

19



OFICINA ESPAÑOLA DE  
PATENTES Y MARCAS

ESPAÑA



11 Número de publicación: **2 625 334**

51 Int. Cl.:

**H04N 19/503** (2014.01)

12

TRADUCCIÓN DE PATENTE EUROPEA

T3

96 Fecha de presentación y número de la solicitud europea: **26.09.2003** **E 10180810 (3)**

97 Fecha y número de publicación de la concesión europea: **26.04.2017** **EP 2268035**

54 Título: **Método de codificación de vídeo, método 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:

**19.07.2017**

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 2 625 334 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

Método de codificación de vídeo, método de descodificación de vídeo, aparato de codificación de vídeo, aparato de descodificación de vídeo, programa de codificación de vídeo y programa de descodificación de vídeo

5 **Antecedentes de la invención**

### Campo de la invención

10 La presente invención se refiere a un método de descodificación de vídeo y a un aparato de descodificación de vídeo.

### Técnica anterior relacionada

15 Se usan técnicas de codificación de señal de vídeo para transmisión y almacenamiento-regeneración de señales de vídeo. Las técnicas bien conocidas incluyen, por ejemplo, los métodos de codificación de vídeo de normas internacionales tales como la recomendación ITU-T H.263 (a continuación en el presente documento denominada H.263), la norma internacional ISO/IEC 14496-2 (MPEG-4 Visual, a continuación en el presente documento denominada MPEG-4), etc.. Otro método de codificación más nuevo conocido es un método de codificación de vídeo  
20 previsto para cumplir la normalización internacional mediante ITU-T e ISO/IEC; la recomendación ITU-T H.264 y la norma internacional ISO/IEC 14496-10 (Proyecto de Comité Final Común de Especificación de Vídeo Común, a continuación en el presente documento denominada H.26L).

25 El documento US 6205177 divulga un método de descodificación de una secuencia de imágenes codificadas, comprendiendo el método etapas implementadas por ordenador realizadas mediante un procesador de un sistema informático para implementar las etapas de recibir las imágenes codificadas, un tiempo de descodificación de cada una de las imágenes codificadas; descodificar las imágenes codificadas recibidas en imágenes descodificadas, basándose en el tiempo de descodificación recibido; mantener al menos alguna de las imágenes descodificadas en una memoria intermedia de imágenes descodificadas antes de emitir al menos alguna de las imágenes  
30 descodificadas desde la memoria intermedia de imágenes descodificadas en tiempos de emisión respectivos, en el que la memoria intermedia de imágenes descodificadas mantiene al menos una imagen descodificada usada como referencia para descodificar al menos una de las imágenes codificadas; y calcular los tiempos de emisión respectivos de la imagen descodificada.

35 Debido a que una señal de vídeo en movimiento consiste en una serie de imágenes (tramas) que varían poco a poco con el tiempo, es una práctica común en estos métodos de codificación de vídeo implementar predicción inter-tramas entre una trama recuperada como un objetivo para la codificación (trama actual) y otra trama (trama de referencia) y de ese modo se reduce la redundancia temporal en la señal de vídeo. En este caso, cuando la predicción inter-tramas se lleva a cabo entre la trama actual y una trama de referencia menos diferente de la trama  
40 actual, la redundancia puede reducirse más y puede aumentarse la eficacia de codificación.

Por este motivo, tal como se muestra en la figura 6, la trama de referencia para la trama A1 actual puede ser o bien una trama A0 temporalmente anterior o bien una trama A2 temporalmente posterior con respecto a la trama A1 actual. La predicción con la trama anterior se denomina predicción hacia delante, mientras que la predicción con la  
45 trama posterior se denomina predicción hacia atrás. Predicción bidireccional se define como una predicción en la que se selecciona arbitrariamente uno de los dos métodos de la predicción, o como una predicción en la que se usan simultáneamente ambos métodos.

En general, con el uso de tal predicción bidireccional, tal como en el ejemplo mostrado en la figura 6, una trama temporalmente anterior como una trama de referencia para la predicción hacia delante y una trama temporalmente posterior como una trama de referencia para la predicción hacia atrás se almacenan previamente antes de la trama actual.

Las figuras 7A y 7B son diagramas que muestran (A) la descodificación y (B) la emisión de las tramas en el caso de la predicción bidireccional mostrada en la figura 6. Por ejemplo, en la descodificación de MPEG-4, cuando la trama A1 actual se descodifica mediante predicción bidireccional inter-tramas, siendo la trama A0 una trama temporalmente anterior y siendo la trama A2 una trama temporalmente posterior con respecto a la trama A1 actual se descodifican en primer lugar como tramas descodificadas mediante predicción inter-tramas sin uso de la predicción inter-tramas o como tramas descodificadas mediante predicción hacia delante inter-tramas, antes de la descodificación de la trama A1 actual, y se retienen como tramas de referencia. Tras ello, la trama A1 actual se descodifica mediante predicción bidireccional usando estas dos tramas A0, A2 retenidas de ese modo (figura 7A).

En este caso, por tanto, el orden de tiempos de descodificación de la trama temporalmente posterior de referencia A2 y la trama A1 actual es inverso al orden de tiempos de emisión de sus respectivas imágenes descodificadas.  
65 Cada una de estas tramas A0, A1, y A2 está unida con información de tiempo de emisión 0, 1, ó 2, y por tanto puede conocerse la secuencia temporal de las tramas según esta información. Por este motivo, las imágenes

descodificadas se emiten en el orden correcto (figura 7B). En MPEG-4, la información de tiempo de emisión está descrita como valores absolutos.

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

Las figuras 9A y 9B son diagramas que muestran (A) la descodificación y (B) la emisión de las tramas en el caso de la predicción bidireccional mostrada en la figura 8. Por ejemplo, en la descodificación de H.26L, una pluralidad de tramas de referencia puede retenerse dentro de un intervalo hasta un límite superior predeterminado del número de tramas de referencia y, en el caso de llevar a cabo la predicción inter-tramas, una trama de referencia óptima se designa arbitrariamente entre ellas. En este caso, en el que la trama B2 actual se descodifica como una trama predicha de manera bidireccional, las tramas de referencia se descodifican en primer lugar antes de la descodificación de la trama B2 actual; las tramas de referencia incluyen una pluralidad de tramas temporalmente anteriores (por ejemplo, dos tramas B0, B1) y una pluralidad de tramas temporalmente posteriores (por ejemplo, dos tramas B3, B4) con respecto a la trama B2 actual, que están descodificadas y retenidas como tramas de referencia. La trama B2 actual puede predecirse a partir de una trama designada arbitrariamente como la usada para la predicción entre esas tramas B0, B1, B3, y B4 (figura 9A).

En este caso, por tanto, el orden de tiempos de descodificación de las tramas B3, B4 de referencia temporalmente posteriores y la trama B2 actual se vuelve inverso al orden de sus respectivos tiempos de emisión. Cada una de estas tramas B0-B4 está unida con información de tiempo de emisión o información 0-4 de orden de emisión, y la secuencia temporal de las tramas puede conocerse según esta información. Por este motivo, las imágenes descodificadas se emiten en el orden correcto (figura 9B). La información de tiempo de emisión se describe a menudo como valores absolutos. El orden de emisión se usa cuando los intervalos de trama son constantes.

Para llevar a cabo la descodificación mediante la predicción hacia atrás usando tramas temporalmente posteriores como tramas predictivas, es necesario cumplir la condición de que la descodificación de las tramas temporalmente posteriores esté completa antes de la descodificación de la trama actual para que estén disponibles como tramas predictivas. En este caso, se provoca un retraso antes de que la imagen descodificada de la trama actual se vuelva disponible, en comparación con una trama a la que no se aplica la predicción hacia atrás.

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

Se supone en el presente documento que un tiempo en el que se obtiene una imagen descodificada de trama A0 sin retraso debido a la predicción hacia atrás y sin inversión de los órdenes de tiempos de descodificación y tiempos de emisión con respecto a cualquier otra trama (una trama sin retraso y sin inversión se denominará a continuación en el presente documento trama no asociada con predicción hacia atrás) se define como un tiempo de emisión correlacionado con la imagen descodificada, y la imagen descodificada se emite en el tiempo de emisión. Suponiendo que la trama posterior es la trama A1 predicha hacia atrás, la imagen descodificada de la misma se descodificará después de la trama A2 temporalmente posterior, y se da por tanto un retraso antes de que se obtenga la imagen descodificada.

Por este motivo, si el tiempo en el que se obtiene la imagen descodificada para la trama A0 no asociada con predicción hacia atrás se define como una referencia de tiempo de emisión, la imagen descodificada de la trama A1 predicha hacia atrás no se obtiene mediante el tiempo de emisión correlacionado con la misma (figura 10B). Concretamente, un intervalo de tiempo de emisión entre la imagen descodificada de la trama A0 no asociada con predicción hacia atrás y la imagen descodificada de la trama A1 predicha hacia atrás se vuelve más largo debido al tiempo de retraso necesario para la ejecución de la predicción hacia atrás que el intervalo original, lo cual lleva a una emisión de vídeo forzada.

Por tanto, en el caso en el que se aplica la predicción hacia atrás inter-tramas en codificación de vídeo, tal como se muestra en la figura 10C, es necesario retrasar previamente el tiempo de emisión de la imagen descodificada de la

trama A0 no asociada con predicción hacia atrás el tiempo de retraso necesario para la ejecución de la predicción hacia atrás así como para permitir manejar de manera correcta el intervalo de tiempo de emisión para la trama A1 predicha hacia atrás.

5 Habitualmente, la predicción hacia atrás inter-tramas se aplicó a la codificación de vídeo bajo las condiciones en las que se llevaba a cabo la codificación a una alta tasa de bits y la velocidad de trama fija de 30 tramas/segundo al igual que las de señales de radiodifusión de televisión que se usaban siempre, como radiodifusión o acumulación de  
10 televisión de las mismas, debido a que la predicción hacia atrás inter-tramas lleva a más opciones para la predicción y por tanto a un incremento de complejidad computacional, lo cual hace la implementación de la misma difícil en un equipo simple y debido a que no se deseaba el incremento de tiempo de retraso en comunicación en tiempo real que conlleva interlocución bidireccional como conferencias de vídeo.

En este caso, por ejemplo, como en MPEG-4, en el que se usa una trama temporalmente posterior como trama de referencia para la predicción hacia atrás, el tiempo de retraso necesario en ejecución de la predicción hacia atrás es constante. Por ejemplo, cuando la velocidad de trama es de 30 tramas/segundo tal como se describió anteriormente,  
15 el tiempo de retraso es un intervalo de tiempo de cada trama, es decir, de 1/30 segundo. Por consiguiente, el tiempo que debe retrasarse el tiempo de emisión de la imagen descodificada de la trama no asociada con predicción hacia atrás puede establecerse de igual manera a 1/30 segundo.

## 20 Sumario de la invención

En los últimos años, sin embargo, siguiendo la mejora en rendimiento informático y el progreso en diversificación de servicios de vídeo, el retraso es tolerable en envío de video a través de Internet y comunicaciones móviles, y se ha  
25 aumentado el uso de codificación de vídeo que requiere codificación a baja tasa de bits. Para implementar la codificación a baja tasa de bits, se aplican velocidades de trama más pequeñas de 30 tramas/segundo, o se usan velocidades de trama variables para cambiar de manera dinámica la velocidad de trama con el fin de controlar la tasa de bits de codificación.

