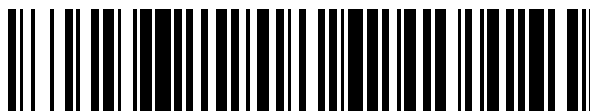


19



OFICINA ESPAÑOLA DE  
PATENTES Y MARCAS

ESPAÑA



11 Número de publicación: **2 728 323**

51 Int. Cl.:

**H04N 19/50** (2014.01)

**H04N 19/46** (2014.01)

**H04N 19/44** (2014.01)

12

TRADUCCIÓN DE PATENTE EUROPEA

T3

96 Fecha de presentación y número de la solicitud europea: **26.09.2003 E 18162792 (8)**

97 Fecha y número de publicación de la concesión europea: **01.05.2019 EP 3355584**

54 Título: **Procedimiento de codificación de vídeo, procedimiento de decodificación de vídeo, aparato de codificación de vídeo, aparato de decodificación de vídeo, programa de codificación de vídeo y programa de decodificación de vídeo**

30 Prioridad:

**03.10.2002 JP 2002291610**

**02.07.2003 JP 2003190634**

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

**23.10.2019**

73 Titular/es:

**NTT DOCOMO, INC. (100.0%)**

**11-1, Nagatacho 2-chome,  
Chiyoda-ku, Tokyo 100-6150, JP**

72 Inventor/es:

**ADACHI, SATORU;  
BOON, CHOONG SENG;  
KATO, SADAATSU;  
ETOH, MINORU y  
TAN, THIEW KENG**

74 Agente/Representante:

**FÚSTER OLAGUIBEL, Gustavo Nicolás**

**ES 2 728 323 T3**

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

## DESCRIPCIÓN

Procedimiento de codificación de vídeo, procedimiento de decodificación de vídeo, aparato de codificación de vídeo, aparato de decodificación de vídeo, programa de codificación de vídeo y programa de decodificación de vídeo

5

### Antecedentes de la invención

#### Campo de la invención

La presente invención se refiere a un procedimiento de codificación de vídeo, a un procedimiento de decodificación de vídeo, a un aparato de codificación de vídeo, a un aparato de decodificación de vídeo, a un sistema de procesamiento de vídeo, a un programa de codificación de vídeo y a un programa de decodificación de vídeo.

#### Técnica anterior relacionada

Se usan técnicas de codificación de señal de vídeo para la transmisión y almacenamiento-regeneración de señales de vídeo. Las técnicas bien conocidas incluyen, por ejemplo, los procedimientos de codificación de vídeo de normas internacionales tales como la recomendación UIT-T H.263 (denominada en lo sucesivo H.263), la norma internacional ISO/IEC 14496-2 (MPEG-4 Visual, denominada en lo sucesivo MPEG-4), etcétera. Otro procedimiento de codificación más nuevo conocido es un procedimiento de codificación de vídeo programado para una normalización internacional conjunta mediante UIT-T e ISO/IEC; la recomendación UIT-T H.264 y la norma internacional ISO/IEC 14496-10 (Borrador del Comité Final Conjunto de Especificación de Vídeo Conjunta, denominada en lo sucesivo H.26L).

Otros ejemplos de decodificación de vídeo conocidos incluyen las técnicas descritas en el documento US 6 205 177 B1 para decodificar imágenes de vídeo para su transmisión desde una primera estación hasta una segunda estación que incluye usar múltiples imágenes de referencia para determinar un vector de movimiento e información de retardo de tiempo en la primera estación con el fin de predecir una imagen nueva para su visualización en la segunda estación, y las técnicas descritas en el documento JP H10 248066 para un procedimiento de decodificación que usa señales de sincronización para impedir perturbaciones en imágenes visualizadas debido a retardos en el procesamiento de decodificación. El documento US 6 205 177 B1 divulga un procedimiento de decodificación de imágenes codificadas, comprendiendo el procedimiento: recibir una imagen codificada, y una temporización de decodificación  $Tr(n)$  que representa una temporización según la cual va a decodificarse la imagen codificada; decodificar la imagen codificada según la temporización de decodificación  $Tr(n)$  y derivar una imagen decodificada, en el que una memoria intermedia de imágenes decodificadas contiene al menos una imagen decodificada usada como referencia para decodificar al menos una de las imágenes codificadas.

Dado que una señal de vídeo en movimiento consiste en una serie de imágenes (cuadros) que varían poco a poco con el tiempo, es una práctica común en estos procedimientos de codificación de vídeo implementar una predicción inter-cuadro entre un cuadro recuperado como un objetivo para la codificación (cuadro actual) y otro cuadro (cuadro de referencia) y reducir de ese modo la redundancia temporal en la señal de vídeo. En este caso, cuando la predicción inter-cuadro se lleva a cabo entre el cuadro actual y un cuadro de referencia menos diferente del cuadro actual, la redundancia puede reducirse más y la eficiencia de codificación puede aumentarse.

Por este motivo, tal como se muestra en la Fig. 6, el cuadro de referencia para el cuadro actual A1 puede ser o bien un cuadro temporalmente anterior A0 o un cuadro temporalmente posterior A2 con respecto al cuadro actual A1. La predicción con el cuadro anterior se denomina predicción hacia delante, mientras que la predicción con el cuadro posterior se denomina predicción hacia atrás. La predicción bidireccional se define como una predicción en la que se selecciona arbitrariamente uno de los dos procedimientos de predicción, o como una predicción en la que se usan simultáneamente ambos procedimientos.

En general, con el uso de tal predicción bidireccional, como en el ejemplo mostrado en la Fig. 6, se almacenan previamente antes del cuadro actual cada uno de un cuadro temporalmente anterior como un cuadro de referencia para la predicción hacia delante y un cuadro temporalmente posterior como un cuadro de referencia para la predicción hacia atrás.

Las Fig. 7A y 7B son diagramas que muestran (A) la decodificación y (B) la emisión de los cuadros en el caso de la predicción bidireccional mostrada en la Fig. 6. Por ejemplo, en la decodificación de MPEG-4, en la que el cuadro actual A1 se decodifica mediante predicción inter-cuadro bidireccional, el cuadro A0, que es un cuadro temporalmente anterior, y el cuadro A2, que es un cuadro temporalmente posterior con respecto al cuadro actual A1, se decodifican primero como cuadros decodificados mediante predicción intra-cuadro sin usar la predicción inter-cuadro o como cuadros decodificados mediante predicción inter-cuadro hacia delante, antes de decodificar el cuadro actual A1, y se conservan como cuadros de referencia. Después, el cuadro actual A1 se decodifica mediante predicción bidireccional usando estos dos cuadros A0, A2 conservados de ese modo (Fig. 7A).

En este caso, por tanto, el orden de los tiempos de decodificación del cuadro de referencia temporalmente posterior

A2 y del cuadro actual A1 es inverso al orden de los tiempos de emisión de sus imágenes decodificadas respectivas. Cada uno de estos cuadros A0, A1 y A2 está asociado a información de tiempo de emisión 0, 1 o 2 y, por tanto, la secuencia temporal de los cuadros puede conocerse según esta información. Por este motivo, las imágenes decodificadas se emiten en el orden correcto (Fig. 7B). En MPEG-4, la información de tiempo de emisión se describe como valores absolutos.

Algunos de los procedimientos de codificación de vídeo recientes permiten que la predicción inter-cuadro mencionada anteriormente se lleve a cabo usando múltiples cuadros de referencia, en lugar de un cuadro de referencia en el sentido hacia delante y un cuadro de referencia en el sentido hacia atrás, para permitir una predicción a partir de un cuadro con un cambio más pequeño con respecto al cuadro actual, como se muestra en la Fig. 8. La Fig. 8 muestra un ejemplo que usa dos cuadros temporalmente anteriores B0, B1 y dos cuadros temporalmente posteriores B3, B4 con respecto al cuadro actual B2 como cuadros de referencia para el cuadro actual B2.

Las Fig. 9A y 9B son diagramas que muestran (A) la decodificación y (B) la emisión de los cuadros en el caso de la predicción bidireccional mostrada en la Fig. 8. Por ejemplo, en la decodificación de la H.26L, puede conservarse una pluralidad de cuadros de referencia dentro de un intervalo hasta un límite superior predeterminado del número de cuadros de referencia y, en caso de llevar a cabo una predicción inter-cuadro, se designa arbitrariamente entre ellos un cuadro de referencia óptimo. En este caso, donde el cuadro actual B2 se decodifica como un cuadro predicho de manera bidireccional, los cuadros de referencia se decodifican en primer lugar antes de la decodificación del cuadro actual B2; los cuadros de referencia incluyen una pluralidad de cuadros temporalmente anteriores (por ejemplo, dos cuadros B0, B1) y una pluralidad de cuadros temporalmente posteriores (por ejemplo, dos cuadros B3, B4) con respecto al cuadro actual B2, que se descodifican y conservan como cuadros de referencia. El cuadro actual B2 puede predecirse a partir de un cuadro designado arbitrariamente como el usado para una predicción de entre esos cuadros B0, B1, B3 y B4 (Fig. 9A).

En este caso, por tanto, el orden de los tiempos de decodificación de los cuadros de referencia temporalmente posteriores B3, B4 y del cuadro actual B2 se vuelve inverso al orden de sus tiempos de emisión respectivos. Cada uno de estos cuadros B0-B4 está asociado a información de tiempo de emisión o a información de orden de emisión 0-4, y la secuencia temporal de los cuadros puede conocerse según esta información. Por este motivo, las imágenes decodificadas se emiten en el orden correcto (Fig. 9B). La información de tiempo de emisión se describe a menudo como valores absolutos. El orden de emisión se usa cuando los intervalos de cuadro son constantes.

Para llevar a cabo la decodificación mediante la predicción hacia atrás usando cuadros temporalmente posteriores como cuadros predictivos, es necesario cumplir la condición de que la decodificación de los cuadros temporalmente posteriores se complete antes de la decodificación del cuadro actual para que estén disponibles como cuadros predictivos. En este caso, se provoca un retardo antes de que la imagen decodificada del cuadro actual se vuelva disponible, en comparación con un cuadro en el que no se aplica la predicción hacia atrás.

Esto se describirá específicamente más adelante con referencia a las Fig. 10A a 10C. Las Fig. 10A a 10C corresponden al ejemplo mostrado en las Fig. 6, 7A y 7B. En primer lugar, se decodifican los datos codificados de cada cuadro A0-A2 en un orden necesario para la ejecución de una predicción inter-cuadro, y se supone que los intervalos de los cuadros son intervalos de tiempo constantes según una velocidad de cuadro y que el tiempo necesario para la operación de decodificación es despreciable para cada cuadro A0-A2, independientemente de si se aplica la predicción inter-cuadro, e independientemente de los sentidos de la predicción inter-cuadro (Fig. 10A). En la práctica, los intervalos de decodificación de los cuadros A0-A2 no tienen que ser constantes y pueden cambiar dependiendo de factores tales como una variación en los bits de codificación de los cuadros A0-A2 o similares; sin embargo, puede suponerse que son constantes en promedio. El tiempo necesario para la operación de decodificación no es cero tampoco, pero no planteará un problema importante en la descripción que se ofrece a continuación en el presente documento si la diferencia del mismo no es tan grande entre los cuadros A0-A2.

