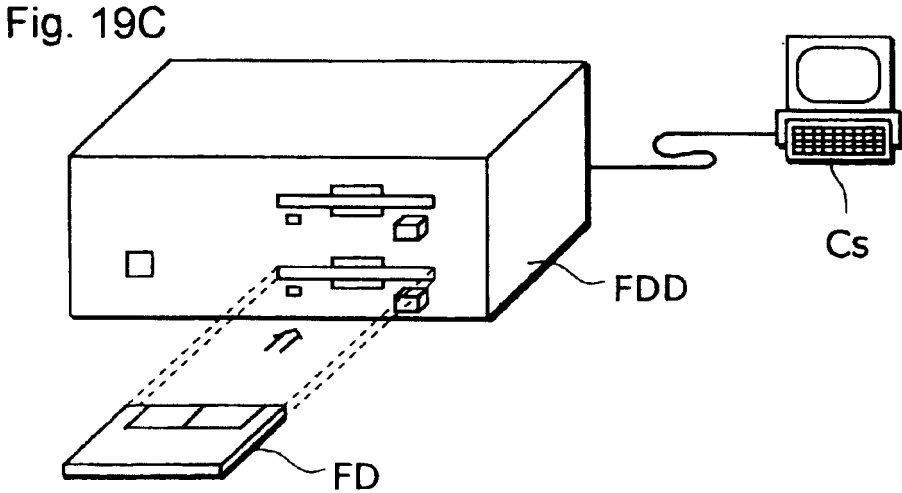
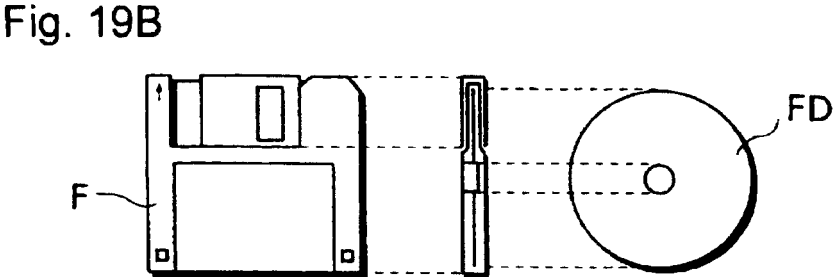
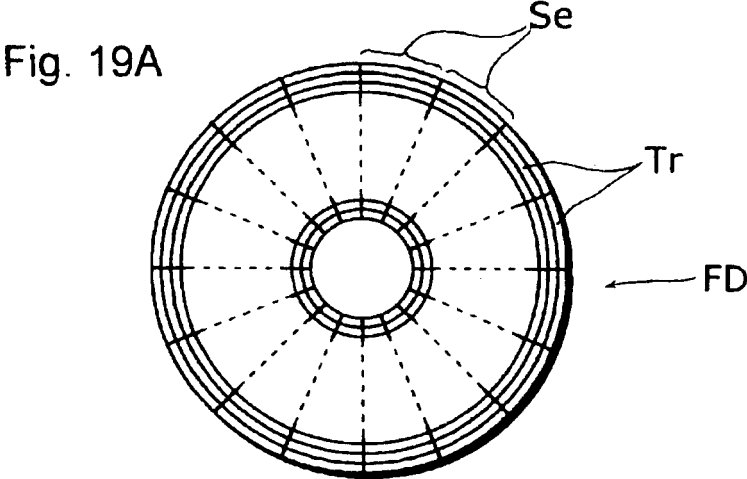


FIG. 18



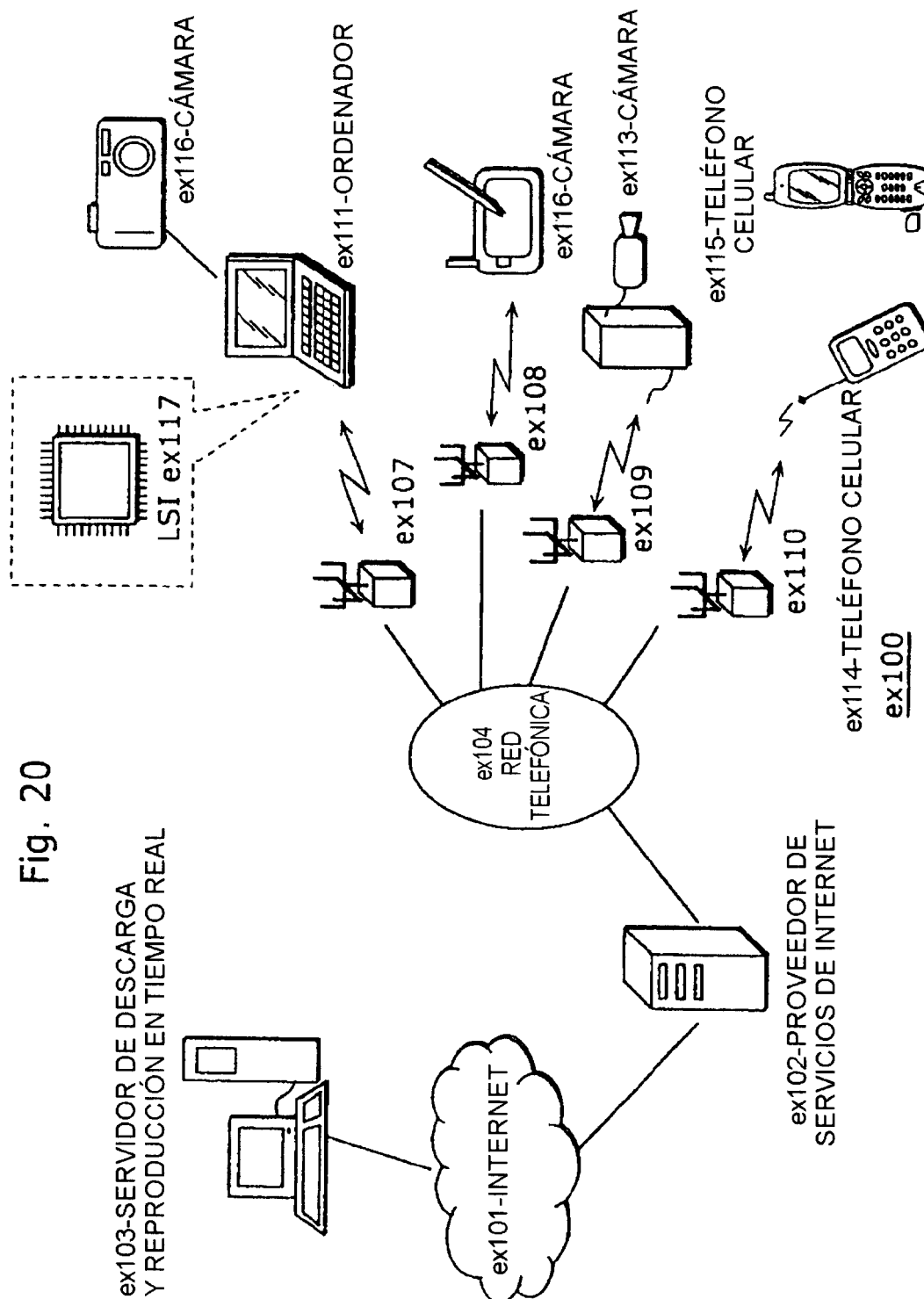


Fig. 20

Fig. 21