En tal codificación de vídeo, en la que se aplica la anteriormente mencionada predicción hacia atrás con el fin de  
30 incrementar más la eficacia de codificación, el tiempo de retraso debido a la predicción hacia atrás no es siempre 1/30 segundo tal como se usaba anteriormente. En la aplicación de velocidades de trama variables, las velocidades de trama no son constantes. Por ejemplo, en el caso en el que se use una pequeña velocidad de trama sobre una base temporal, el intervalo de tiempo de cada trama se vuelve grande, y por tanto el tiempo que debe retrasarse el tiempo de emisión de la imagen descodificada de la trama no asociada con predicción hacia atrás no se determina  
35 de manera única. Por este motivo, se vuelve inviable manejar de manera correcta el intervalo de tiempo de emisión entre la imagen descodificada de la trama no asociada con predicción hacia atrás y la imagen descodificada de la trama predicha hacia atrás.

En este caso, existen posibles medios en los que un tiempo de retraso grande permisible está permitido previamente  
40 para la predicción hacia atrás y en los que el tiempo de emisión de la imagen descodificada de la trama no asociada con predicción hacia atrás se retrasa siempre por este tiempo de retraso, manejando de manera correcta de ese modo el intervalo de tiempo de emisión relativo a la imagen descodificada de la trama predicha hacia atrás. En este caso, sin embargo, se añade siempre el retraso grande al tiempo de emisión de la imagen descodificada, independientemente del tiempo de retraso en la predicción hacia atrás práctica.

45 Cuando se usan tramas de referencia múltiples en la predicción hacia atrás como en H.26L, la descodificación de todas las tramas de referencia que son tramas temporalmente posteriores debe completarse antes de la descodificación de la trama actual. Esto aumenta además el tiempo de retraso necesario para la ejecución de la predicción hacia atrás.

50 En este caso, debido a que el número de tramas de referencia usadas en la predicción hacia atrás se determina de manera única como un número de tramas temporalmente posteriores para la trama actual, que se descodificaron antes de la trama actual, el número de tramas de referencia puede cambiarse opcionalmente dentro del intervalo hasta el límite superior predeterminado del número de tramas de referencia máximo.

55 Por ejemplo, suponiendo que el límite superior del número de tramas de referencia es 4, el número de tramas de referencia usado en la predicción hacia atrás puede ser 2 tal como se muestra en la figura 8, o 1 tal como se muestra en la figura 11A, o 3 tal como se muestra en la figura 11B. Debido a que el número de tramas de referencia puede cambiarse de esta manera, el tiempo de retraso necesario para la ejecución de la predicción hacia atrás  
60 puede variar enormemente. Esto lleva a un fallo en el manejo de manera correcta del intervalo de tiempo de emisión entre la imagen descodificada de la trama no asociada con predicción hacia atrás y la imagen descodificada de la trama predicha hacia atrás.

65 En este caso, debido a que el número de tramas de referencia máximo que puede usarse en la predicción hacia atrás no supera el límite superior del número de tramas de referencia, el tiempo de retraso según el límite superior del número de tramas de referencia es un tiempo de retraso máximo que puede realizarse en ejecución de la

predicción hacia atrás. Por tanto, si el tiempo de emisión de la imagen descodificada de la trama no asociada con predicción hacia atrás siempre se retrasa este tiempo de retraso, puede manejarse el intervalo de tiempo de emisión relativo a la imagen descodificada de la trama predicha hacia atrás de manera correcta.

5 En este caso, sin embargo, se añade un retraso grande siempre al tiempo de emisión de la imagen descodificada, independientemente del número de tramas de referencia usadas realmente para la trama predicha hacia atrás. En la aplicación de velocidades de trama variables tal como se describió anteriormente, mientras que el número de tramas de referencia máximo puede determinarse de manera única, el tiempo de retraso máximo no puede determinarse de manera única.

10 Hasta ahora, en la aplicación de la predicción hacia atrás para la codificación de vídeo, era inviable determinar de manera única el tiempo de retraso necesario para la ejecución de la predicción hacia atrás, excepto el caso en el que usar una velocidad de trama fija estaba claro. Esto daba como resultado un fallo en el manejo de manera correcta del intervalo de tiempo de emisión entre la imagen descodificada de la trama no asociada con predicción hacia atrás y la imagen descodificada de la trama predicha hacia atrás, planteando por tanto el problema de que la emisión de vídeo se volviera forzada.

15 En el caso en el que se usan tramas de referencia múltiples en la predicción hacia atrás, también puede cambiarse el número de tramas de referencia, lo cual posiblemente varía el tiempo de retraso. Por tanto, existe el problema del fallo en el manejo de manera correcta del intervalo de tiempo entre la imagen descodificada de la trama no asociada con predicción hacia atrás y la imagen descodificada de la trama predicha hacia atrás. En el caso en el que el tiempo de retraso máximo se suponga siempre con el propósito de afrontar este problema, esto hace surgir el problema de que se añade siempre el retraso grande al tiempo de emisión de la imagen descodificada.

20 La presente invención se ha realizado con el fin de resolver los problemas anteriores, y un objeto de la invención es proporcionar un método de descodificación de vídeo y un aparato de descodificación de vídeo que sea capaz de lograr la emisión de imágenes descodificadas en intervalos de tiempo apropiados cuando se emplea predicción hacia atrás inter-tramas.

30 El objeto de la invención se logra mediante las reivindicaciones 1 y 2.

#### **Breve descripción de los dibujos**

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

La figura 2 es un diagrama que muestra un ejemplo de codificación de tramas en caso de llevarse a cabo predicción bidireccional.

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

La figura 4 es un diagrama de bloques que muestra un ejemplo de la configuración del aparato de descodificación de vídeo.

45 Las figuras 5A y 5B son diagramas que muestran (A) la descodificación y (B) la emisión de tramas en caso de llevarse a cabo la predicción bidireccional mostrada en la figura 2.

50 La figura 6 es un diagrama que muestra codificación de tramas en caso de llevarse a cabo predicción bidireccional.

Las figuras 7A y 7B son diagramas que muestran (A) la descodificación y (B) la emisión de tramas en caso de llevarse a cabo la predicción bidireccional mostrada en la figura 6.

55 La figura 8 es un diagrama que muestra codificación de tramas en caso de llevarse a cabo predicción bidireccional.

Las figuras 9A y 9B son diagramas que muestran (A) la descodificación y (B) la emisión de tramas en caso de llevarse a cabo la predicción bidireccional mostrada en la figura 8.

60 Las figuras 10A a 10C son diagramas que muestran (A) la descodificación, (B) la emisión, y (C) la emisión retrasada de tramas en caso de llevarse a cabo predicción bidireccional.

Las figuras 11A y 11B son diagramas que muestran codificación de tramas en caso de llevarse a cabo predicción bidireccional.

65 **Descripción de las realizaciones preferidas**

Las realizaciones preferidas del método de codificación de vídeo y aparato de descodificación de vídeo según la presente invención se describirán a continuación en detalle con referencia a los dibujos. Los mismos elementos se designarán con los mismos símbolos de referencia a lo largo de toda la descripción de los dibujos, sin descripción redundante de los mismos.

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

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

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

30 En correspondencia con tal aparato 1 de codificación de vídeo, el aparato 2 de descodificación de vídeo está configurado para efectuar la entrada del tiempo de retraso máximo que está provocado por la predicción hacia atrás, además de los datos D1 codificados del aparato 1 de codificación de vídeo. A continuación, el aparato 2 de descodificación de vídeo descodifica los datos D1 codificados con referencia al tiempo de retraso máximo de entrada para generar los datos D2 de vídeo.

35 Mediante el aparato 1 de codificación de vídeo y el método de codificación de vídeo configurado para emitir el tiempo de retraso máximo, el aparato 2 de descodificación de vídeo y el método de descodificación de vídeo configurados para efectuar la entrada del tiempo de retraso máximo, y el sistema de procesamiento de vídeo equipado con estos aparatos 1, 2, que están adaptados para la predicción hacia atrás inter-tramas tal como se describió anteriormente, se vuelve viable lograr la emisión de imágenes descodificadas en intervalos de tiempo apropiados en ejecución de la predicción inter-tramas usando la predicción hacia atrás inter-tramas.

45 Sobre el tiempo de retraso máximo emitido en la codificación de vídeo, por ejemplo, puede definirse el tiempo de retraso máximo como una diferencia de tiempo entre un tiempo de aparición de una trama que va a someterse a la predicción hacia atrás inter-tramas y un tiempo de aparición de una trama posterior temporalmente última que puede usarse como una trama de referencia para la predicción hacia atrás.

50 En cuanto al tiempo de retraso máximo introducido en la descodificación de vídeo, por ejemplo, el tiempo de retraso máximo (a continuación en el presente documento denominado `dpb_output_delay`) puede definirse como una diferencia de tiempo entre un tiempo de descodificación de una trama sin retraso debido a la predicción hacia atrás inter-tramas y sin inversión de órdenes de tiempos de descodificación y tiempos de emisión con respecto a otra trama (el tiempo de descodificación se denominará a continuación en el presente documento  $T_r$ ) y un tiempo de emisión de imagen descodificada correlacionada con la trama pertinente (el tiempo de emisión se denominará a continuación en el presente documento  $T_o$ ). En este caso se establece, preferiblemente, una referencia para los  
55 tiempos de emisión de imagen descodificada tras a partir de ahí basándose en el tiempo de retraso máximo.

La aplicación del tiempo de retraso máximo puede ser un método de aplicación a la totalidad de datos codificados o un método de aplicación a cada trama. Otro método de aplicación es un método de aplicación del tiempo de retraso máximo a cada una de las tramas posteriores a la comunicación de la información del tiempo de retraso máximo, es decir, a la trama para la que se indica el tiempo de retraso máximo y a cada una de las tramas temporalmente posteriores a esta trama. La emisión, introducción, aplicación, etc. del tiempo de retraso máximo en estos métodos se detallará específicamente más adelante.

60 El procesamiento correspondiente al método de codificación de vídeo ejecutado en el aparato 1 de codificación de vídeo anterior puede confirmarse mediante el programa de codificación de vídeo para autorizar a un ordenador a ejecutar la codificación de vídeo. El procesamiento correspondiente al método de descodificación de vídeo ejecutado

en el aparato 2 de decodificación de vídeo puede confirmarse mediante el programa de decodificación de vídeo para autorizar a un ordenador a ejecutar la decodificación de vídeo.

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

10 De manera similar, el aparato 2 de decodificación de vídeo 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 2 de decodificación de vídeo puede confirmarse autorizando a la CPU a ejecutar el programa de decodificación de vídeo predeterminado.

15 El programa anteriormente señalado puede distribuirse para autorizar a la CPU a ejecutar los procedimientos para la codificación de vídeo o para la decodificación de vídeo de forma que se graba en medios de grabación legibles por ordenador. Tales medios de grabación incluyen, por ejemplo, medios magnéticos tales como discos duros y discos flexibles, medios ópticos tal como CD-ROM y DVD-ROM, medios magneto-ópticos 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 programa.

20 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 figura 1, y el método de codificación de vídeo y el método de decodificación de vídeo correspondientes al mismo se describirán con realizaciones específicas. La descripción siguiente en el presente documento se basará en la suposición de que las operaciones de codificación y decodificación de vídeo en movimiento están implementadas basándose en H.26L, y partes no descritas específicamente sobre la operación en codificación de vídeo serán conforme a la operación en H.26L. Sin embargo, ha de observarse que la presente invención no se limita a H.26L.