Se supone en el presente documento que un tiempo en el que se obtiene una imagen decodificada de un cuadro A0 sin retardo debido a una predicción hacia atrás y sin inversión de los órdenes de los tiempos de codificación y los tiempos de emisión con respecto a cualquier otro cuadro (un cuadro sin retardo y sin inversión se denominará en lo sucesivo un cuadro no asociado con predicción hacia atrás), se define como un tiempo de emisión correlacionado con la imagen decodificada, y la imagen decodificada se emite en el tiempo de emisión. Suponiendo que el cuadro posterior es el cuadro predicho hacia atrás A1, la imagen decodificada del mismo se decodificará después del cuadro temporalmente posterior A2, y se produce por tanto un retardo antes de que se obtenga la imagen decodificada.

Por este motivo, si el tiempo en el que se obtiene la imagen decodificada para el cuadro no asociado con predicción hacia atrás A0 se define como una referencia de tiempo de emisión, la imagen decodificada del cuadro predicho hacia atrás A1 no se obtiene mediante el tiempo de emisión correlacionado con la misma (Fig. 10B). Concretamente, un intervalo de tiempo de emisión entre la imagen decodificada del cuadro no asociado con predicción hacia atrás A0 y la imagen decodificada del cuadro predicho hacia atrás A1 se vuelve más largo debido al tiempo de retardo necesario para la ejecución de la predicción hacia atrás que el intervalo original, lo que da lugar a una emisión de

vídeo forzada.

Por tanto, en caso de aplicar la predicción inter-cuadro hacia atrás en la codificación de vídeo, tal como se muestra en la Fig. 10C, es necesario retardar previamente el tiempo de emisión de la imagen decodificada del cuadro no asociado con predicción hacia atrás A0 con el tiempo de retardo necesario para la ejecución de la predicción hacia atrás así como para poder manejar correctamente el intervalo de tiempo de emisión para el cuadro predicho hacia atrás A1.

De manera convencional, la predicción inter-cuadro hacia atrás se ha aplicado a la codificación de vídeo en las condiciones en las que se ha llevado a cabo la codificación a una alta tasa de bit y a la velocidad de cuadro fija de 30 cuadros/segundo al igual que las de señales de radiodifusión de televisión que se usaban siempre, tal como radiodifusión de televisión o acumulación de las mismas, puesto que la predicción inter-cuadro hacia atrás aporta más opciones para la predicción y, por tanto, un aumento de la complejidad computacional, de modo que hace difícil su implementación en un equipo simple, y puesto que no se deseaba aumentar el tiempo de retardo en comunicaciones en tiempo real que implican una interlocución bidireccional, como videoconferencias.

En este caso, por ejemplo, como en MPEG-4, donde se usa un cuadro temporalmente posterior como un cuadro de referencia para la predicción hacia atrás, el tiempo de retardo requerido en la ejecución de la predicción hacia atrás es constante. Por ejemplo, cuando la velocidad de cuadro es de 30 cuadros/segundo tal como se describió anteriormente, el tiempo de retardo es un intervalo de tiempo de cada cuadro, es decir, 1/30 segundos. Por consiguiente, el tiempo que debe retardarse el tiempo de emisión de la imagen decodificada del cuadro no asociado con predicción hacia atrás puede establecerse de igual manera a 1/30 segundos.

### Sumario de la invención

En los últimos años, sin embargo, como resultado de la mejora en el rendimiento informático y el progreso en la diversificación de los servicios de vídeo, el retardo es tolerable en la distribución de vídeo a través de Internet y comunicaciones móviles, y se ha aumentado el uso de codificación de vídeo que requiere una codificación a bajas tasas de bit. Para implementar la codificación a bajas tasas de bit, se aplican velocidades de cuadro inferiores a 30 cuadros/segundo, o se usan velocidades de cuadro variables para cambiar dinámicamente la velocidad de cuadro con el fin de controlar la tasa de bit de codificación.

En tal codificación de vídeo en la que se aplica la predicción hacia atrás mencionada anteriormente con el fin de aumentar más la eficiencia de codificación, el tiempo de retardo debido a la predicción hacia atrás no es siempre de 1/30 segundos, tal como se usó anteriormente. En la aplicación de velocidades de cuadro variables, las velocidades de cuadro no son constantes. Por ejemplo, en caso de usarse una pequeña velocidad de cuadro con carácter temporal, el intervalo de tiempo de cada cuadro se vuelve grande y, por tanto, el tiempo que debe retardarse el tiempo de emisión de la imagen decodificada del cuadro no asociado con predicción hacia atrás no se determina de manera exclusiva. Por este motivo, se vuelve inviable manejar correctamente el intervalo de tiempo de emisión entre la imagen decodificada del cuadro no asociado con predicción hacia atrás y la imagen decodificada del cuadro predicho hacia atrás.

En este caso, existen posibles medios en los que un tiempo de retardo grande permisible se permite previamente para la predicción hacia atrás y en los que el tiempo de emisión de la imagen decodificada del cuadro no asociado con predicción hacia atrás se retarda siempre con este tiempo de retardo, manejando por tanto correctamente el intervalo de tiempo de emisión con relación a la imagen decodificada del cuadro predicho hacia atrás. En este caso, sin embargo, el retardo grande se añade siempre al tiempo de emisión de la imagen decodificada, independientemente del tiempo de retardo en la predicción hacia atrás práctica.

Cuando se usan múltiples cuadros de referencia en la predicción hacia atrás como en H.26L, la decodificación de todos los cuadros de referencia que son cuadros temporalmente posteriores debe completarse antes de la decodificación del cuadro actual. Esto aumenta adicionalmente el tiempo de retardo necesario para la ejecución de la predicción hacia atrás.

En este caso, dado que el número de cuadros de referencia usados en la predicción hacia atrás se determina de manera exclusiva como un número de cuadros temporalmente posteriores al cuadro actual, que se decodificaron antes del cuadro actual, el número de cuadros de referencia puede cambiarse opcionalmente dentro del intervalo hasta el límite superior predeterminado del número máximo de cuadros de referencia.

Por ejemplo, suponiendo que el límite superior del número de cuadros de referencia sea 4, el número de cuadros de referencia usados en la predicción hacia atrás puede ser 2 tal como se muestra en la Fig. 8, o 1 tal como se muestra en la Fig. 11A, o 3 tal como se muestra en la Fig. 11B. Dado que el número de cuadros de referencia puede cambiarse de esta manera, el tiempo de retardo necesario para la ejecución de la predicción hacia atrás puede variar ampliamente. Esto da lugar a un fallo a la hora de manejar correctamente el intervalo de tiempo de emisión entre la imagen decodificada del cuadro no asociado con predicción hacia atrás y la imagen decodificada del cuadro predicho hacia atrás.

5 En este momento, dado que el número máximo de cuadros de referencia que puede usarse en la predicción hacia atrás no excede el límite superior del número de cuadros de referencia, el tiempo de retardo según el límite superior del número de cuadros de referencia es un tiempo de retardo máximo que puede darse en la ejecución de la predicción hacia atrás. Por tanto, si el tiempo de emisión de la imagen decodificada del cuadro no asociado con predicción hacia atrás siempre se retarda con este tiempo de retardo, el intervalo de tiempo de emisión con relación a la imagen decodificada del cuadro predicho hacia atrás puede manejarse correctamente.

10 En este caso, sin embargo, siempre se añade un retardo grande al tiempo de emisión de la imagen decodificada, independientemente del número de cuadros de referencia usados realmente para el cuadro predicho hacia atrás. En la aplicación de velocidades de cuadro variables tal como se describió anteriormente, mientras que el número máximo de cuadros de referencia puede determinarse de manera exclusiva, el tiempo de retardo máximo no puede determinarse de manera exclusiva.

15 Hasta ahora, en la aplicación de la predicción hacia atrás para la codificación de vídeo, era inviable determinar de manera exclusiva el tiempo de retardo necesario para la ejecución de la predicción hacia atrás, excepto en el caso en el que el uso de una velocidad de cuadro fija estaba claro. Esto dio como resultado un fallo en el correcto manejo del intervalo de tiempo de emisión entre la imagen decodificada del cuadro no asociado con predicción hacia atrás y la imagen decodificada del cuadro predicho hacia atrás, planteando por tanto el problema de que la emisión de vídeo se volviera forzada.

20 En caso de usar múltiples cuadros de referencia en la predicción hacia atrás, el número de cuadros de referencia también puede cambiarse para variar posiblemente el tiempo de retardo. Por tanto, existe el problema del fallo a la hora de manejar correctamente el intervalo de tiempo entre la imagen decodificada del cuadro no asociado con predicción hacia atrás y la imagen decodificada del cuadro predicho hacia atrás. En caso de que el tiempo de retardo máximo se suponga siempre con el fin de hacer frente a este problema, surge el problema de que siempre se añade el retardo grande al tiempo de emisión de la imagen decodificada.

25 La presente invención se ha llevado a cabo con el fin de solucionar los problemas anteriores, y un objeto de la invención es proporcionar un procedimiento de decodificación de imágenes codificadas capaz de lograr la emisión de imágenes decodificadas en intervalos de tiempo apropiados cuando se emplea una predicción inter-cuadro hacia atrás. El objeto de la invención se logra mediante la reivindicación 1.

35 **Breve descripción de los dibujos**

La Fig. 1 es un diagrama de bloques que muestra la estructura esquemática del aparato de codificación de vídeo, del aparato de decodificación de vídeo y del sistema de procesamiento de vídeo.

40 La Fig. 2 es un diagrama que muestra un ejemplo de codificación de cuadros en caso de que se lleve a cabo la predicción bidireccional.

La Fig. 3 es un diagrama de bloques que muestra un ejemplo de la configuración del aparato de codificación de vídeo.

45 La Fig. 4 es un diagrama de bloques que muestra un ejemplo de la configuración del aparato de decodificación de vídeo.

50 Las Fig. 5A y 5B son diagramas que muestran (A) la decodificación y (B) la emisión de cuadros en caso de que se lleve a cabo la predicción bidireccional mostrada en la Fig. 2.

La Fig. 6 es un diagrama que muestra la codificación de cuadros en caso de que se lleve a cabo la predicción bidireccional.

55 Las Fig. 7A y 7B son diagramas que muestran (A) la decodificación y (B) la emisión de cuadros en caso de que se lleve a cabo la predicción bidireccional mostrada en la Fig. 6.

La Fig. 8 es un diagrama que muestra la codificación de cuadros en caso de que se lleve a cabo la predicción bidireccional.

60 Las Fig. 9A y 9B son diagramas que muestran (A) la decodificación y (B) la emisión de cuadros en caso de que se lleve a cabo la predicción bidireccional mostrada en la Fig. 8.

Las Fig. 10A a 10C son diagramas que muestran (A) la decodificación, (B) la emisión y (C) la emisión retardada de cuadros en caso de que se lleve a cabo la predicción bidireccional.

65 Las Fig. 11A y 11B son diagramas que muestran la codificación de cuadros en caso de que se lleve a cabo la

predicción bidireccional.

### Descripción de las realizaciones preferidas

5 Las realizaciones preferidas del procedimiento de decodificación de imágenes codificadas según la presente invención se describirán a continuación en detalle con referencia a los dibujos. Los mismos elementos se denotarán mediante los mismos símbolos de referencia a lo largo de toda la descripción de los dibujos, sin descripción redundante de los mismos.

10 En primer lugar, se describirá esquemáticamente la codificación y decodificación de una imagen en movimiento en la presente invención. La Fig. 1 es un diagrama de bloques que muestra la estructura esquemática del aparato de codificación de vídeo, del aparato de decodificación de vídeo y del sistema de procesamiento de vídeo según la presente invención. El sistema de procesamiento de vídeo comprende un aparato de codificación de vídeo 1 y un aparato de decodificación de vídeo 2. El aparato de codificación de vídeo 1, el aparato de decodificación de vídeo 2 y el sistema de procesamiento de vídeo se describirán a continuación junto con el procedimiento de codificación de vídeo y el procedimiento de decodificación de vídeo ejecutados en los mismos.