30 (Primera realización)

35 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 trama fija. En la codificación según la presente realización, el número de tramas de referencia máximo usado para la predicción hacia atrás se determina en primer lugar, tras ello se calcula el tiempo de retraso máximo a partir de este número de tramas de referencia máximo y la velocidad de trama usada en codificación, y entonces se emite el tiempo de retraso máximo. En la decodificación según la presente realización, en el caso de decodificar una trama no asociada con predicción hacia atrás, un tiempo de emisión de una imagen descodificada de la misma se retrasa el tiempo de retraso máximo de entrada. Tras ello, se aplica el tiempo de retraso al tiempo de emisión de manera uniforme a cada trama, para impedir que el intervalo de tiempo de emisión entre la imagen descodificada de la trama no asociada con predicción hacia atrás y la imagen descodificada de la trama predicha hacia atrás se desvíe del intervalo original.

45 En la codificación, debido a que el límite superior del número de tramas de referencia usado se determina previamente, el número de tramas de referencia máximo usado para la predicción hacia atrás se determina en primer lugar dentro del intervalo que no supera el límite superior. A continuación, basándose en la velocidad de trama usada para la codificación, que también se determina previamente, se calcula el tiempo de retraso máximo como un intervalo de tiempo de una trama o dos o más tramas según el número de tramas de referencia máximo usado para la predicción hacia atrás.

50 La figura 2 es un diagrama que muestra un ejemplo de codificación de una trama en ejecución de la predicción bidireccional. En el presente documento, la figura 2 muestra el ejemplo en el que las tramas de referencia usadas para la trama F2 actual son dos tramas F0, F1 temporalmente anteriores antes de la trama F2 actual y dos tramas F3, F4 temporalmente posteriores después de la trama F2 actual.

55 En el caso de que el número de tramas de referencia máximo usado para la predicción hacia atrás sea de 2 y de que la velocidad de trama sea de 15 tramas/segundo, tal como se muestra en la figura 2, el intervalo de tiempo de una trama es de  $1/15$  segundos. En este caso, por tanto, el tiempo de retraso máximo es  $2 \times (1/15) = 2/15$  segundos.

60 En la operación de codificación, se controla la codificación de cada trama a continuación en el presente documento para no llevar a cabo la predicción hacia atrás que requiere un tiempo de retraso sobre el tiempo de retraso máximo. Específicamente, una secuencia de codificación de tramas se controla de manera que cualquier trama de referencia usada en la predicción hacia atrás, es decir, cualquier trama temporalmente posterior después de la trama actual, no se codifique y emita antes de la trama actual por encima del número de tramas de referencia máximo usado en la predicción hacia atrás.

65

La figura 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 1 de codificación de vídeo mostrado en la figura 3 se compone de un codificador 10 para codificar una trama (imagen) mediante el método predeterminado, un controlador 15 (CPU) para controlar operaciones de partes respectivas en el aparato 1 de codificación, una memoria 11 de trama dispuesta entre el terminal 1a de entrada y el codificador 10, y un multiplexor 12 dispuesto entre el terminal 1b de salida y el codificador 10. El controlador 15 tiene un dispositivo 16 de cálculo de tiempo de retraso máximo para calcular el tiempo de retraso máximo, como una función del mismo. El codificador 10 está dotado de una memoria 13 intermedia de salida.

En la codificación de vídeo en el presente aparato 1 de codificación, se introducen condiciones para la codificación de vídeo a través del terminal 1c de entrada. En esta introducción de las condiciones, las condiciones de codificación se seleccionan o 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 una trama como objetivo para la codificación, la velocidad de trama, y la tasa de bits y, además de estas, las condiciones de codificación incluyen también una estructura de referencia predictiva del vídeo (si se aplica predicción hacia atrás), el número de tramas almacenadas temporalmente y usadas como tramas de referencia (correspondientes a la capacidad de memoria 13 intermedia de salida), y el número de tramas de referencia usado en la predicción hacia atrás. Estas condiciones pueden establecerse para variar con el tiempo. Las condiciones de codificación introducidas a través del terminal 1c de entrada se almacenan en el controlador 15.

Con un inicio de la operación de codificación, el controlador 15 envía las condiciones de codificación al codificador 10, en el que se establecen las condiciones de codificación. Por otra parte, una trama tal como un objeto codificado se introduce a través del terminal 1a de entrada y se alimenta a través de la memoria 11 de trama al codificador 10 para codificarse en el mismo. La trama de entrada se guarda temporalmente en la memoria 11 de trama, debido a que se cambia el orden de tramas para la ejecución de la predicción hacia atrás. Por ejemplo, en el ejemplo mostrado en la figura 2, se introduce la trama F2 a través del terminal 1a de entrada antes de las tramas F3, F4, pero se codifica después de las tramas F3, F4; por tanto, la trama F2 se guarda temporalmente en la memoria 11 de trama.

El codificador 10 codifica la trama basándose en el algoritmo de H.26L. A continuación los datos codificados se alimentan al multiplexor 12 para multiplexarse con otra información relacionada y a continuación los datos multiplexados se emiten a través del terminal 1b de salida. La trama usada para la predicción se reproduce en el codificador 10 y se almacena como una trama de referencia para la codificación de la siguiente trama en la memoria 13 intermedia.

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

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

Se supone en la presente realización que se añade una sintaxis para transmitir el tiempo de retraso máximo a la sintaxis de datos codificados en H.26L, con el propósito de implementar la emisión del tiempo de retraso máximo en la codificación y la entrada del tiempo de retraso máximo en la decodificación. En este ejemplo la nueva sintaxis está añadida en el conjunto de parámetro de secuencia que es una sintaxis para transmitir la información que va a aplicarse a la totalidad de datos codificados.

El parámetro `dpb_output_delay` se define como una sintaxis para llevar el tiempo de retraso máximo. Se supone en el presente documento 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 retraso máximo en la unidad de tiempo de 90 kHz. Una cifra indicada en la unidad de tiempo se codifica y transmite mediante un código de longitud fija sin asignar de 32 bits. Por ejemplo, cuando el tiempo de retraso máximo es de 2/15 segundos tal como se describió anteriormente, `dpb_output_delay` es  $(2/15) \times 90000 = 12000$ .

En la operación de decodificación, se descodifica el tiempo de retraso máximo llevado por `dpb_output_delay`, y se retrasa un tiempo de emisión de una imagen descodificada usándolo.

La figura 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 2 de decodificación de vídeo mostrado en la figura 4 se compone de un descodificador 20 para descodificar datos codificados para generar una imagen regenerada, un controlador 25 (CPU) para controlar operaciones de partes respectivas en el aparato 2 de decodificación, una memoria 21 intermedia de entrada dispuesta entre el terminal 2a de entrada y el descodificador 20, y una memoria 22 intermedia



de salida dispuesta entre el terminal 2b de salida y el descodificador 20. El controlador 25 tiene un dispositivo 26 de cálculo de tiempo de emisión de imagen para calcular un tiempo de emisión para la visualización de una trama, como una función de la misma.

5 En la descodificación de vídeo en el presente aparato 2 de descodificación, se introducen datos como objeto descodificado a través del terminal 2a de entrada. Estos datos son datos multiplexados de los datos codificados de cada trama codificada por el aparato 1 de codificación mostrado en la figura 3, indicando el tiempo de retraso máximo `dpb_output_delay`, y el identificador (N) el momento de visualización de cada trama.

10 Los datos de entrada se almacenan en la memoria 21 intermedia de entrada. Cuando un comando desde el controlador 25 indica la llegada de un tiempo de descodificación, se introducen datos de una trama desde la memoria 21 intermedia de entrada en el descodificador 20 y se descodifican a continuación según el algoritmo de H.26L. La trama regenerada se almacena de esta manera en la memoria 22 intermedia de salida. La trama en la memoria 22 intermedia de salida se alimenta de nuevo a través de la línea 23 al descodificar 20 para usarse como  
15 una trama de referencia para la descodificación de la siguiente trama.

Por otra parte, el tiempo de retraso máximo `dpb_output_delay`, la velocidad de trama, y el identificador (N) de cada trama descodificada en el descodificador 20 se alimentan en el controlador 25. A continuación el dispositivo 26 de  
20 cálculo de tiempo de emisión de imagen del controlador 25 calcula el tiempo de emisión de cada trama a partir de estos datos según la siguiente ecuación.

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

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

Suponiendo que `dpb_output_delay` es de 2/15 segundos y el intervalo de trama es de 1/15 segundos como en el ejemplo mostrado en la figura 2, los tiempos de emisión de las respectivas tramas se calculan como sigue según la ecuación anterior.

30  $N = 0, T_o(0) = 2/15$

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

35  $N = 2, T_o(2) = 4/15$

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

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

45 Las figuras 5A y 5B son diagramas que muestran (A) la descodificación y (B) la emisión de las tramas en caso de la predicción bidireccional mostrada en la figura 2. Se supone en la operación de descodificación que los datos codificados de las tramas se descodifican en el orden necesario para la ejecución de la predicción inter-tramas, los intervalos de la misma son intervalos de tiempo constantes según la velocidad de trama, y el tiempo necesario para la operación de descodificación es despreciable para cada trama, independientemente de si se aplica predicción inter-tramas e independientemente de los sentidos de la predicción inter-tramas. En este caso, el tiempo de retraso máximo necesario para la ejecución de la predicción hacia atrás en la trama predicha hacia atrás es igual a un  
50 intervalo de tiempo de una trama o tramas según el número de tramas de referencia máximo usado para la predicción hacia atrás. Este tiempo se lleva como un tiempo de retraso máximo mediante `dpb_output_delay`. Por consiguiente, para emitir una imagen descodificada, se retrasa un tiempo de emisión de la misma el tiempo de retraso máximo.

55 En la práctica, los intervalos de descodificación de las respectivas tramas no son constantes, y pueden variar dependiendo de tales factores como la variación en bits de codificación de las tramas. El tiempo necesario para la operación de descodificación de cada trama puede variar también según si la trama es una trama predicha hacia atrás o según los bits de codificación de cada trama.

60 Para retrasar el tiempo de emisión, por tanto, la referencia se establece en el tiempo en el que se obtiene la imagen descodificada para la trama F0 no asociada con predicción hacia atrás sin retraso debido a la predicción hacia atrás y sin inversión de órdenes de tiempos de descodificación y tiempos de emisión con respecto a cualquier otra trama, tal como se muestra en las figuras 5A y 5B. Concretamente, un tiempo obtenido retrasando el tiempo en el que se obtiene la imagen descodificada, por el tiempo de retraso máximo comunicado mediante `dpb_output_delay` se define

como un tiempo igual al tiempo de emisión correlacionado con esta imagen descodificada, y se usa como un tiempo en emisión de referencia de imágenes descodificadas. Tras ello, las imágenes F1-F4 descodificadas se emiten cuando este tiempo de referencia coincide con un tiempo igual a un tiempo de emisión correlacionado con cada una imagen descodificada.

Por ejemplo, cuando el tiempo de retraso máximo es de 2/15 segundos tal como se describió anteriormente, un tiempo en un retraso de 2/15 segundos desde el tiempo cuando se obtiene la imagen descodificada para la trama no asociada con predicción hacia atrás, se define como un tiempo igual al tiempo de emisión correlacionado con esta imagen descodificada y se usa como un tiempo en emisión de referencia de imágenes descodificadas después.

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

En caso de que se omita la comunicación del tiempo de retraso máximo, la operación de codificación puede estipularse previamente, por ejemplo, para no usar la predicción hacia atrás o de manera que el número de tramas de referencia usado en la predicción hacia atrás pueda alterarse opcionalmente dentro de un intervalo que no supere el límite superior del número de tramas de referencia.

La operación de descodificación puede configurarse para realizarse en conformidad con la estipulación en la operación de codificación, por ejemplo, cuando no se aplica la predicción hacia atrás, se produce retraso innecesario para la ejecución de la predicción hacia atrás; o, también puede configurarse la operación de descodificación de manera que el número de tramas de referencia usado en la predicción hacia atrás pueda alterarse opcionalmente dentro del intervalo que no supere el límite superior del número de tramas de referencia, es decir, el tiempo de retraso puede variar enormemente. En este caso, la operación de descodificación puede configurarse para realizar siempre un procesamiento que asuma un tiempo de retraso máximo esperado, o la operación de descodificación puede configurarse para permitir una variación de intervalos de tiempo de emisión de imágenes descodificadas y realizar procesamiento simplificado sin consideración del tiempo de retraso de cada trama.

La presente realización se describió suponiendo que las operaciones se implementaron basándose en H.26L, pero se observa que los métodos de codificación de vídeo a los que puede aplicarse la presente invención no se limitan a H.26L y que la presente invención puede aplicarse a diversos métodos de codificación de vídeo usando la predicción hacia atrás inter-tramas.

En la presente realización, se añadió la sintaxis por códigos de longitud fija como una sintaxis para transmitir el tiempo de retraso máximo en el conjunto de parámetros de secuencia, pero se observa que los códigos y la sintaxis para transmitirlo, o la unidad de tiempo para expresar el tiempo de retraso máximo, por supuesto, no se limitan a estos. Los códigos de longitud fija pueden sustituirse por códigos de longitud variable, y el tiempo de retraso máximo puede transmitirse mediante cualquiera de diversas sintaxis que pueden transportar información que va a aplicarse a la totalidad de datos codificados.

Por ejemplo, en H.26L, puede añadirse una sintaxis para dar un mensaje de información de mejora complementario. En un caso que usa otro método de codificación de vídeo, puede transmitirse el tiempo de retraso máximo mediante una sintaxis para transmitir la información que va a aplicarse a la totalidad de datos codificados en el método de codificación pertinente. En otro caso, también puede transmitirse el tiempo de retraso máximo fuera de los datos codificados en el método de codificación de vídeo como en la Recomendación ITU-T H.245 usada para transportar información de control en comunicación usando H.263.

(Segunda realización)

Se describirá a continuación la segunda realización de la presente invención. La presente realización describirá una forma realizada de codificación a velocidades de trama variables. Las operaciones en la codificación y descodificación según la presente realización son básicamente muy similares a las de la primera realización. Debido a que la presente realización usa velocidades de trama variables, implica una operación a velocidades de trama bajas para evitar la ejecución de la predicción hacia atrás que requiere el tiempo de retraso por encima del tiempo de retraso máximo calculado previamente, además de la operación en codificación en la primera realización, para impedir que el intervalo de tiempo de emisión entre la imagen descodificada de la trama no asociada con predicción hacia atrás y la imagen descodificada de la trama predicha hacia atrás se desvíe del intervalo original incluso con variación de velocidades de trama.

Debido a que en la operación de codificación el límite superior del número de tramas de referencia se determina previamente, el número de tramas de referencia máximo usado para la predicción hacia atrás se determina en primer lugar dentro del intervalo que no supera el límite superior. A continuación el intervalo de tiempo de trama máximo se determina basándose en una velocidad de trama objetivo previamente determinada en control de tasas de bits de codificación, y se calcula el tiempo de retraso máximo como un intervalo de tiempo de una trama o dos o

más tramas según el número de tramas de referencia máximo usado en la predicción hacia atrás y el intervalo de tiempo de trama máximo.

5 En la operación de codificación, tras ello, se controla la codificación de cada trama para evitar la predicción hacia atrás que requiere un tiempo de retraso más allá del tiempo de retraso máximo. Específicamente, el orden de codificación de tramas se controla para impedir que cualquier trama de referencia usada en la predicción hacia atrás, es decir, cualquier trama temporalmente posterior después de la trama actual, que va más allá del número de tramas de referencia máximo usado en la predicción hacia atrás, se codifique y emita antes de la trama actual.

10 Además, cuando la velocidad de trama de codificación se vuelve temporalmente pequeña debido al control de tasas de bits de codificación, para hacer que el intervalo de trama de tiempo en este caso sea más grande que el intervalo de tiempo de trama máximo, se controla la codificación de cada trama para no aplicar predicción hacia atrás a la codificación de la trama.

15 La presente realización es sustancialmente idéntica a la primera realización porque el tiempo de retraso máximo se emite en la codificación, porque la sintaxis `dpb_output_delay` para transmitir el tiempo de retraso máximo se añade a la sintaxis de datos codificados con el propósito de efectuar la entrada de la misma en la decodificación, y por la definición de la sintaxis.

20 En la presente realización, la operación de decodificación se dispone para decodificar el tiempo de retraso máximo comunicado mediante `dpb_output_delay` y el retraso el tiempo de emisión de la imagen decodificada mediante uso de este. Este procesamiento también es el mismo que el de la primera realización.

(Tercera realización)

25 Se describirá a continuación la tercera realización de la presente invención. La presente realización describirá una forma realizada en la que el tiempo de retraso máximo se comunica opcionalmente para cada trama y por tanto puede cambiarse de manera flexible. Las operaciones en la codificación y decodificación según la presente realización son básicamente similares a aquellas en la primera realización o la segunda realización.

30 En la presente realización, la sintaxis `dpb_output_delay` para transmitir el tiempo de retraso máximo, que se definió en la primera realización, se dispone para añadirse en el conjunto de parámetro de imagen que es una sintaxis para llevar la información aplicada a cada trama en lugar de la sintaxis para llevar la información aplicada a la totalidad de datos codificados. La sintaxis `dpb_output_delay` en el presente documento está configurada para indicar el tiempo de retraso 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 transmite mediante un código de longitud fija sin asignar de 32 bits.

35 La presente realización es muy similar a la primera realización, en cuanto al cálculo del tiempo de retraso máximo en codificación y en cuanto al retraso del tiempo de emisión de la imagen decodificada por el uso del tiempo de retraso máximo en la decodificación. Las configuraciones del aparato de codificación de vídeo y el aparato de decodificación de vídeo usado en la presente realización son muy similares a aquellas mostradas en las figuras 3 y 4 sobre la primera realización.

40 Va a explicarse cómo determinar el tiempo de retraso máximo `dpb_output_delay` de cada trama en la presente realización. En el aparato 1 de codificación mostrado en la figura 3, el controlador 15 calcula el tiempo (D) de retraso debido a la predicción hacia atrás mediante el método tal como se describió en la primera realización y determina el tiempo  $Tr(n)$  de codificación de cada trama. Cuando un tiempo  $Tin(n)$  de visualización de cada trama se alimenta desde la memoria 11 de trama, se calcula `dpb_output_delay(n)` de esta trama tal como sigue.

50 
$$dpb\_output\_delay(n) = Tin(n) + D - Tr(n)$$

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

55 En la presente realización, el tiempo  $Tr(n)$  para la codificación de cada trama se codifica también a la vez. Tomando la figura 2 como ejemplo,  $D = 2/15$  segundos, y  $Tin(n) = 0, 1/15, 2/15, 3/15, \text{ ó } 4/15$  ( $n = 0, 1, 2, 3, \text{ ó } 4$ ). Debido al cambio en el orden de codificación,  $Tr(n)$  se vuelve tal como sigue:  $Tr(n) = 0, 1/15, 4/15, 2/15, \text{ ó } 3/15$  ( $n = 0, 1, 2, 3, \text{ ó } 4$ ). En el presente documento se obtiene `dpb_output_delay(n)` de cada trama tal como sigue.

60 
$$n = 0, dpb\_output\_delay(0)$$

$$= 0 + 2/15 - 0 = 2/15$$

$$n = 1, dpb\_output\_delay(1)$$

65 
$$= 1/15 + 2/15 - 1/15 = 2/15$$

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

5

$$n = 3, \text{dpb\_output\_delay}(3) \\ = 3/13 + 2/15 - 2/15 = 3/15$$

10

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

15

Por otra parte, en el aparato 2 de descodificación mostrado en la figura 4, el descodificador 20 envía  $\text{dpb\_output\_delay}(n)$  y  $\text{Tr}(n)$  de cada trama al controlador 25 y el controlador 25 calcula el tiempo  $\text{To}(n)$  de emisión de cada trama basándose en la siguiente ecuación.

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

20

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

25

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

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

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

30

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

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

35

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

40

Debido a que el tiempo de retraso máximo define el tiempo en emisión de referencia de imágenes descodificadas a partir del tiempo cuando se adquiere la imagen descodificada de la trama no asociada con predicción hacia atrás, es suficiente transmitir el tiempo de retraso máximo solamente para la trama no asociada con predicción hacia atrás. Por tanto, se puede emplear, por ejemplo, una configuración en la que la sintaxis para transmitir el tiempo de retraso máximo está dispuesta para poder omitirse en la suposición de que una marca que indica la presencia o ausencia de la sintaxis se transmite antes de la misma. La sintaxis puede estar dispuesta para omitirse opcionalmente para la trama no asociada con predicción hacia atrás, siempre que el tiempo de retraso máximo transmitido antes se aplique en este caso cuando no se transmita el tiempo de retraso máximo.

50

La sintaxis para cada trama en la presente realización también puede usarse simultáneamente con la sintaxis para la totalidad de datos codificados tal como se definió en la primera realización. En este caso, la sintaxis para cada trama 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 retraso máximo transmitido en la sintaxis para la totalidad de datos codificados se aplica de manera continua antes de que el tiempo de retraso máximo se transmita en la sintaxis para cada trama. Después de que se actualice mediante la sintaxis para cada trama, el tiempo retrasado basado en la misma se usa como tiempo de referencia en emisión de cada imagen descodificada tras la misma.

55

La presente realización se describió suponiendo que estaba basada de manera confirmada en H.26L, pero se observa que los métodos de codificación de vídeo a los que puede aplicarse la presente invención no se limitan a H.26L y que la presente invención puede aplicarse a diversos métodos de codificación de vídeo usando la predicción hacia atrás inter-tramas.

60

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

65

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

10 El método de codificación de vídeo, método de descodificación de vídeo, aparato de codificación de vídeo, aparato de descodificación de vídeo, sistema de procesamiento de vídeo, programa de codificación de vídeo y programa de descodificación de vídeo según la presente invención proporcionan el siguiente efecto, tal como se detalla a continuación. Concretamente, cuando se codifica una imagen en movimiento que consiste en una serie de tramas mediante la predicción hacia atrás inter-tramas para emitirse, se vuelve viable lograr la emisión de imágenes descodificadas en intervalos de tiempo apropiados al emplear la predicción hacia atrás inter-tramas, mediante el método de codificación de vídeo, aparato de codificación y programa de codificación configurados para emitir el tiempo de retraso máximo debido a la predicción hacia atrás, el método de descodificación de vídeo, aparato de descodificación y programa de descodificación configurados para efectuar la entrada del tiempo de retraso máximo, y el sistema de procesamiento de vídeo que usa los mismos.