15 El aparato de codificación de vídeo 1 es un dispositivo configurado para codificar datos de vídeo D0 que consisten en una serie de imágenes (cuadros) y emitir datos codificados D1 para la transmisión, el almacenamiento y la regeneración de imágenes en movimiento. El aparato de decodificación de vídeo 2 es un dispositivo configurado para decodificar datos de entrada codificados D1 para generar datos de imagen en movimiento decodificados D2 que consisten en una serie de cuadros. El aparato de codificación de vídeo 1 y el aparato de decodificación de vídeo 2 están conectados mediante una línea de transmisión de datos inalámbrica o por cable predeterminada, con el fin de transmitir datos necesarios tales como los datos codificados D1 y otros.

20 En la codificación de la imagen en movimiento llevada a cabo en el aparato de codificación de vídeo 1, tal como se describió anteriormente, la predicción inter-cuadro se lleva a cabo entre un cuadro de datos de vídeo D0 introducido como un objetivo para la codificación, y otro cuadro como un cuadro de referencia, reduciendo de ese modo la redundancia en los datos de vídeo. En el sistema de procesamiento de vídeo mostrado en la Fig. 1, el aparato de codificación de vídeo 1 lleva a cabo la predicción inter-cuadro hacia atrás a partir de un cuadro temporalmente posterior para la predicción inter-cuadro. Además, este aparato de codificación de vídeo 1 emite el tiempo de retardo máximo que está provocado por la predicción hacia atrás, además de los datos codificados D1.

25 De manera correspondiente a tal aparato de codificación de vídeo 1, el aparato de decodificación de vídeo 2 está configurado para efectuar la entrada del tiempo de retardo máximo que está provocado por la predicción hacia atrás, además de los datos codificados D1 desde el aparato de codificación de vídeo 1. Después, el aparato de decodificación de vídeo 2 decodifica los datos codificados D1 con referencia al tiempo de retardo máximo de entrada para generar los datos de vídeo D2.

30 Mediante el aparato de codificación de vídeo 1 y el procedimiento de codificación de vídeo configurados para emitir el tiempo de retardo máximo, el aparato de decodificación de vídeo 2 y el procedimiento de decodificación de vídeo configurados para efectuar la entrada del tiempo de retardo máximo, y el sistema de procesamiento de vídeo equipados con los aparatos 1, 2, que están adaptados para la predicción inter-cuadro hacia atrás tal como se describió anteriormente, se vuelve viable lograr la emisión de imágenes decodificadas en intervalos de tiempo apropiados en la ejecución de la predicción inter-cuadro usando la predicción inter-cuadro hacia atrás.

35 En cuanto al tiempo de retardo máximo emitido en la codificación de vídeo, por ejemplo, el tiempo de retardo máximo puede definirse como una diferencia de tiempo entre un tiempo de aparición de un cuadro que va a someterse a la predicción inter-cuadro hacia atrás y un tiempo de aparición de un último cuadro temporalmente posterior que puede usarse como un cuadro de referencia para la predicción hacia atrás.

40 En cuanto al tiempo de retardo máximo introducido en la decodificación de vídeo, por ejemplo, el tiempo de retardo máximo (denominado en lo sucesivo  $dpb\_output\_delay$ ) puede definirse como una diferencia de tiempo entre un tiempo de decodificación de un cuadro sin retardo debido a una predicción inter-cuadro hacia atrás y sin inversión de órdenes de tiempos de codificación y de tiempos de emisión con respecto a otro cuadro (el tiempo de decodificación se denominará en lo sucesivo  $Tr$ ) y un tiempo de emisión de imagen decodificada correlacionada con el cuadro pertinente (el tiempo de emisión se denominará en lo sucesivo  $To$ ). En este caso se establece después, preferiblemente, una referencia para los tiempos de emisión de imagen decodificada en función del tiempo de retardo máximo.

45 Una aplicación del tiempo de retardo máximo puede ser un procedimiento que lo aplica a todos los datos codificados o un procedimiento que lo aplica a cada cuadro. Otro procedimiento de aplicación es un procedimiento de aplicación del tiempo de retardo máximo a cada uno de los cuadros posteriores a la comunicación de la información del tiempo de retardo máximo, es decir, al cuadro para el que se indica el tiempo de retardo máximo y a cada uno de los cuadros temporalmente posteriores a ese cuadro. La emisión, introducción, aplicación, etc. del tiempo de retardo máximo en estos procedimientos se detallarán específicamente más adelante.

El procesamiento correspondiente al procedimiento de codificación de vídeo ejecutado en el aparato de codificación de vídeo 1 anterior puede implementarse mediante el programa de codificación de vídeo para permitir que un ordenador ejecute la codificación de vídeo. El procesamiento correspondiente al procedimiento de decodificación de vídeo ejecutado en el aparato de decodificación de vídeo 2 puede implementarse mediante el programa de decodificación de vídeo para permitir que un ordenador ejecute la decodificación de vídeo.

Por ejemplo, el aparato de codificación de vídeo 1 puede estar formado por una CPU conectada a una ROM que almacena programas de software necesarios para operaciones de codificación de vídeo respectivas y a una RAM que guarda temporalmente datos durante la ejecución de un programa. En esta configuración, el aparato de codificación de vídeo 1 puede implementarse permitiendo que la CPU ejecute el programa de codificación de vídeo predeterminado.

De manera similar, el aparato de decodificación de vídeo 2 puede estar formado por una CPU conectada a una ROM que almacena programas de software necesarios para operaciones de decodificación de vídeo respectivas y una RAM que guarda temporalmente datos durante la ejecución de un programa. En esta configuración, el aparato de decodificación de vídeo 2 puede implementarse permitiendo que la CPU ejecute el programa de decodificación de vídeo predeterminado.

El programa expuesto anteriormente para permitir que la CPU ejecute los procesos para la codificación de vídeo o para la decodificación de vídeo puede distribuirse de una forma en la que se graba en un medio de grabación legible por ordenador. Tales medios de grabación incluyen, por ejemplo, medios magnéticos tales como discos duros y discos flexibles, medios ópticos tales como CD-ROM y DVD-ROM, medios magnetoópticos tales como discos flópticos, o dispositivos de hardware, por ejemplo, tales como RAM, ROM y memorias no volátiles semiconductoras, montadas especialmente para ejecutar o almacenar comandos de programas.

Se describirán con realizaciones específicas el aparato de codificación de vídeo, el aparato de decodificación de vídeo, el sistema de procesamiento de vídeo dotado de los mismos mostrado en la Fig. 1, y el procedimiento de codificación de vídeo y el procedimiento de decodificación de vídeo correspondientes a los mismos. La siguiente descripción se basará en la suposición de que las operaciones de codificación y decodificación de vídeo en movimiento se implementan en función de H.26L, y las partes no descritas específicamente acerca del funcionamiento en la codificación de vídeo serán conformes al funcionamiento en H.26L. Sin embargo, ha de observarse que la presente invención no está limitada a H.26L.

(Primera realización)

En primer lugar, se describirá la primera realización de la presente invención. La presente realización describirá una forma realizada de codificación a una velocidad de cuadro fija. En la codificación según la presente realización, se determina en primer lugar el número máximo de cuadros de referencia usado para una predicción hacia atrás, después se calcula el tiempo de retardo máximo a partir de este número máximo de cuadros de referencia y la velocidad de cuadro usada en la codificación, y después se emite el tiempo de retardo máximo. En la decodificación según la presente realización, con motivo de la decodificación de un cuadro no asociado con predicción hacia atrás, un tiempo de emisión de una imagen decodificada del mismo se retarda con el tiempo de retardo máximo de entrada. Después, el tiempo de retardo para el tiempo de emisión se aplica uniformemente a cada cuadro para impedir que el intervalo de tiempo de emisión entre la imagen decodificada del cuadro no asociado con predicción hacia atrás y la imagen decodificada del cuadro predicho hacia atrás se desvíe del intervalo original.

En la codificación, dado que el límite superior del número de cuadros de referencia usados se determina previamente, el número máximo de cuadros de referencia usados para la predicción hacia atrás se determina en primer lugar dentro del intervalo que no excede el límite superior. Después, en función de la velocidad de cuadro usada para la codificación, que también se determina previamente, el tiempo de retardo máximo se calcula como un intervalo de tiempo de un cuadro o dos o más cuadros según el número máximo de cuadros de referencia usados para la predicción hacia atrás.

La Fig. 2 es un diagrama que muestra un ejemplo de codificación de un cuadro en la ejecución de la predicción bidireccional. En este caso, esta Fig. 2 muestra el ejemplo en el que los cuadros de referencia usados para el cuadro actual F2 son dos cuadros temporalmente anteriores F0, F1 antes del cuadro actual F2 y dos cuadros temporalmente posteriores F3, F4 después del cuadro actual F2.

En caso de que el número máximo de cuadros de referencia usados para la predicción hacia atrás sea de 2 y de que la velocidad de cuadro sea de 15 cuadros/segundo, tal como se muestra en la Fig. 2, el intervalo de tiempo de un cuadro es de 1/15 de segundos. En este caso, por tanto, el tiempo de retardo máximo es de  $2 \times (1/15) = 2/15$  segundos.

En la operación de codificación, la codificación de cada cuadro se controla en lo sucesivo de modo que no lleva a cabo la predicción hacia atrás que requiere un tiempo de retardo por encima del tiempo de retardo máximo.

Específicamente, se controla una secuencia de codificación de cuadros de modo que ningún cuadro de referencia usado en la predicción hacia atrás, es decir, ningún cuadro temporalmente posterior después del cuadro actual, no se codifica ni se emite antes del cuadro actual por encima del número máximo de cuadros de referencia usados en la predicción hacia atrás.

5 La Fig. 3 es un diagrama de bloques que muestra un ejemplo de la configuración del aparato de codificación de vídeo usado en la presente realización. El aparato de codificación de vídeo 1 mostrado en la Fig. 3 comprende un codificador 10 para codificar un cuadro (imagen) mediante el procedimiento predeterminado, un controlador (CPU) 15 para controlar operaciones de partes respectivas en el aparato de codificación 1, una memoria de cuadros 11 dispuesta entre un terminal de entrada 1a y un codificador 10, y un multiplexor 12 dispuesto entre un terminal de salida 1b y un codificador 10. El controlador 15 tiene un dispositivo de cálculo de tiempo de retardo máximo 16 para calcular el tiempo de retardo máximo, como una función del mismo. El codificador 10 está dotado de una memoria intermedia de salida 13.

15 En la codificación de vídeo en el presente aparato de codificación 1, se introducen condiciones para la codificación de vídeo a través de un terminal de entrada 1c. En esta introducción de las condiciones, las condiciones de codificación se seleccionan o se introducen generalmente a través de un dispositivo de entrada tal como un teclado. En la presente realización, específicamente, las condiciones de codificación introducidas incluyen el tamaño de un cuadro como un objetivo para la codificación, la velocidad de cuadro y la tasa de bit y, además de éstas, las condiciones de codificación incluyen también una estructura de referencia predictiva del vídeo (si se aplica predicción hacia atrás), el número de cuadros almacenados temporalmente y usados como cuadros de referencia (correspondiente a la capacidad de la memoria intermedia de salida 13), y el número de cuadros de referencia usados en la predicción hacia atrás. Estas condiciones pueden establecerse de modo que varíen con el tiempo. Las condiciones de codificación introducidas a través del terminal de entrada 1c se almacenan en el controlador 15.