20 Particularmente, a diferencia de la técnica anterior, los tiempos de emisión no son valores absolutos, sino valores relativos del tiempo  $T_r$  de descodificación; por tanto, la invención proporciona el efecto de capacidad de describir y transmitir de manera precisa el valor del tiempo de retraso máximo `dpb_output_delay` mediante un pequeño número de bits, incluso en el caso de que la velocidad de trama sea variable. Incluso si el tiempo  $T_r$  de descodificación tiene un desplazamiento o si no se recibe, se emitirá una imagen correspondiente con un retraso de `dpb_output_delay` del tiempo de terminación de descodificación, presentando por tanto la ventaja de que se emitan imágenes a intervalos correctos.

25

**REIVINDICACIONES**

1. Un método de descodificación de una secuencia de imágenes codificadas, comprendiendo el método etapas implementadas por ordenador realizadas mediante un procesador de un sistema informático para implementar las siguientes etapas:
- 5
- recibir las imágenes codificadas, un tiempo de descodificación de cada una de las imágenes codificadas, y sintaxis de tiempo de retraso asociadas respectivamente con las imágenes codificadas, en el que las sintaxis de tiempo de retraso representan al menos un tiempo de retraso máximo que está provocado por predicción hacia atrás definido de manera única para las imágenes codificadas individuales;
- 10
- descodificar las imágenes codificadas recibidas para dar imágenes descodificadas, basándose en el tiempo de descodificación recibido;
- 15
- mantener al menos alguna de las imágenes descodificadas en una memoria intermedia de imágenes descodificadas antes de emitir al menos alguna de las imágenes descodificadas desde la memoria intermedia de imágenes descodificadas en respectivos tiempos de emisión, en el que la memoria intermedia de imágenes descodificadas mantiene al menos una imagen descodificada usada como referencia para descodificar al menos una de las imágenes codificadas; y
- 20
- calcular los respectivos tiempos de emisión de la imagen descodificada, basándose en el respectivo tiempo de descodificación y los respectivos tiempos de retraso máximos.
2. Un aparato de descodificación para la descodificación de una secuencia de imágenes codificadas, que comprende:
- 25
- un receptor (21) configurado para recibir las imágenes codificadas, un tiempo de descodificación de cada una de las imágenes codificadas y sintaxis de tiempo de retraso asociadas respectivamente con las imágenes codificadas, en el que las sintaxis de tiempo de retraso representan al menos un tiempo de retraso máximo que está provocado por predicción hacia atrás definido de manera única para las imágenes codificadas individuales;
- 30
- un descodificador (20) configurado para descodificar las imágenes codificadas recibidas para dar imágenes descodificadas, basándose en el tiempo de descodificación recibido;
- 35
- una memoria (22) intermedia de imágenes descodificadas configurada para poder mantener las imágenes descodificadas; y
- 40
- un control (25) de salida configurado para mantener al menos alguna de las imágenes descodificadas en la memoria (22) intermedia de imágenes descodificadas antes de emitir al menos alguna de las imágenes descodificadas desde la memoria intermedia de imágenes descodificadas en respectivos tiempos de emisión, en el que la memoria (22) intermedia de imágenes descodificadas mantiene al menos una imagen descodificada usada como referencia para descodificar al menos una de las imágenes codificadas,
- 45
- en el que el control (25) de salida está configurado además para calcular los respectivos tiempos de emisión de al menos algunas imágenes descodificadas, basándose en los respectivos tiempos de descodificación y los respectivos tiempos de retraso máximos.

**Fig.1**

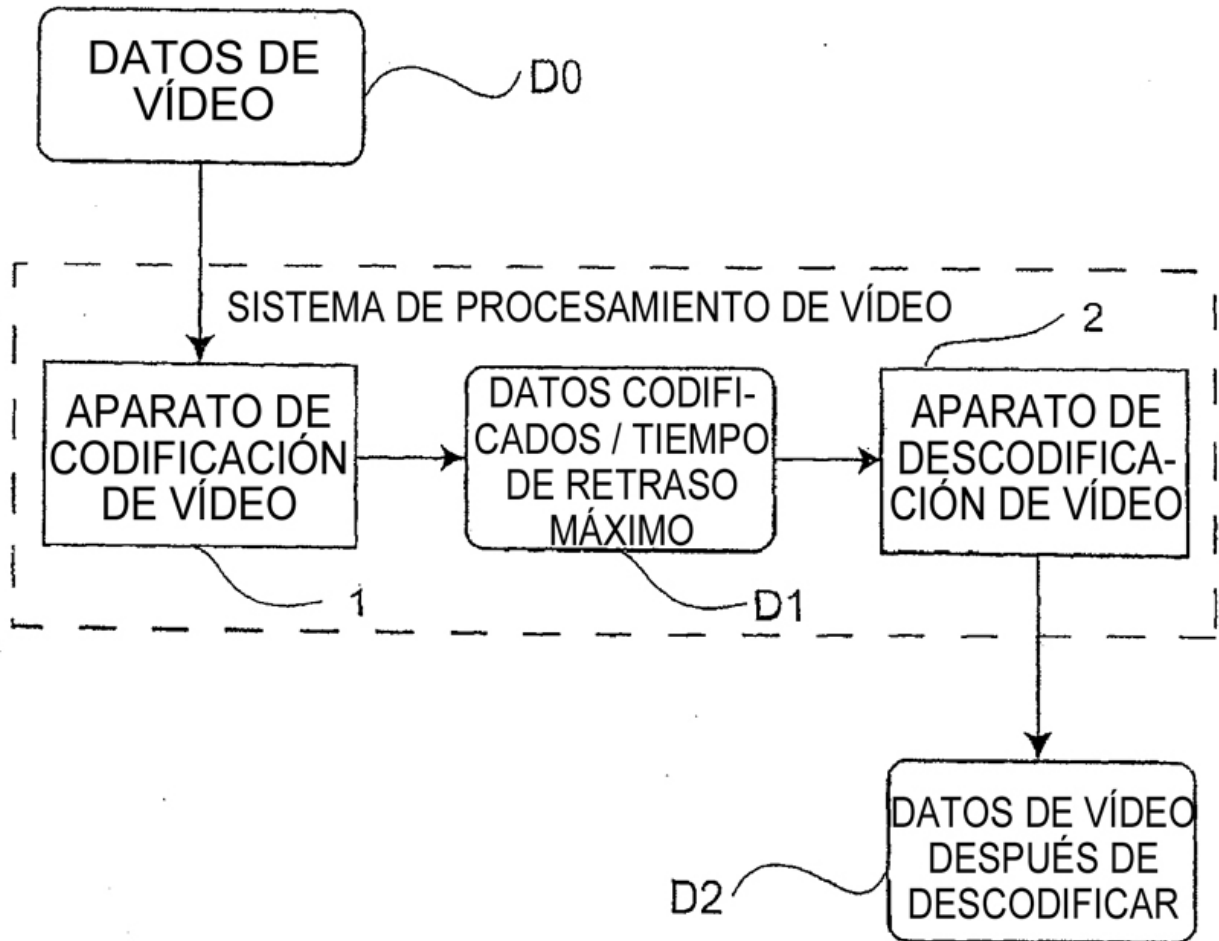


Fig.2

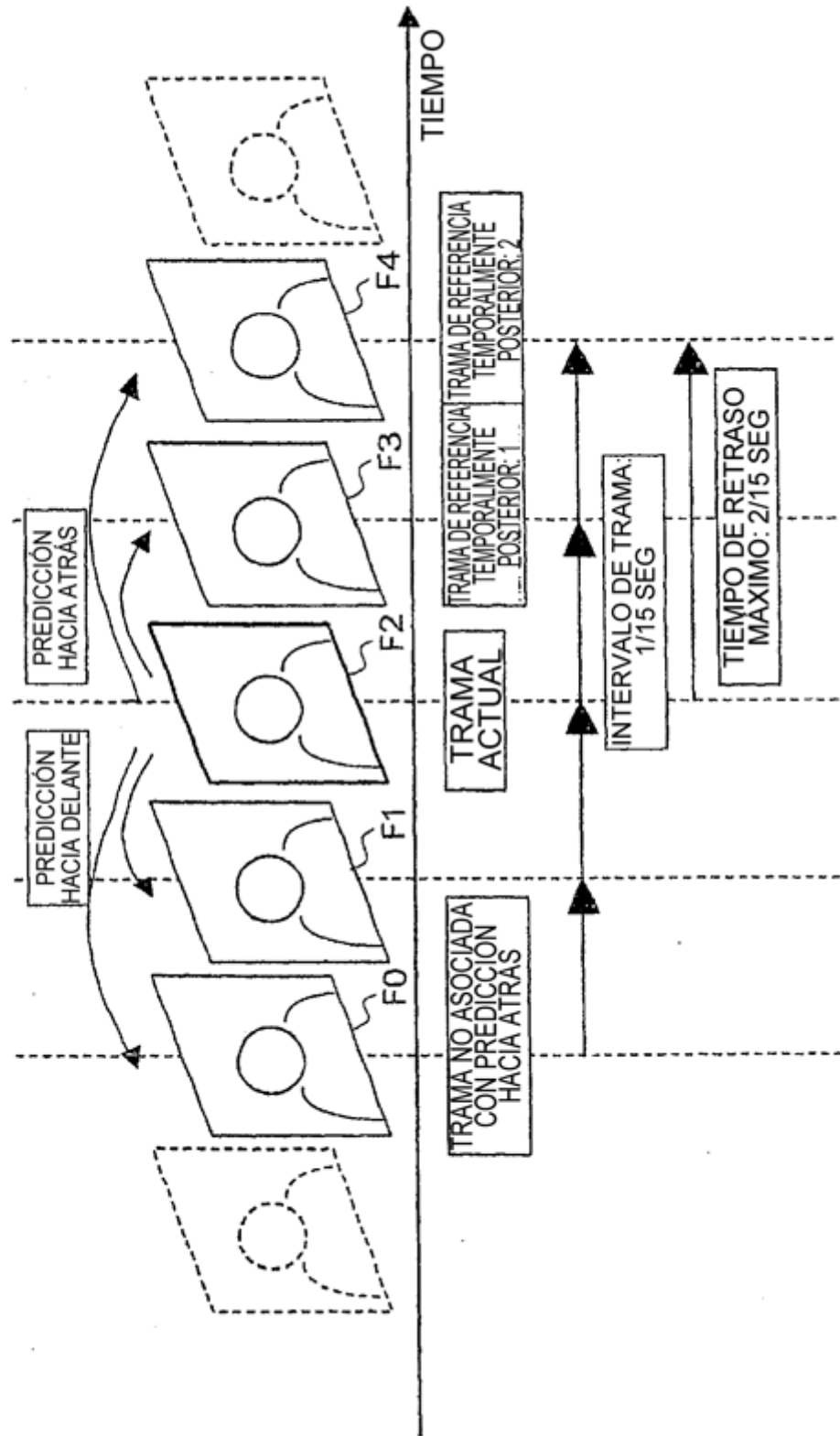




Fig.3

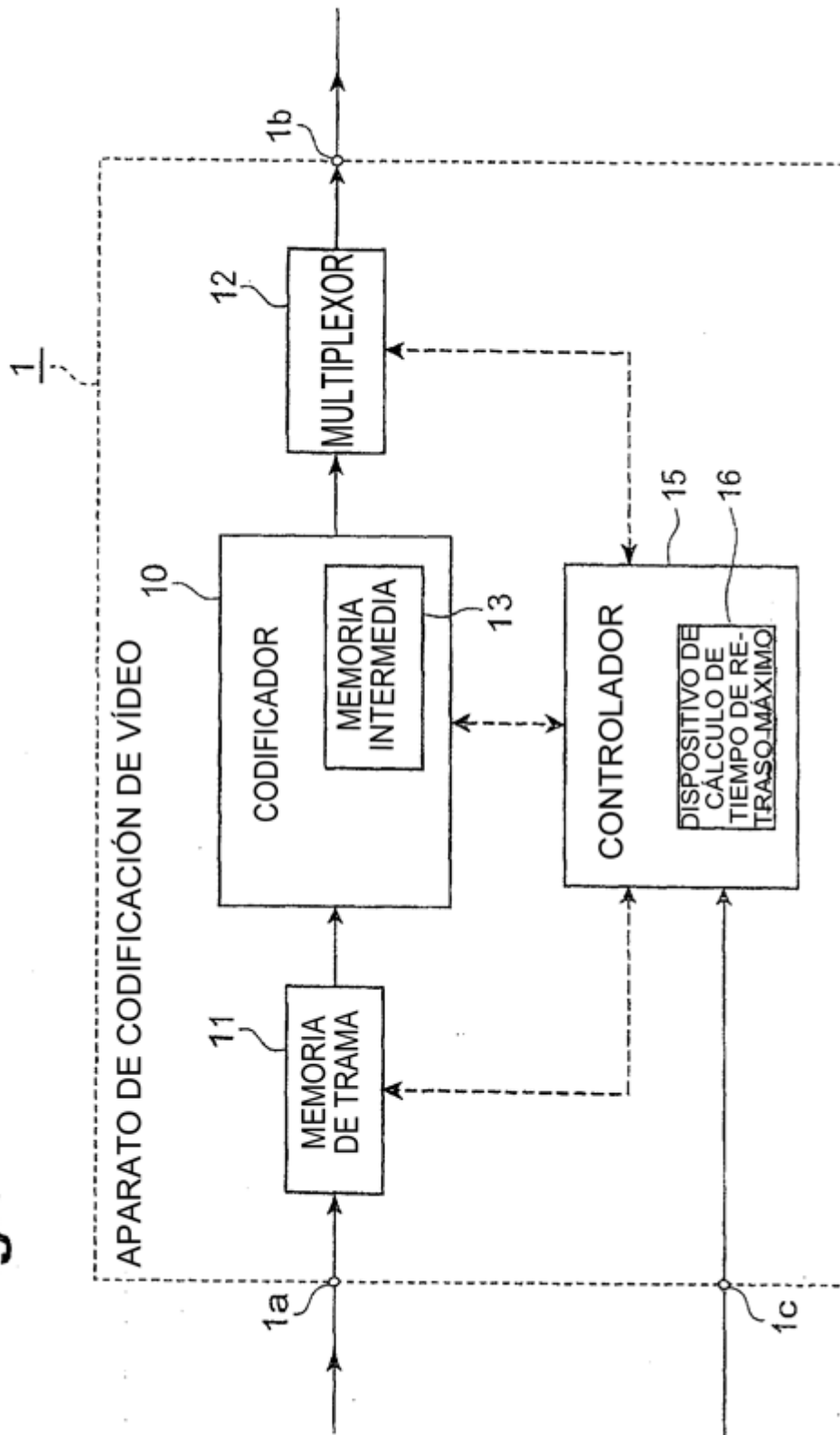
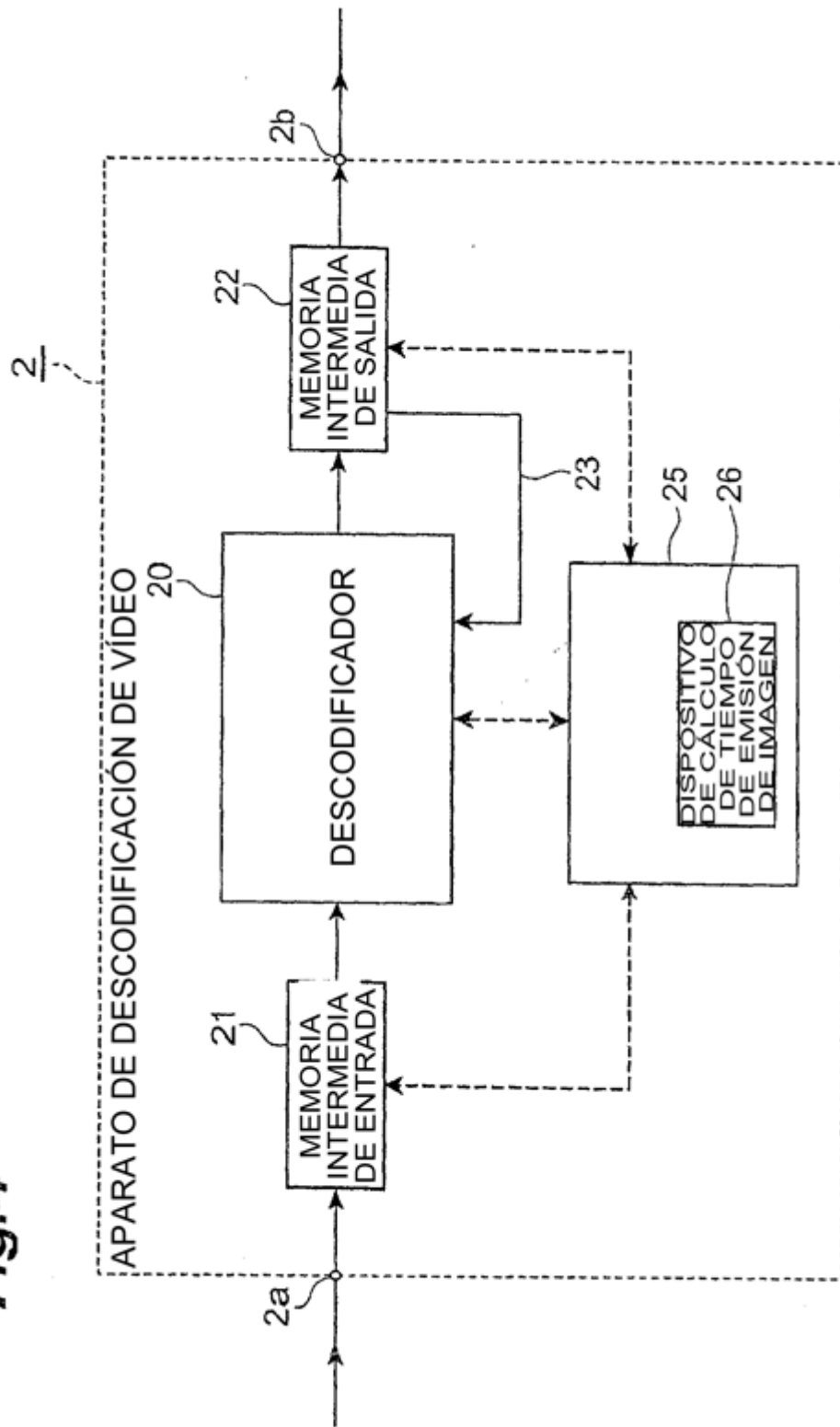


Fig.4



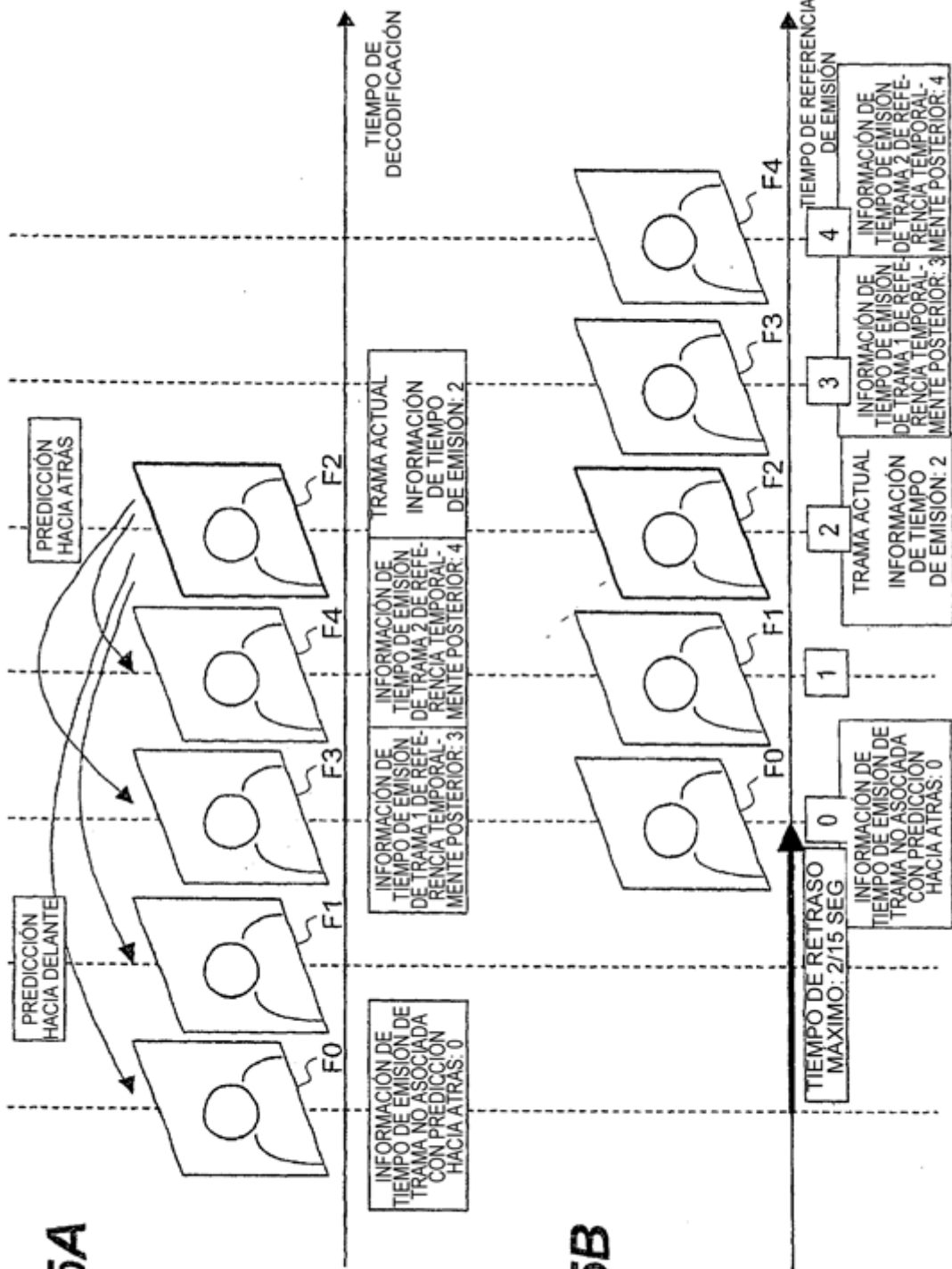
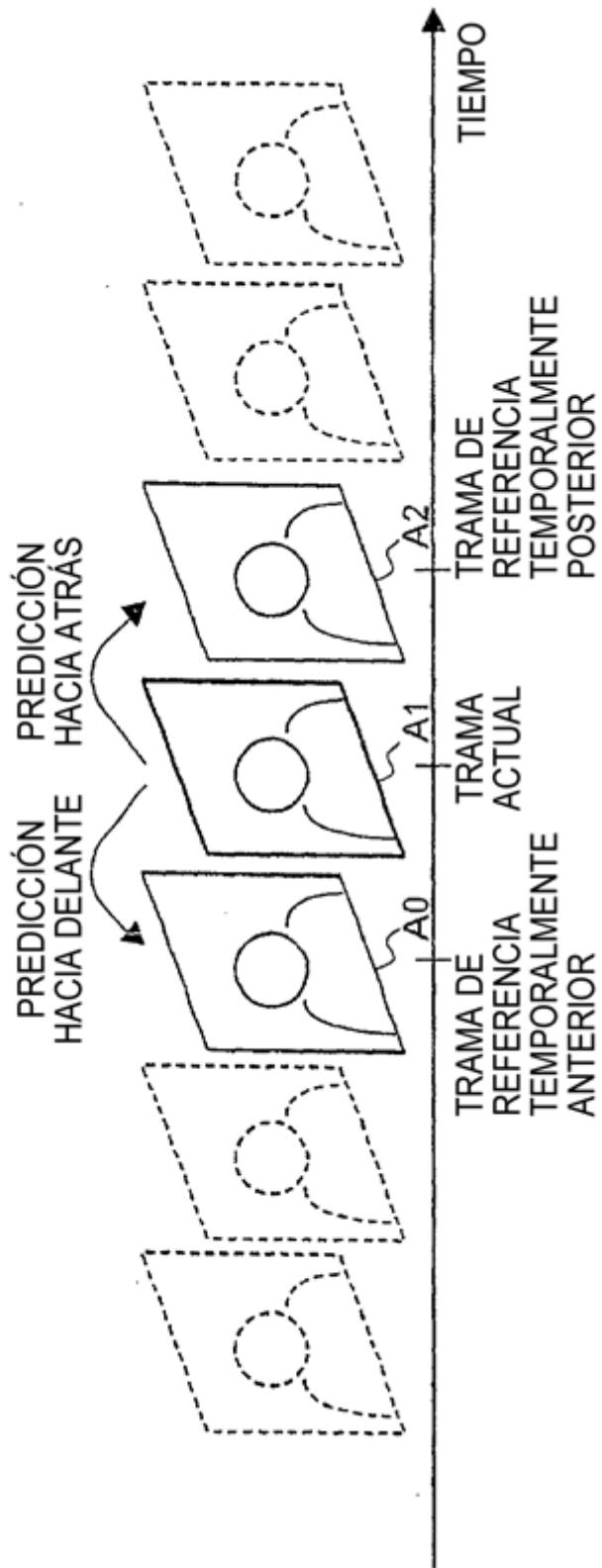