25 Con el inicio de la operación de codificación, el controlador 15 envía las condiciones de codificación al codificador 10, en el que se establece la condición de codificación. Por otro lado, se introduce un cuadro como un objeto codificado a través del terminal de entrada 1a y se alimenta a través de la memoria de cuadros 11 al codificador 10 para codificarse en el mismo. El cuadro de entrada se guarda temporalmente en la memoria de cuadros 11, puesto que el orden de los cuadros se cambia para la ejecución de la predicción hacia atrás. Por ejemplo, en el ejemplo mostrado en la Fig. 2, se introduce el cuadro F2 a través del terminal de entrada 1a antes de los cuadros F3, F4, pero se codifica después de los cuadros F3, F4; por tanto, el cuadro F2 se guarda temporalmente en la memoria de cuadros 11.

35 El codificador 10 codifica el cuadro basándose en el algoritmo de H.26L. Después, los datos codificados se alimentan al multiplexor 12 para multiplexarse con otra información relacionada y, posteriormente, los datos multiplexados se emiten a través del terminal de salida 1b. El cuadro usado para la predicción se reproduce en el codificador 10 y se almacena como un cuadro de referencia para la codificación del siguiente cuadro en la memoria intermedia 13.

40 En la presente realización, el dispositivo de cálculo de tiempo de retardo máximo 16 del controlador 15 calcula el tiempo de retardo máximo `dpb_output_delay`, basándose en el número de cuadros de referencia y la velocidad de cuadro introducida a través del terminal de entrada 1c y usada para la predicción hacia atrás. Después, el multiplexor 12 añade el tiempo de retardo máximo a datos de imagen codificados. Además, se añade también un identificador (N) que indica un turno de visualización para la identificación de cada cuadro, junto con los datos codificados de cada cuadro.

45 Es evidente que cuando no se aplica la predicción hacia atrás, el número de cuadros de referencia usados es cero y, por tanto, el valor de `dpb_output_delay` es cero.

50 En la presente realización, se supone que una sintaxis para transmitir el tiempo de retardo máximo se añade a la sintaxis de datos codificados en H.26L con el fin de implementar la emisión del tiempo de retardo máximo en la codificación y la introducción del tiempo de retardo máximo en la decodificación. En este ejemplo se añade la nueva sintaxis al conjunto de parámetros de secuencia, que es una sintaxis para transmitir la información que va a aplicarse a todos los datos codificados.

55 El parámetro `dpb_output_delay` se define como una sintaxis para transportar el tiempo de retardo máximo. En este caso, se supone que el parámetro `dpb_output_delay` usa la misma unidad de tiempo usada en las otras sintaxis que indica el tiempo en H.26L y que indica el tiempo de retardo máximo en la unidad de tiempo de 90 kHz. Una cifra indicada en la unidad de tiempo se codifica y se transmite mediante un código de 32 bits de longitud fija sin signo. Por ejemplo, cuando el tiempo de retardo máximo es de 2/15 segundos tal como se describió anteriormente, `dpb_output_delay` es  $(2/15) \times 90000 = 12000$ .

60 En la operación de decodificación se decodifica el tiempo de retardo máximo transportado por `dpb_output_delay`, y el tiempo de emisión de una imagen decodificada se retarda usándolo.



La Fig. 4 es un diagrama de bloques que muestra un ejemplo de la configuración del aparato de decodificación de vídeo usado en la presente realización. El aparato de decodificación de vídeo 2 mostrado en la Fig. 4 comprende un decodificador 20 para decodificar datos codificados para generar una imagen regenerada, un controlador (CPU) 25 para controlar operaciones de partes respectivas en el aparato de decodificación 2, una memoria intermedia de entrada 21 dispuesta entre un terminal de entrada 2a y un decodificador 20, y una memoria intermedia de salida 22 dispuesta entre un terminal de salida 2b y un decodificador 20. El controlador 25 tiene un dispositivo de cálculo de tiempo de emisión de imagen 26 para calcular el tiempo de emisión para la visualización de un cuadro, como una función del mismo.

En la decodificación de vídeo en el presente aparato de decodificación 2, se introducen datos como un objeto decodificado a través del terminal de entrada 2a. Estos datos son datos multiplexados de los datos codificados de cada cuadro codificado por el aparato de codificación 1 mostrado en la Fig. 3, donde el tiempo de retardo máximo  $dpb\_output\_delay$  y el identificador (N) indican el turno de visualización de cada cuadro.

Los datos de entrada se almacenan en la memoria intermedia de entrada 21. Cuando un comando del controlador 25 indica la llegada de un tiempo de decodificación, se introducen datos de un cuadro desde la memoria intermedia de entrada 21 en el decodificador 20 y, a continuación, se decodifican según el algoritmo de H.26L. El cuadro regenerado de esta manera se almacena en la memoria intermedia de salida 22. El cuadro en la memoria intermedia de salida 22 se alimenta de nuevo a través de una línea 23 al decodificador 20 para usarse como un cuadro de referencia para la decodificación del siguiente cuadro.

Por otro lado, el tiempo de retardo máximo  $dpb\_output\_delay$ , la velocidad de cuadro y el identificador (N) de cada cuadro decodificado en el decodificador 20 se alimentan en el controlador 25. Después, el dispositivo de cálculo de tiempo de emisión de imagen 26 del controlador 25 calcula el tiempo de emisión de cada cuadro a partir de estos datos según la siguiente ecuación.

$$To(n) = dpb\_output\_delay + N \times \text{intervalo de cuadro}$$

En esta ecuación, el intervalo de cuadro se determina a partir de la velocidad de cuadro.

Suponiendo que  $dpb\_output\_delay$  es de 2/15 segundos y que el intervalo de cuadro es de 1/15 segundos como en el ejemplo mostrado en la Fig. 2, los tiempos de emisión de los respectivos cuadros se calculan de la siguiente manera según la ecuación anterior.

$$N = 0, To(0) = 2/15$$

$$N = 1, To(1) = 3/15$$

$$N = 2, To(2) = 4/15$$

$$N = 3, To(3) = 5/15$$

Según los tiempos de emisión  $To(n)$  obtenidos de esta manera por el controlador 25, las cuadros en la memoria intermedia de salida 22 se emiten a intervalos constantes al terminal de salida 2b, tal como se indica mediante los cuadros F0, F1, F2 y F3 mostrados en la Fig. 5B. Aunque no se ilustra, el terminal de salida 2b está conectado a un dispositivo de visualización tal como un monitor.

Las Fig. 5A y 5B son diagramas que muestran (A) la decodificación y (B) la emisión de los cuadros en el caso de la predicción bidireccional mostrada en la Fig. 2. Se supone que en la operación de decodificación los datos codificados de los cuadros se decodifican en el orden necesario para la ejecución de la predicción inter-cuadro, los intervalos de los mismos son intervalos de tiempo constantes según la velocidad de cuadro, y el tiempo necesario para la operación de decodificación es despreciable para cada cuadro, independientemente de si se aplica la predicción inter-cuadro e independientemente de los sentidos de predicción inter-cuadro. En este caso, el tiempo de retardo máximo necesario para la ejecución de la predicción hacia atrás en el cuadro predicho hacia atrás es igual a un intervalo de tiempo de un cuadro o cuadros según el número máximo de cuadros de referencia usados para la predicción hacia atrás. Este tiempo se transporta como un tiempo de retardo máximo mediante  $dpb\_output\_delay$ . Por consiguiente, para emitir una imagen decodificada, un tiempo de emisión de la misma se retarda con el tiempo de retardo máximo.

En la práctica, los intervalos de decodificación de los respectivos cuadros no son constantes, y pueden variar dependiendo de factores tales como la variación en bits de codificación de los cuadros. El tiempo necesario para la operación de decodificación de cada cuadro puede variar también en función de si el cuadro es un cuadro predicho hacia atrás o en función de los bits de codificación de cada cuadro.

Para retardar el tiempo de emisión, por tanto, se establece la referencia en el tiempo en el que se obtiene la imagen decodificada para el cuadro no asociado con predicción hacia atrás F0 sin retardo debido a una predicción hacia atrás y sin inversión de órdenes de tiempos de decodificación y tiempos de emisión con respecto a cualquier otro cuadro, tal como se muestra en las Fig. 5A y 5B. Concretamente, un tiempo obtenido retardando el tiempo en el que se obtiene la imagen decodificada con el tiempo de retardo máximo comunicado mediante `dpb_output_delay` se define como un tiempo igual al tiempo de emisión correlacionado con esta imagen decodificada, y se usa como un tiempo de referencia en la emisión de imágenes decodificadas. Después, las imágenes decodificadas F1-F4 se emiten cuando este tiempo de referencia coincide con un tiempo igual a un tiempo de emisión correlacionado con cada imagen decodificada.

Por ejemplo, cuando el tiempo de retardo máximo es de 2/15 segundos tal como se describió anteriormente, un tiempo en un retardo de 2/15 segundos desde el tiempo en que se obtiene la imagen decodificada para el cuadro no asociado con predicción hacia atrás, se define como un tiempo igual al tiempo de emisión correlacionado con esta imagen decodificada y, después, se usa como un tiempo de referencia en la emisión de imágenes decodificadas.

Según las circunstancias, posiblemente, el tiempo de retardo máximo no se comunica a propósito, con el fin de simplificar la operación de codificación o decodificación. En tales casos, la sintaxis para comunicar el tiempo de retardo máximo puede disponerse para poder omitirse en la suposición de que una marca para indicar la presencia o ausencia de la sintaxis se transmita antes que la sintaxis para transmitir el tiempo de retardo máximo.

En caso de que se omita la comunicación del tiempo de retardo máximo, la operación de codificación puede estipularse previamente, por ejemplo, para no usar la predicción hacia atrás o de modo que el número de cuadros de referencia usados en la predicción hacia atrás pueda alterarse opcionalmente dentro del intervalo que no excede el límite superior del número de cuadros de referencia.

La operación de decodificación puede configurarse para realizarse de conformidad con lo estipulado en la operación de codificación; por ejemplo, cuando no se aplica la predicción hacia atrás, no se produce ningún retardo necesario para la ejecución de predicción hacia atrás; como alternativa, la operación de decodificación puede configurarse también de modo que el número de cuadros de referencia usados en la predicción hacia atrás pueda alterarse opcionalmente dentro del intervalo que no excede el límite superior del número de cuadros de referencia, es decir, el tiempo de retardo puede variar ampliamente. En este caso, la operación de decodificación puede configurarse para realizar siempre un procesamiento que asuma un tiempo de retardo máximo esperado, o la operación de decodificación puede configurarse para permitir una variación de intervalos de tiempo de emisión de imágenes decodificadas y realizar un procesamiento simplificado sin tomar en consideración el tiempo de retardo de cada cuadro.

La presente realización se ha descrito suponiendo que las operaciones se han implementado según la H.26L, pero debe observarse que los procedimientos de codificación de vídeo a los que puede aplicarse la presente invención no están limitados a H.26L y que la presente invención puede aplicarse a diversos procedimientos de codificación de vídeo usando la predicción inter-cuadro hacia atrás.

En la presente realización, la sintaxis por códigos de longitud fija se ha añadido como una sintaxis para transmitir el tiempo de retardo máximo en el conjunto de parámetros de secuencia, pero debe observarse que los códigos y la sintaxis para transmitirlo, o la unidad de tiempo para expresar el tiempo de retardo máximo, no se limitan, por supuesto, a esto. Los códigos de longitud fija pueden sustituirse por códigos de longitud variable, y el tiempo de retardo máximo puede transmitirse mediante cualquiera de las diversas sintaxis que pueden transportar información que va a aplicarse a todos los datos codificados.

Por ejemplo, en H.26L puede añadirse una sintaxis en un mensaje de información de mejora complementario. En caso de usar otro procedimiento de codificación de vídeo, el tiempo de retardo máximo puede transmitirse mediante una sintaxis para transmitir la información que va a aplicarse a todos los datos codificados en el procedimiento de codificación pertinente. En otro caso, el tiempo de retardo máximo puede transmitirse también fuera de los datos codificados en el procedimiento de codificación de vídeo como en la recomendación UIT-T H.245 usada para transportar información de control en comunicación usando H.263.

(Segunda realización)

A continuación se describirá la segunda realización de la presente invención. La presente realización describirá una forma realizada de codificación a velocidades de cuadro variables. Las operaciones en la codificación y decodificación según la presente realización son muy similares a las de la primera realización. Dado que la presente realización usa velocidades de cuadro variables, esto implica una operación a bajas velocidades de cuadro para evitar la ejecución de la predicción hacia atrás que requiere el tiempo de retardo por encima del tiempo de retardo máximo calculado previamente, además de la operación de codificación en la primera realización, de modo que se impide que el intervalo de tiempo de emisión entre la imagen decodificada del cuadro no asociado con predicción hacia atrás y la imagen decodificada del cuadro predicho hacia atrás se desvíe del intervalo original incluso con una variación de las velocidades de cuadro.

5 Dado que en la operación de codificación el límite superior del número de cuadros de referencia se determina previamente, el número máximo de cuadros de referencia usados para la predicción hacia atrás se determina en primer lugar dentro del intervalo que no excede el límite superior. Entonces, el intervalo de tiempo de cuadro máximo se determina en función de una velocidad de cuadro objetivo determinada previamente en un control de tasas de bits de codificación, y el tiempo de retardo máximo se calcula como un intervalo de tiempo de un cuadro o dos o más cuadros según el número máximo de cuadros de referencia usados en la predicción hacia atrás y el intervalo de tiempo de cuadro máximo.

10 En la operación de codificación, la codificación de cada cuadro se controla después para evitar la predicción hacia atrás que requiere el tiempo de retardo más allá del tiempo de retardo máximo. Específicamente, el orden de codificación de cuadros se controla para impedir que cualquier cuadro de referencia usado en la predicción hacia atrás, es decir, cualquier cuadro temporalmente posterior después del cuadro actual, que va más allá del número máximo de cuadros de referencia usados en la predicción hacia atrás, se codifique y se emita antes que el cuadro actual.

15 Además, cuando la velocidad de cuadro de codificación se vuelve temporalmente pequeña a causa del control de tasas de bits de codificación, para hacer que el intervalo de tiempo de cuadro en ese caso sea mayor que el intervalo de tiempo de cuadro máximo, se controla la codificación de cada cuadro de modo que no se aplique la predicción hacia atrás para codificar ahí el cuadro.

20 La presente realización es sustancialmente idéntica a la primera realización en que el tiempo de retardo máximo se emite en la codificación, en que la sintaxis `dpb_output_delay` para transmitir el tiempo de retardo máximo se añade a la sintaxis de datos codificados con el fin de efectuar la entrada de la misma en la decodificación, y en la definición de la sintaxis.

25 En la presente realización, la operación de decodificación se dispone para decodificar el tiempo de retardo máximo comunicado mediante `dpb_output_delay` y el retardo del tiempo de emisión de la imagen decodificada mediante su uso. Este procesamiento es igual también al de la primera realización.

30 (Tercera realización)

35 A continuación se describirá la tercera realización de la presente invención. La presente realización describirá una forma realizada en la que el tiempo de retardo máximo se comunica opcionalmente para cada cuadro y, por tanto, puede cambiarse con flexibilidad. Las operaciones en la codificación y decodificación según la presente realización son básicamente similares a las de la primera realización o la segunda realización.

40 En la presente realización, la sintaxis `dpb_output_delay` para transmitir el tiempo de retardo máximo, que se definió en la primera realización, está dispuesta para añadirse al conjunto de parámetros de imagen que es una sintaxis para transportar la información aplicada a cada cuadro en lugar de la sintaxis para transportar la información aplicada a todos los datos codificados. La sintaxis `dpb_output_delay` está configurada en este caso para indicar el tiempo de retardo máximo en la unidad de tiempo de 90 kHz, como en el caso de la primera realización, y una cifra expresada en la unidad de tiempo se codifica y se transmite mediante un código de 32 bits de longitud fija sin signo.

45 La presente realización es muy similar a la primera realización, en cuanto al cálculo del tiempo de retardo máximo en la codificación y en cuanto al retardo del tiempo de emisión de la imagen decodificada mediante el uso del tiempo de retardo máximo en la decodificación. Las configuraciones del aparato de codificación de vídeo y del aparato de decodificación de vídeo usados en la presente realización son muy similares a las mostradas en las Fig. 3 y 4 acerca de la primera realización.

50 Va a explicarse cómo determinar el tiempo de retardo máximo `dpb_output_delay` de cada cuadro en la presente realización. En el aparato de codificación 1 mostrado en la Fig. 3, el controlador 15 calcula el tiempo de retardo (D) debido a la predicción hacia atrás mediante el procedimiento descrito en la primera realización, y determina el tiempo de codificación  $Tr(n)$  de cada cuadro. Cuando se alimenta un tiempo de visualización  $Tin(n)$  de cada cuadro desde la memoria de cuadros 11, se calcula `dpb_output_delay(n)` de ese cuadro tal como sigue.

$$\text{dpb\_output\_delay}(n) = Tin(n) + D - Tr(n)$$

55 Este valor de `dpb_output_delay` está correlacionado con el cuadro pertinente y se multiplexa en el multiplexor 12.

60 En la presente realización, el tiempo  $Tr(n)$  para la codificación de cada cuadro también se codifica conjuntamente. Tomando la Fig. 2 como ejemplo,  $D = 2/15$  segundos, y  $Tin(n) = 0, 1/15, 2/15, 3/15$  o  $4/15$  ( $n = 0, 1, 2, 3$  o  $4$ ). A causa del cambio en el orden de codificación,  $Tr(n)$  se vuelve tal como sigue:  $Tr(n) = 0, 1/15, 4/15, 2/15$  o  $3/15$  ( $n = 0, 1, 2, 3$  o  $4$ ). En este caso, `dpb_output_delay(n)` de cada cuadro se obtiene tal como sigue.

65

$$n = 0, \text{dpb\_output\_delay}(0) \\ = 0 + 2/15 - 0 = 2/15$$

$$n = 1, \text{dpb\_output\_delay}(1) \\ = 1/15 + 2/15 - 1/15 = 2/15$$

$$5 \quad n = 2, \text{dpb\_output\_delay}(2) \\ = 2/15 + 2/15 - 4/15 = 0$$

$$n = 3, \text{dpb\_output\_delay}(3) \\ = 3/15 + 2/15 - 2/15 = 3/15$$

$$10 \quad n = 4, \text{dpb\_output\_delay}(4) \\ = 4/15 + 2/15 - 3/15 = 3/15$$

Por otro lado, en el aparato de decodificación 2 mostrado en la Fig. 4, el decodificador 20 envía  $\text{dpb\_output\_delay}(n)$  y  $\text{Tr}(n)$  de cada cuadro al controlador 25, y el controlador 25 calcula el tiempo de emisión  $\text{To}(n)$  de cada cuadro basándose en la siguiente ecuación.

$$15 \quad \text{To}(n) = \text{Tr}(n) + \text{dpb\_output\_delay}$$

Tomando la Fig. 2 como ejemplo,  $\text{To}(n)$  para cada cuadro se calcula tal como sigue según la definición anterior, en función de  $\text{Tr}(n) = 0, 1/15, 4/15, 2/15$  o  $3/15$  ( $n = 0, 1, 2, 3$  o  $4$ ) y  $\text{dpb\_output\_delay}(n) = 2/15, 2/15, 0, 3/15$  o  $3/15$  ( $n = 0, 1, 2, 3$  o  $4$ ).

$$20 \quad n = 0, \text{To}(0) = 0 + 2/15 = 2/15$$

$$n = 1, \text{To}(1) = 1/15 + 2/15 = 3/15$$

$$25 \quad n = 2, \text{To}(2) = 4/15 + 0 = 4/15$$

$$n = 3, \text{To}(3) = 2/15 + 3/15 = 5/15$$

$$30 \quad n = 4, \text{To}(4) = 3/15 + 3/15 = 6/15$$

Concretamente, todas las imágenes se visualizan con el retardo de  $2/15$  segundos y a intervalos constantes en el monitor. Es evidente que cuando no se aplica la predicción hacia atrás, el número de cuadros de referencia usados para la misma es cero y el valor de  $\text{dpb\_output\_delay}(n)$  es, por tanto, cero.

35 Dado que el tiempo de retardo máximo define el tiempo de referencia en la emisión de imágenes decodificadas a partir del tiempo en que se adquiere la imagen decodificada del cuadro no asociado con predicción hacia atrás, basta con transmitir el tiempo de retardo máximo solamente para el cuadro no asociado con predicción hacia atrás. Por tanto, es posible emplear, por ejemplo, una configuración en la que la sintaxis para transmitir el tiempo de retardo máximo esté dispuesta para poder omitirse en la suposición de que una marca que indica la presencia o ausencia de la sintaxis se transmita antes de la misma. La sintaxis puede estar dispuesta también para omitirse opcionalmente para el cuadro no asociado con predicción hacia atrás, siempre que el tiempo de retardo máximo transmitido antes se aplique en ese caso en el que el tiempo de retardo máximo no se transmite.

45 La sintaxis para cada cuadro en la presente realización puede usarse también simultáneamente con la sintaxis para todos los datos codificados tal como se definió en la primera realización. En este caso, la sintaxis para cada cuadro puede omitirse, siempre que una marca que indica la presencia o ausencia de la sintaxis se transmita antes de la misma tal como se describió anteriormente. El tiempo de retardo máximo transmitido en la sintaxis para todos los datos codificados se aplica de manera continua antes de que se transmita el tiempo de retardo máximo en la sintaxis para cada cuadro. Después de que se actualice mediante la sintaxis para cada cuadro, el tiempo retardado basado en la misma se usa después como un tiempo de referencia en la emisión de cada imagen decodificada.

55 La presente realización se ha descrito suponiendo que estaba basada sustancialmente en H.26L, pero debe observarse que los procedimientos de codificación de vídeo a los que puede aplicarse la presente invención no están limitados a H.26L, y que la presente invención puede aplicarse a diversos procedimientos de codificación de vídeo usando la predicción inter-cuadro hacia atrás.

En la presente realización, la sintaxis para transmitir el tiempo de retardo máximo era la sintaxis por códigos de longitud fija añadidos en el conjunto de parámetros de imagen, y es evidente que los códigos y la sintaxis para transmitirlo, o la unidad de tiempo para expresar el tiempo de retardo máximo no se limitan, por supuesto, a esto. Los códigos de longitud fija pueden sustituirse por códigos de longitud variable, y el tiempo de retardo máximo puede comunicarse en cualquiera de diversas sintaxis que sean capaces de comunicar la información que va a aplicarse a cada cuadro.