Fig. 5A

Fig. 5B

**Fig.6**



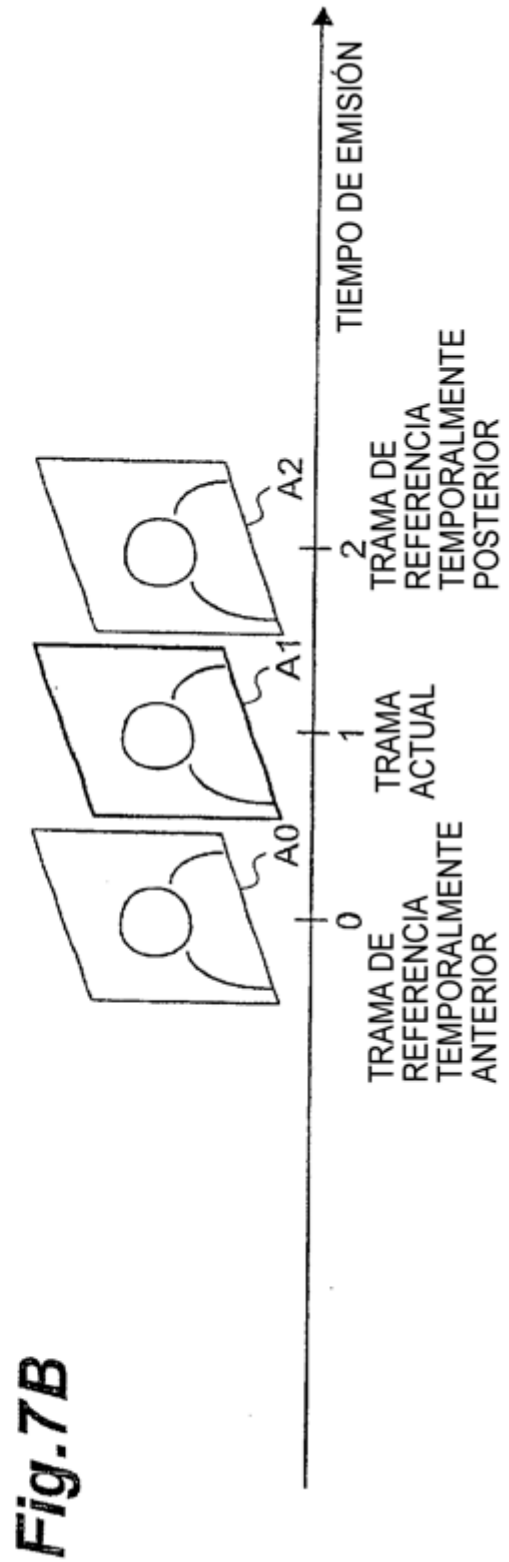
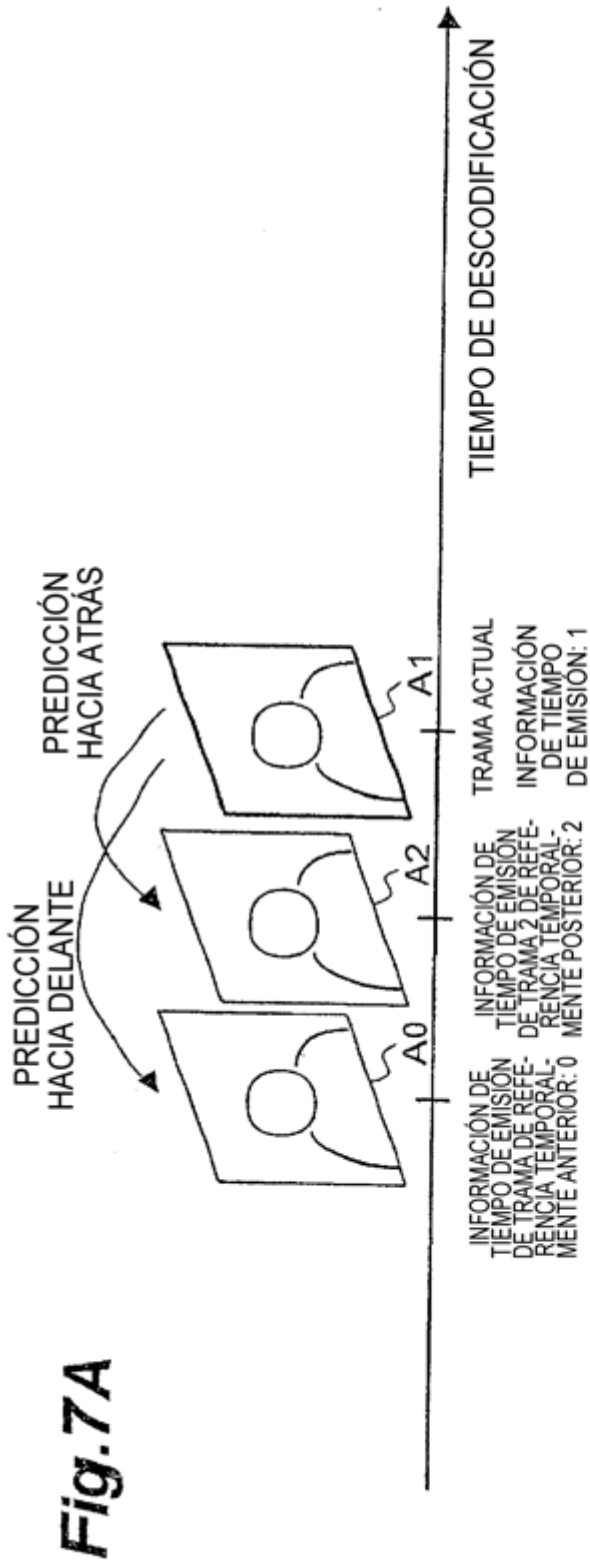
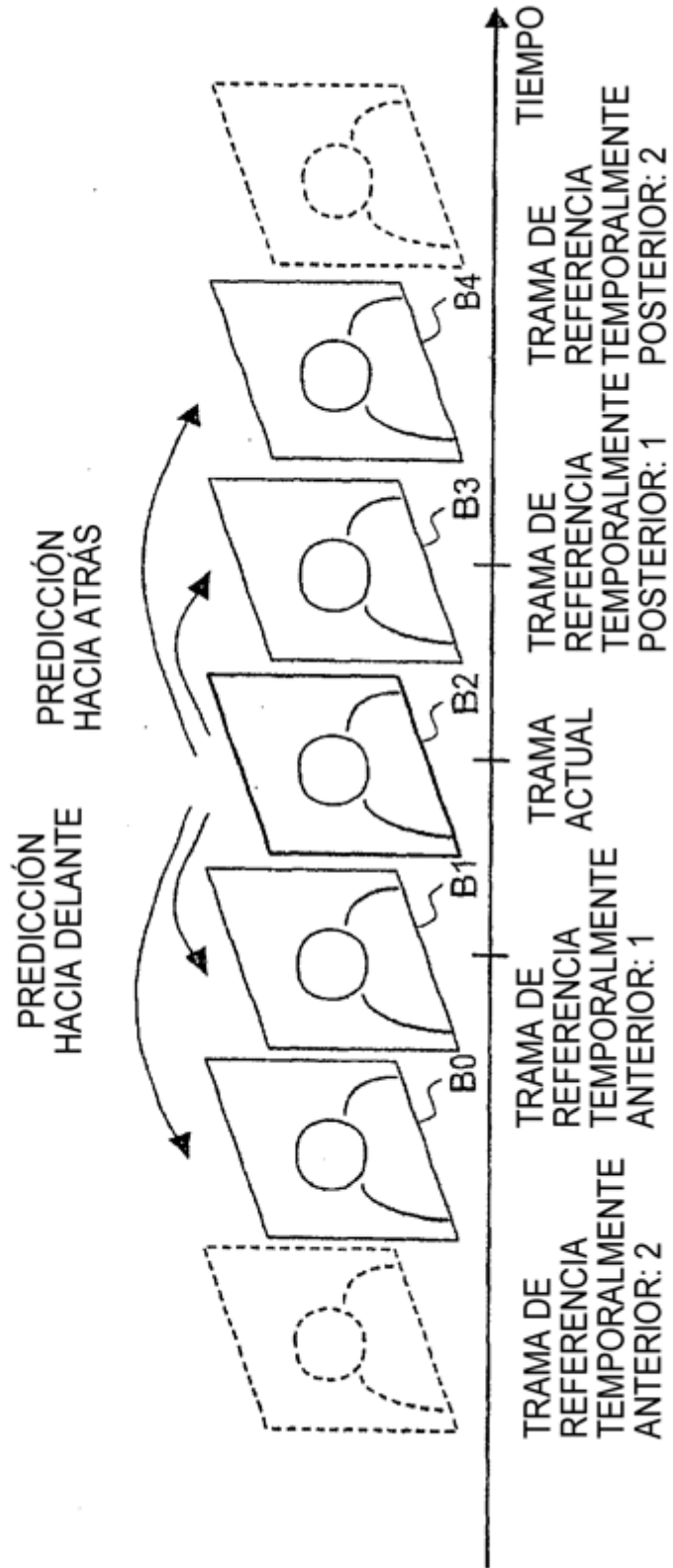
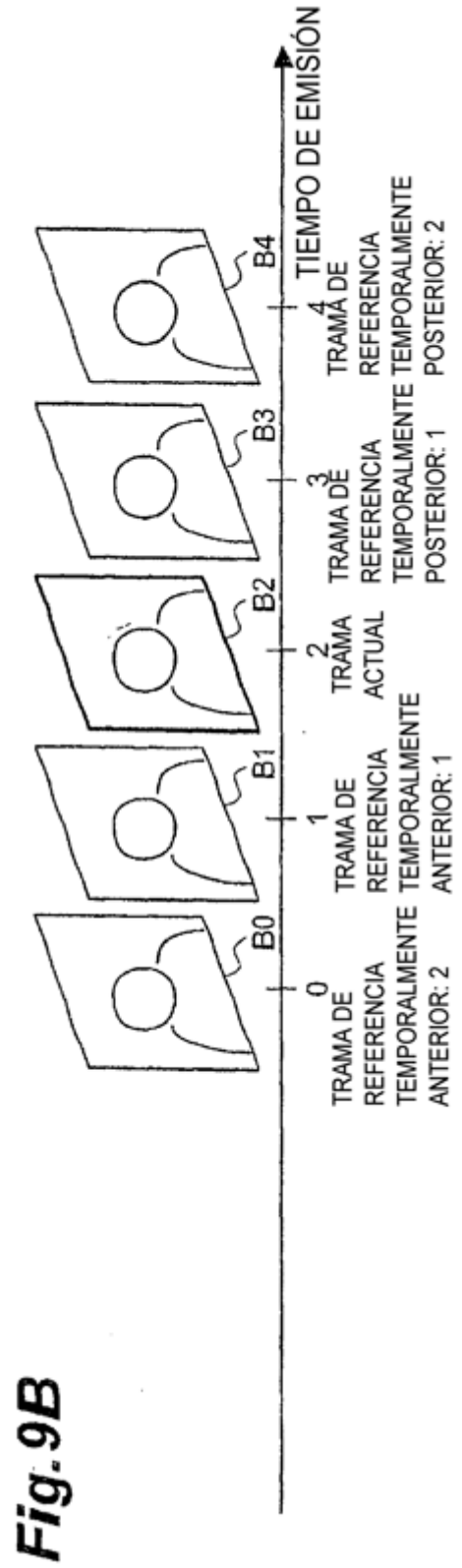
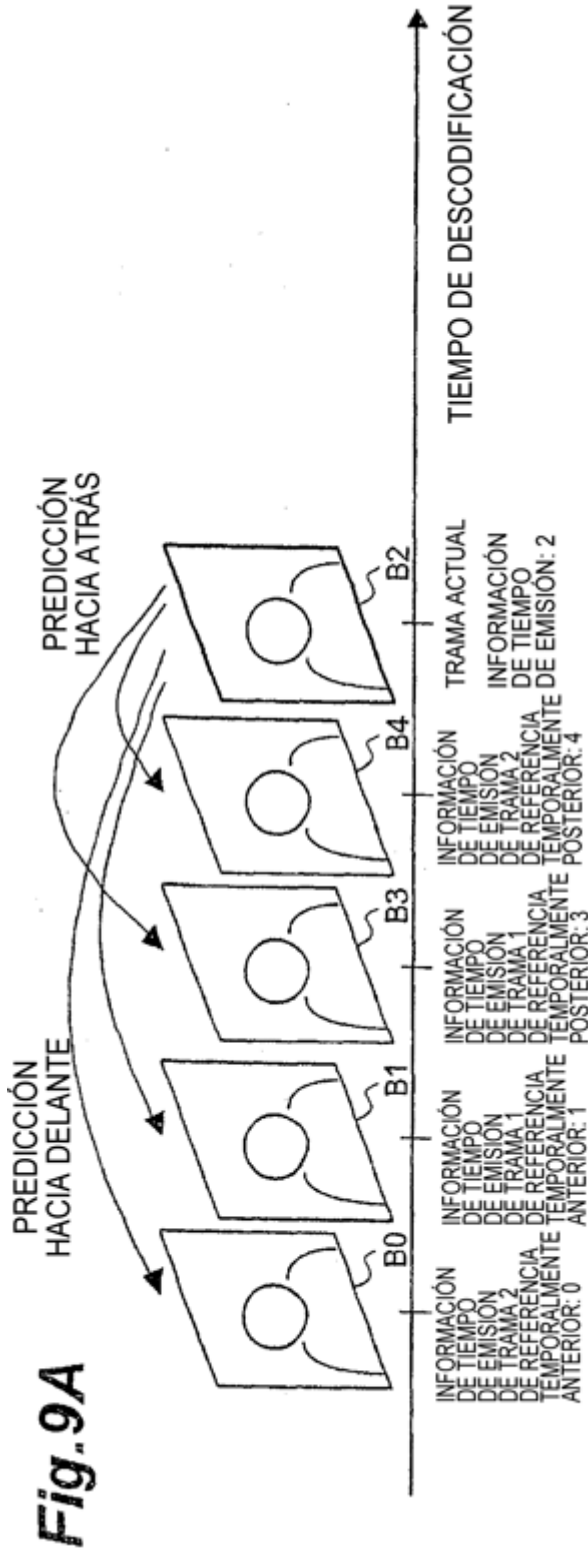


Fig.8





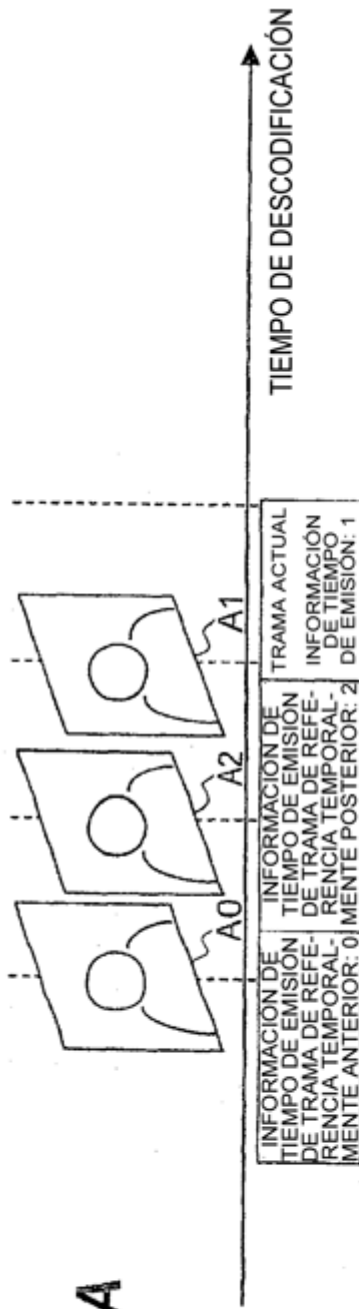


Fig. 10A

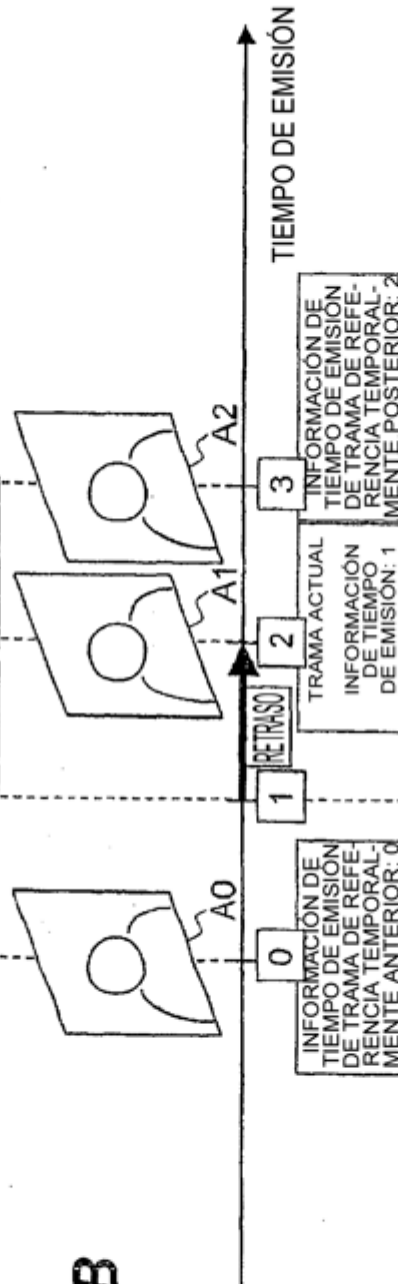


Fig. 10B

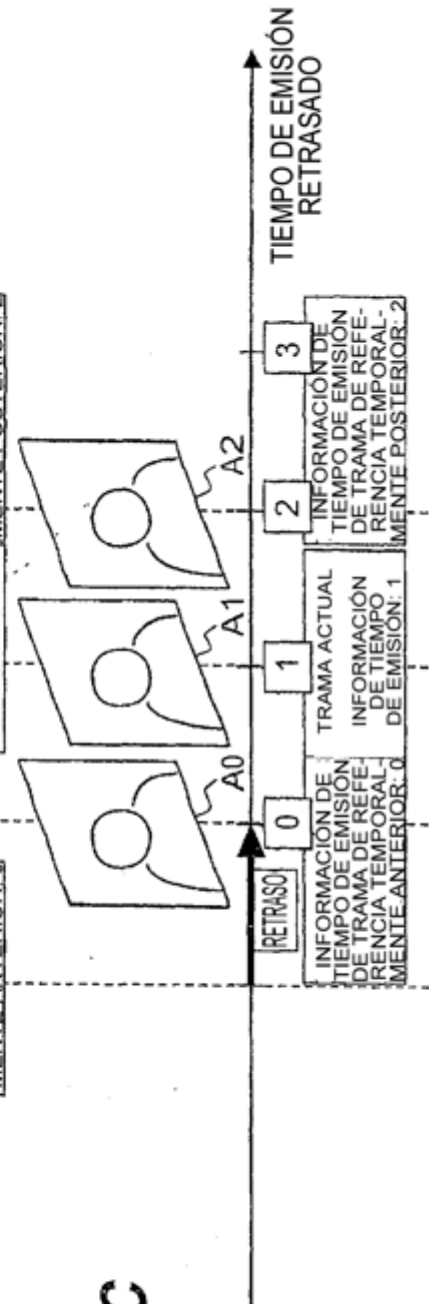


Fig. 10C