Por ejemplo, la sintaxis puede añadirse en un mensaje de información de mejora complementario en H.26L. Cuando se aplica otro procedimiento de codificación de vídeo, es posible usar una sintaxis para comunicar información que va a aplicarse a cada cuadro en el procedimiento de codificación pertinente. Además, la información puede comunicarse también fuera de los datos codificados en el procedimiento de codificación de vídeo, como en la recomendación UIT-T H.245 usada para la comunicación de información de control en comunicación usando la H.263.

El procedimiento de codificación de vídeo, el procedimiento de decodificación de vídeo, el aparato de codificación de vídeo, el aparato de decodificación de vídeo, el sistema de procesamiento de vídeo, el programa de codificación de vídeo y el programa de decodificación de vídeo según la presente invención proporcionan el efecto siguiente, tal como se describió anteriormente. Concretamente, cuando una imagen en movimiento que consiste en una serie de cuadros se codifica mediante la predicción inter-cuadro hacia atrás para su emisión, se vuelve viable lograr la emisión de imágenes decodificadas en intervalos de tiempo apropiados al emplear la predicción inter-cuadro hacia atrás, mediante el procedimiento de codificación de vídeo, el aparato de codificación y el programa de codificación configurados para emitir el tiempo de retardo máximo debido a la predicción hacia atrás, el procedimiento de decodificación de vídeo, el aparato de decodificación y el programa de decodificación configurados para efectuar la entrada del tiempo de retardo máximo, y el sistema de procesamiento de vídeo que los usa.

Particularmente, a diferencia de la técnica anterior, los tiempos de emisión no son valores absolutos, sino valores relativos del tiempo de decodificación  $T_r$ ; por tanto, la invención proporciona el efecto de tener la capacidad de describir y transmitir de manera apropiada el valor del tiempo de retardo máximo  $dpb\_output\_delay$  mediante un pequeño número de bits, incluso en caso de que la velocidad de cuadro sea variable. Incluso si el tiempo de decodificación  $T_r$  tiene una desfase o no se recibe, se emitirá una imagen correspondiente con un retardo de  $dpb\_output\_delay$  desde el tiempo de finalización de la decodificación, presentando por tanto la ventaja de que las imágenes se emiten a intervalos correctos.

**REIVINDICACIONES**

1. Procedimiento de decodificación de imágenes codificadas, comprendiendo el procedimiento:
  - 5 recibir una imagen codificada (n), una temporización de decodificación  $Tr(n)$  que representa una temporización según la cual va a decodificarse la imagen codificada (n), y un tiempo de retardo máximo ( $dpb\_output\_delay(n)$ ) que representa un periodo de tiempo según el cual va a retardarse la emisión de la imagen codificada (n) para su visualización después de que la imagen codificada (n) se decodifica según la temporización de decodificación  $Tr(n)$ ;
  - 10 decodificar la imagen codificada (n) según la temporización de decodificación  $Tr(n)$  y derivar una imagen decodificada (n);
  - 15 calcular, para un cuadro no asociado con predicción hacia atrás ( $F0$ ), una temporización de emisión  $To(n)$  del cuadro no asociado con predicción hacia atrás ( $F0$ ) de tal manera que se añada la temporización de decodificación  $Tr(n)$  con el tiempo de retardo máximo ( $dpb\_output\_delay(n)$ ), en el que la temporización de emisión  $To(n) =$  la temporización de decodificación  $Tr(n) +$  el tiempo de retardo máximo ( $dpb\_output\_delay(n)$ );
  - 20 evitar que la imagen decodificada (n) se emita para su visualización desde una memoria intermedia de imágenes decodificadas (22) durante un periodo de tiempo representado por el tiempo de retardo máximo ( $dpb\_output\_delay(n)$ ) después de que la imagen decodificada (n) se decodifica según la temporización de decodificación  $Tr(n)$  y emitir la imagen decodificada (n) desde la memoria intermedia de imágenes decodificadas (22) según la temporización de emisión  $To(n)$ ,
  - 25 en el que la memoria intermedia de imágenes decodificadas (22) contiene al menos una imagen decodificada usada como referencia para decodificar al menos una de las imágenes codificadas.

**Fig.1**

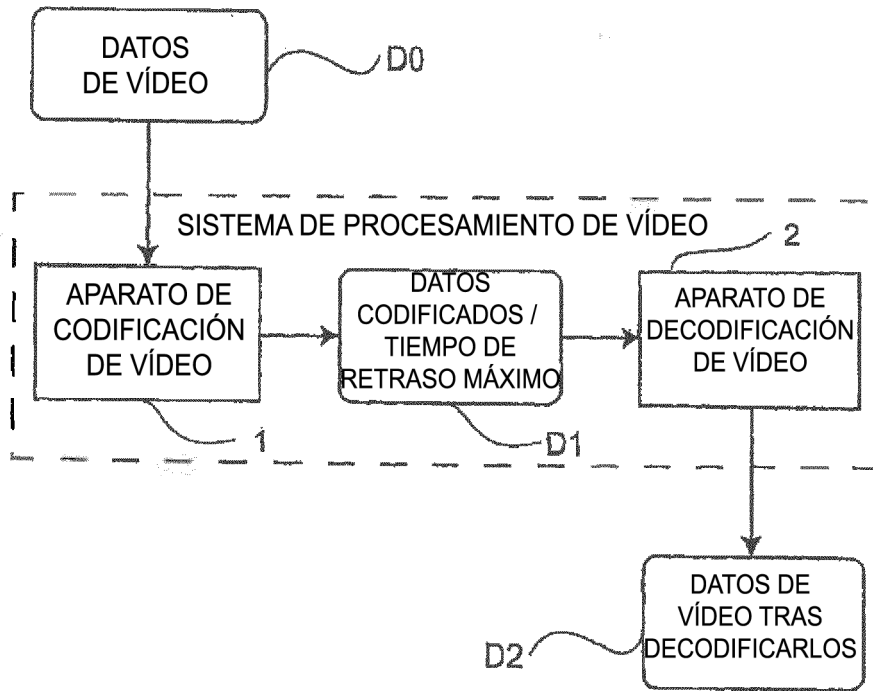


Fig.2

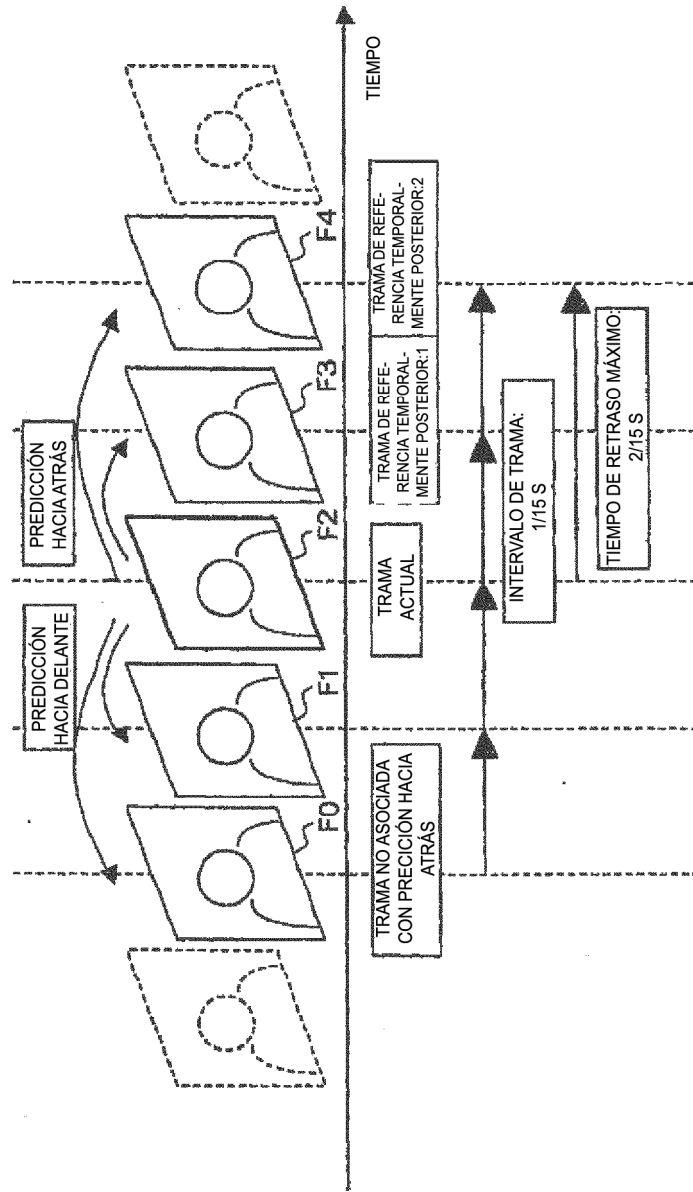




Fig.3

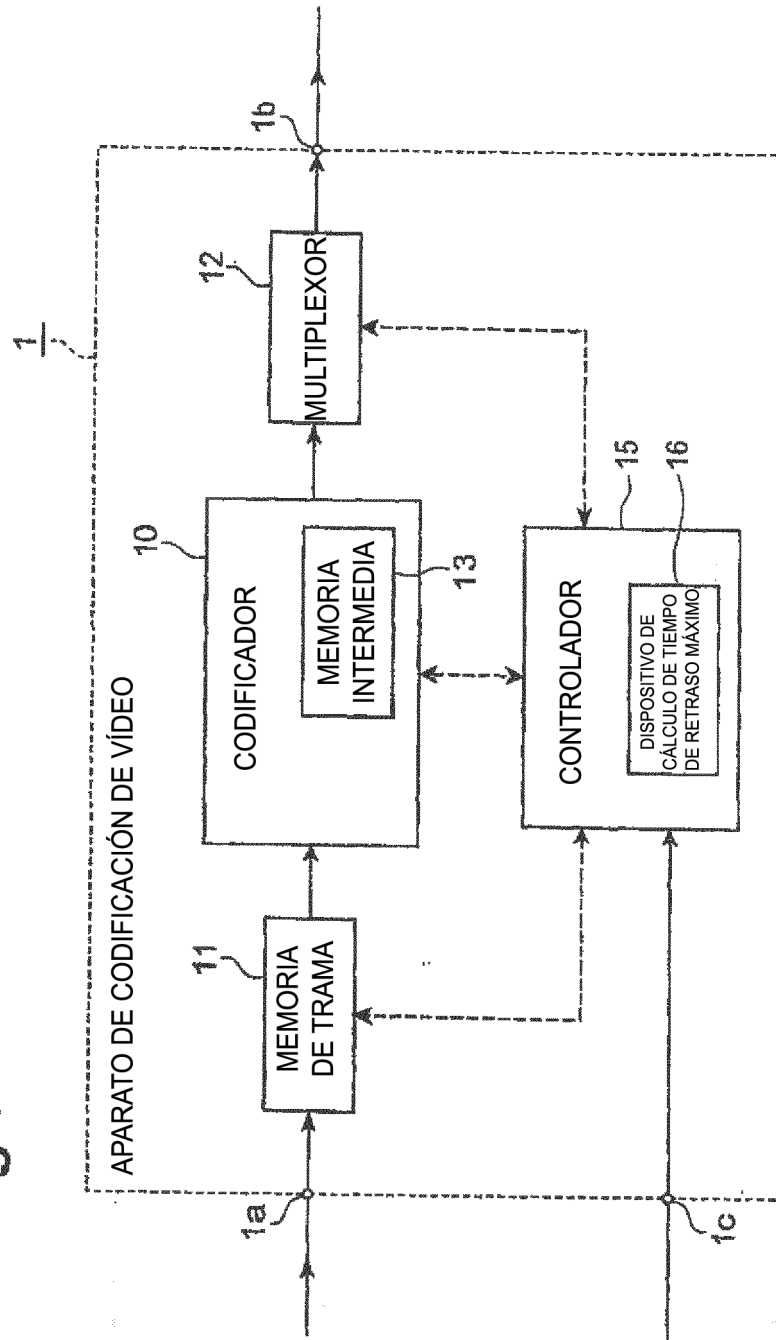
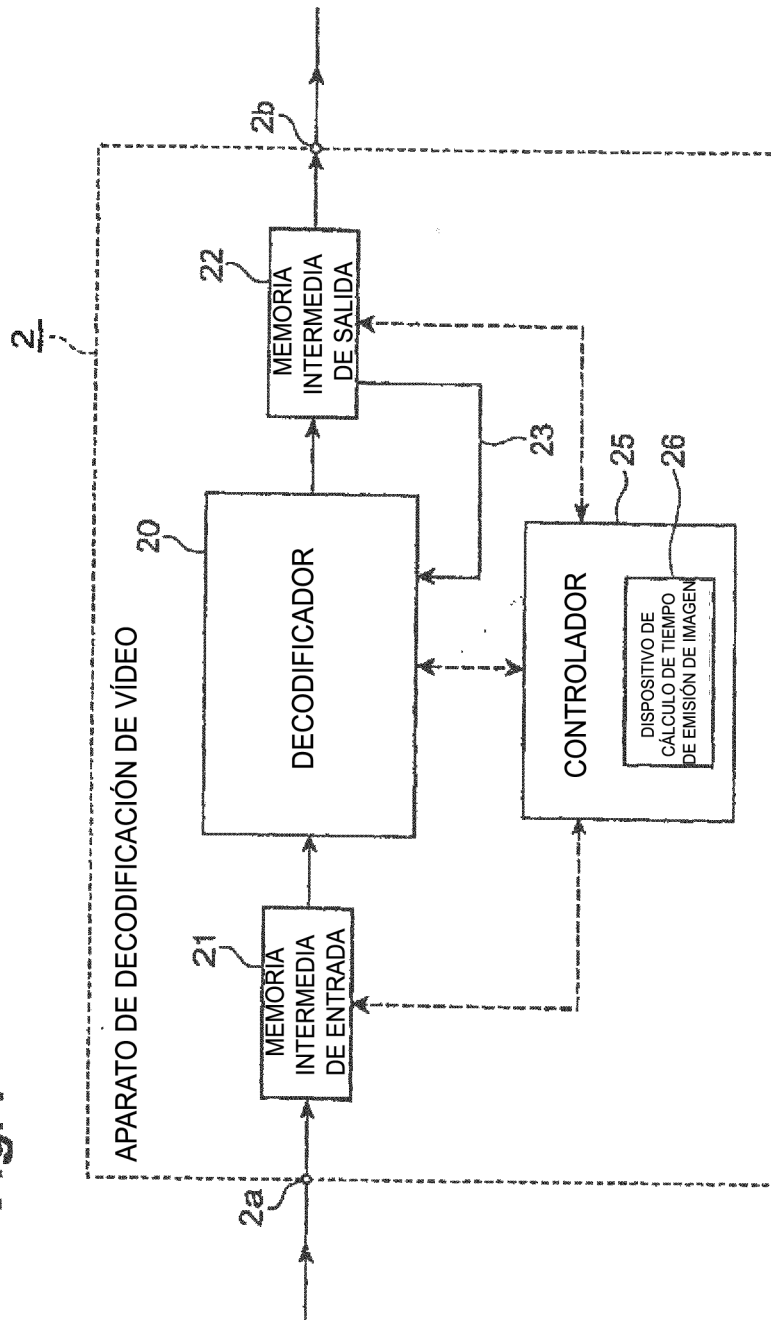
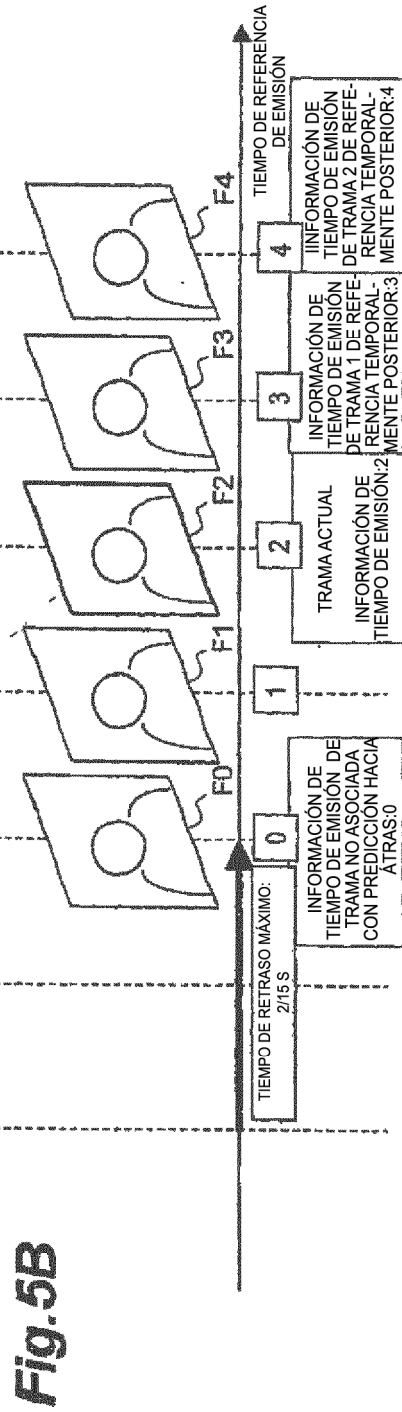
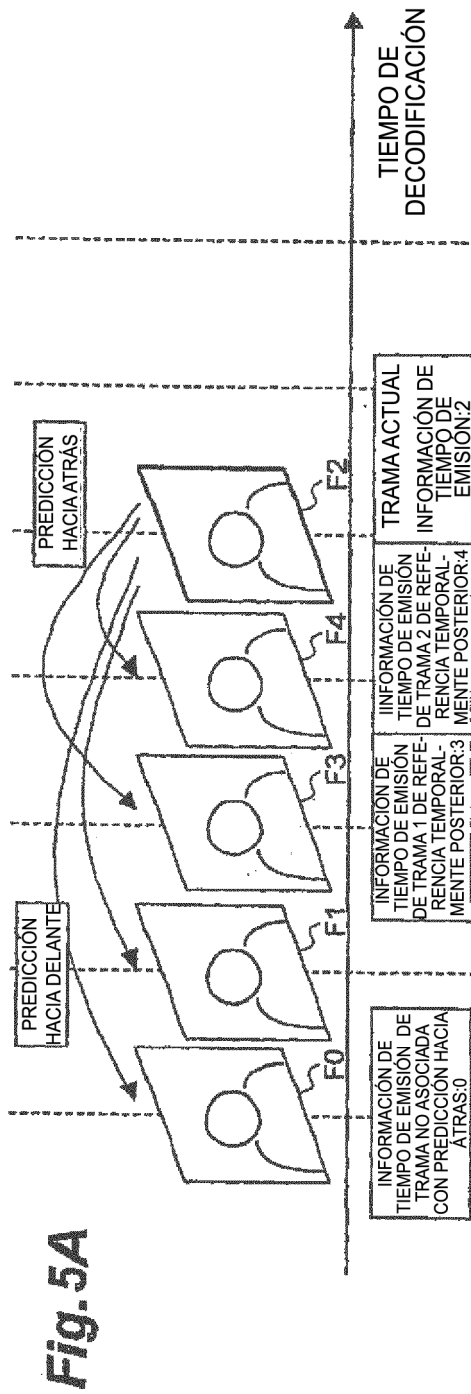
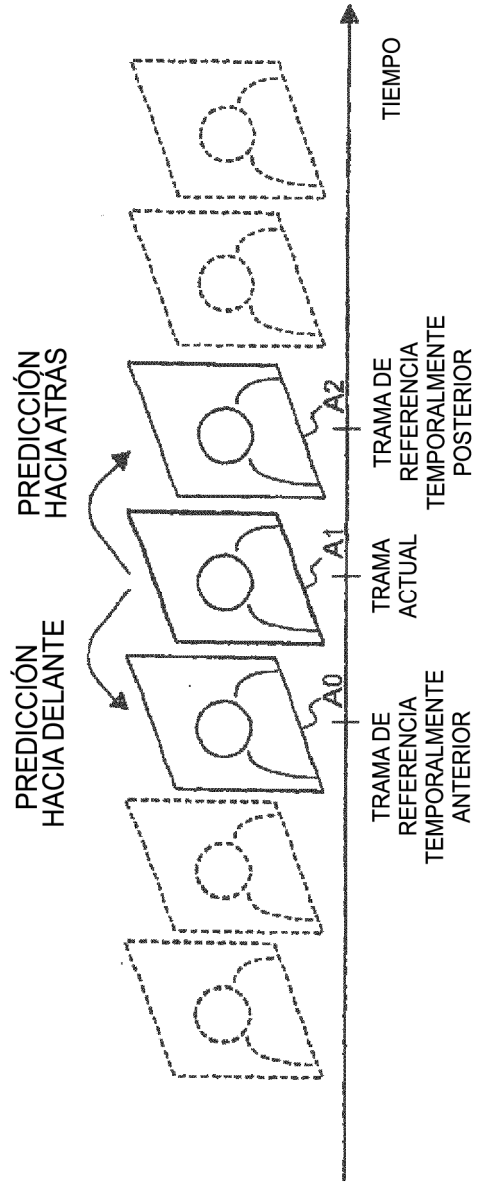


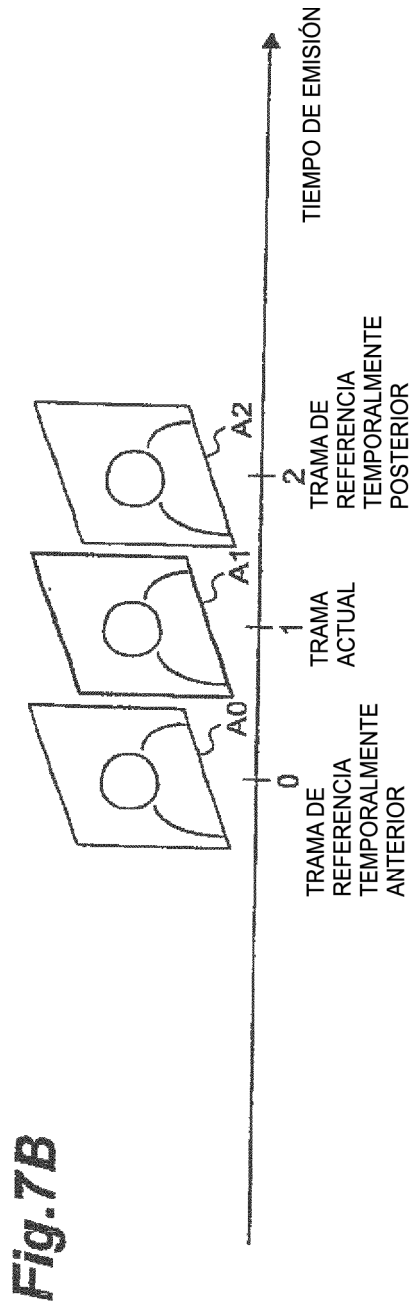
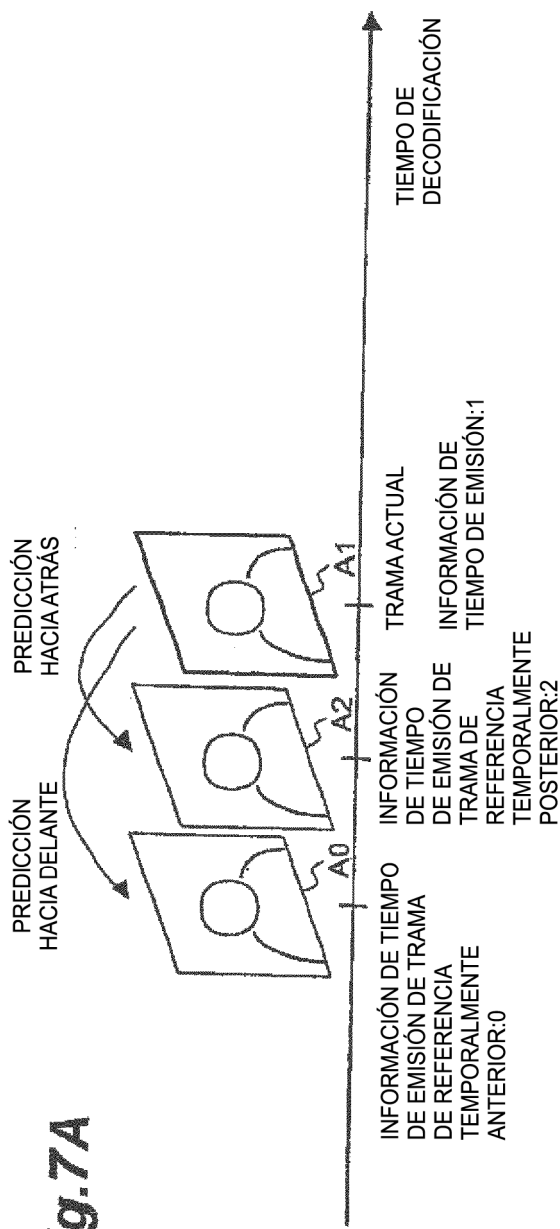
Fig.4



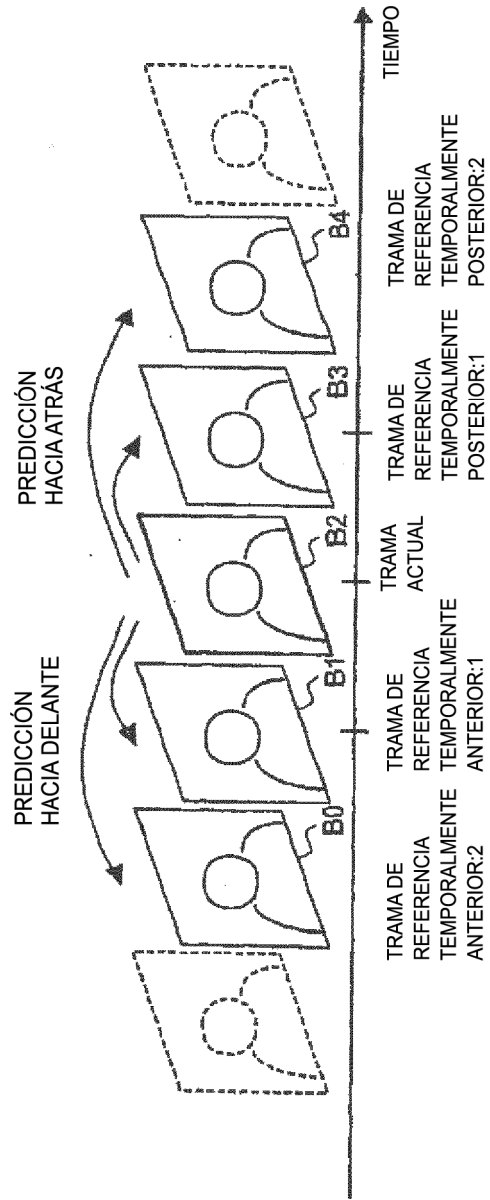


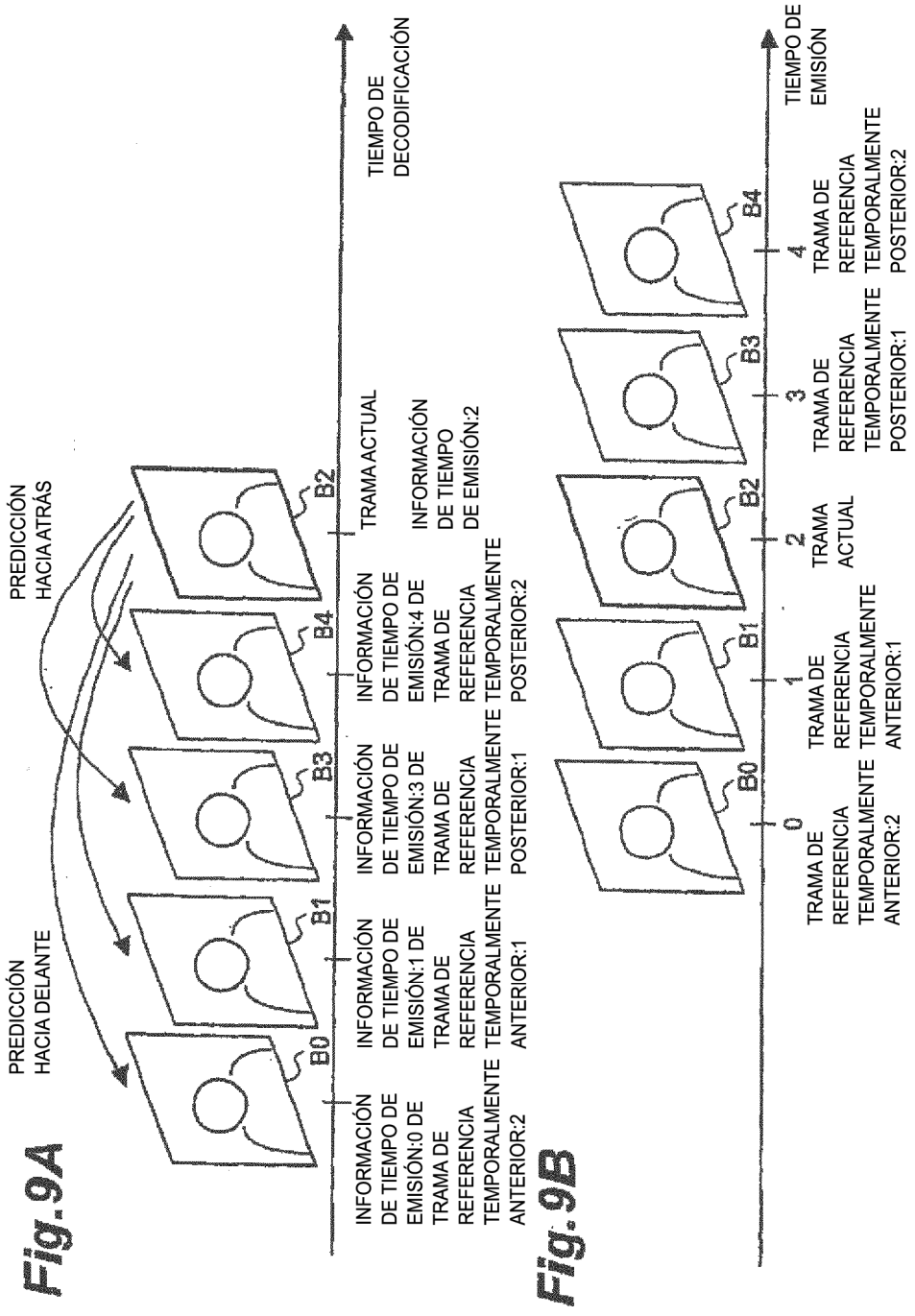
**Fig.6**

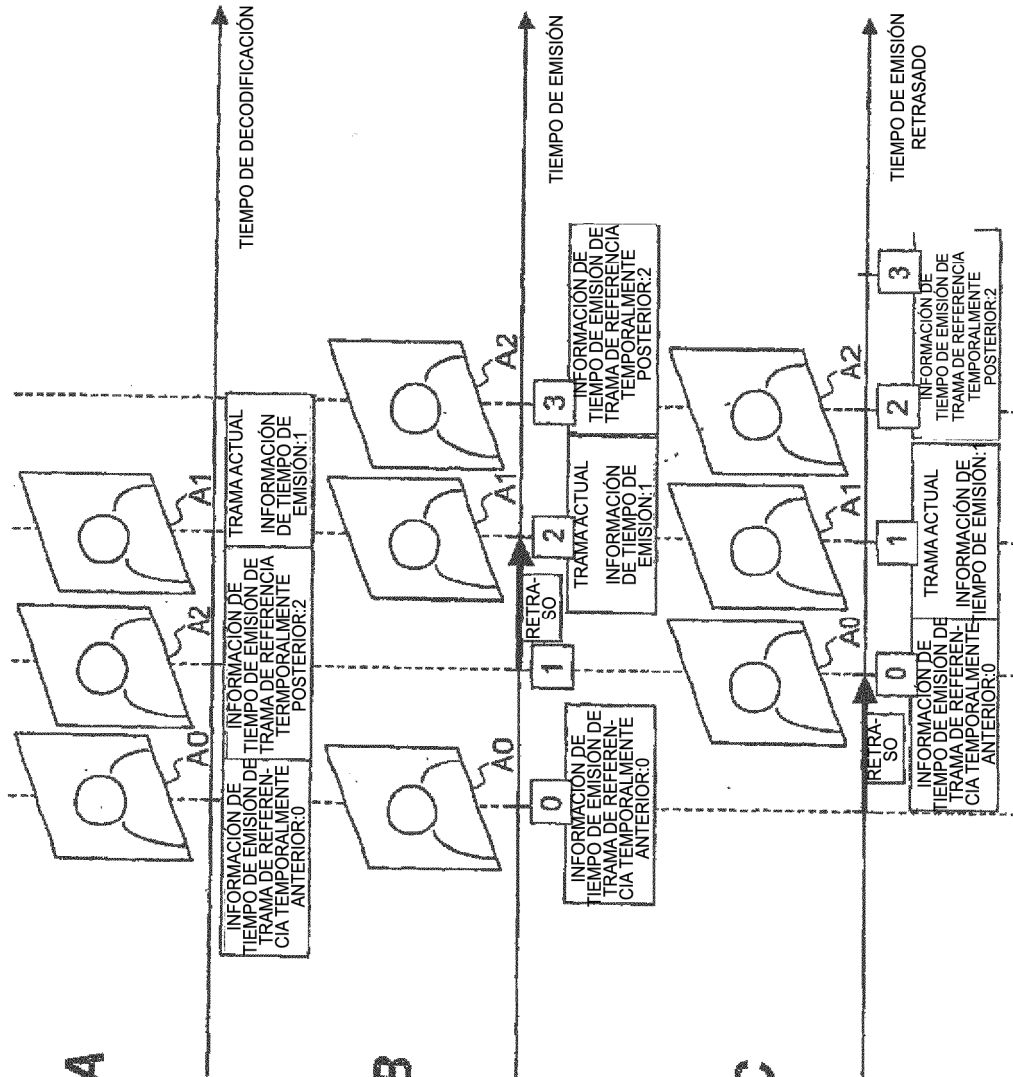




**Fig.8**







**Fig. 10A**

**Fig. 10B**

**Fig. 10C**



