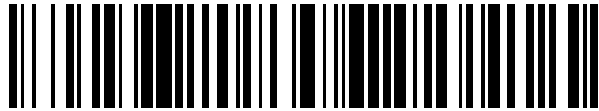


19



OFICINA ESPAÑOLA DE  
PATENTES Y MARCAS

ESPAÑA



11 Número de publicación: **2 572 885**

51 Int. Cl.:

**H04N 19/00** (2014.01)

12

TRADUCCIÓN DE PATENTE EUROPEA

T3

96 Fecha de presentación y número de la solicitud europea: **04.01.2008 E 08700197 (0)**

97 Fecha y número de publicación de la concesión europea: **30.03.2016 EP 2105028**

54 Título: **Método y sistema para codificar una señal de vídeo, señal de vídeo codificada, método y sistema para decodificar una señal de vídeo**

30 Prioridad:

**12.01.2007 EP 07100470**

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

**02.06.2016**

73 Titular/es:

**KONINKLIJKE PHILIPS N.V. (100.0%)  
HIGH TECH CAMPUS 5  
5656 AE EINDHOVEN, NL**

72 Inventor/es:

**BRULS, WILHELMUS, H., A. y  
BOURGE, ARNAUD, P., H.**

74 Agente/Representante:

**ISERN JARA, Jorge**

**ES 2 572 885 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 y sistema para codificar una señal de vídeo, señal de vídeo codificada, método y sistema para decodificar una señal de vídeo

5 Campo de la invención

10 La invención se refiere al campo general de la codificación de vídeo (o compresión de vídeo), y en las realizaciones a la denominada codificación de vídeo escalable. Presenta un método y sistema para codificar una señal de vídeo en el que, se realiza la compresión de imagen, por ejemplo compresión de imagen escalable, y se usan múltiples predicciones temporales en las que se almacenan múltiples fotogramas en una memoria en una unidad de predicción.

15 La invención también se refiere a, por ejemplo, la señal de vídeo escalable codificada y a un método y sistema para decodificar una señal de vídeo.

Descripción de la técnica anterior

20 Los sistemas de visualización de imagen a menudo reciben flujos de datos comprimidos. Se conoce una diversidad de técnicas de compresión de imagen para reducir la cantidad de datos de imagen que deben almacenarse o transmitirse. En compresión de vídeo se hace uso de la predicción, en la que el contenido de un fotograma o parte de un fotograma se predice desde el contenido de uno o más fotogramas recibidos o generados anteriores. En procesamiento de señal de vídeo la señal está comprendida de fotogramas intracodificados e intercodificados, por ejemplo fotogramas I, fotogramas P y fotogramas B. Los fotogramas I están intra-codificados. Los fotogramas P y B se denominan como fotogramas intercodificados. Los fotogramas intra-codificados pueden reconstruirse sin ninguna referencia a otros fotogramas; los fotogramas intercodificados se reconstruyen usando datos de otros fotogramas (predicción hacia delante o hacia atrás). Los fotogramas P y B únicamente contienen información o cambios entre los fotogramas I, a menudo expresados en vectores de movimiento para macrobloques. En procesamiento de señal de vídeo convencional la referencia es relativamente sencilla, y como máximo están referenciados dos fotogramas, los fotogramas P se predicen hacia delante desde los fotogramas I y los fotogramas B se predicen hacia delante y hacia atrás desde fotogramas I y P. Usando vectores de movimiento de estimación de movimiento puede encontrarse que se usan para estimación de movimiento de partes (macrobloques) de un fotograma. Algunas normas de compresión de vídeo más complejas tales como la norma de compresión AVC tienen una posibilidad de muchas múltiples predicciones. Se hace un número relativamente grande de predicciones temporales (es decir hacia delante o hacia atrás). No solamente se consideran los fotogramas más cercanos en tiempo para realizar las predicciones, sino también fotogramas más alejados en el tiempo. En una memoria intermedia se almacenan varios fotogramas en una memoria para usarse para predicción temporal. A medida que el tiempo progresa los fotogramas se desplazan a través de una memoria intermedia en la memoria y los más antiguos acaban fuera de la memoria intermedia a medida que se almacena un nuevo fotograma 'fresco'.

40 Aunque los esquemas existentes son útiles, los inventores se han dado cuenta que existe una posibilidad de mejora de la codificación. Los métodos conocidos tienen una restricción en la predicción. Suspendiendo la restricción se abren nuevas posibilidades sin requerir cambios importantes a la codificación de vídeo existente (escalable), y por lo tanto proporcionar oportunidades novedosas que son fáciles de desarrollar industrialmente.

45 Sumario de la invención

50 Para este fin el método de codificación de acuerdo con la invención está caracterizado como se describe en la reivindicación 1.

Los fotogramas de predicción se producen en la parte de codificación de mejora del codificador, los fotogramas de predicción que se producen en una parte de codificación de flujo de base se producen por separado de la parte de mejora del codificador (aunque aún dentro del codificador, cuando se observa en un sentido más amplio).

55 La invención está basada en la siguiente idea:

En la norma en el decodificador de mejora se usan varias predicciones temporales. Cuanto más antiguo es un fotograma en el almacenamiento, menor de media es la importancia de la predicción temporal basándose en ese fotograma. En la invención se omite el dogma de la predicción temporal usando el espacio reservado para una o más de las predicciones temporales para una predicción separada obtenida del codificador de flujo de base. Esto reducirá ligeramente la precisión de la predicción temporal en el codificador de mejora basándose en los fotogramas de predicción producidos convencionalmente. Sin embargo, la información sobrescrita es relativamente redundante/menos útil. Sobrecribir la información muy a menudo por lo tanto no reduce gravemente la calidad de imagen. La información que puede escribirse en los intervalos sobrescritos, sin embargo, abre posibilidades que no son conseguibles con las normas actuales, o únicamente con cambios muy complejos a los procedimientos existentes de modo que ya no son compatibles con las normas existentes y/o necesitan un gran aumento en bits en

la señal de vídeo. En realizaciones sencillas de la invención la predicción sobrescrita es una de las predicciones más antiguas. Las predicciones más antiguas (más prolongadas en memoria) normalmente son las menos importantes. Como alternativa el método puede comprender un algoritmo para seleccionar la predicción o predicciones menos importantes antes de sobrescribir.

5 El método de acuerdo con la invención no requiere más bits o cambios importantes a la norma AVC existente.

10 En una realización de la invención la memoria es una memoria en un codificador/decodificador de mejora y el fotograma producido por separado es un fotograma sobreescalado/des-entrelazado desde un codificador/decodificador de base. Esto posibilita, sin ningún cambio de sintaxis, usar la AVC convencional para compresión escalable.

En otra realización el fotograma externo comprende vistas de profundidad.

15 Esto permite generar codificación multivista de vídeo en 3D.

20 En otra realización más, ilustrada mediante la Figura 7, el flujo de mejora proporciona una velocidad de fotogramas superior, con una relación diferente de 2. Una aplicación prevista para un caso de este tipo es una capa base a 50 Hz pretendida para aparatos de televisión, y una capa de mejora a 75 Hz pretendida para paneles de LCD de ordenador (o cualquier otra pantalla avanzada). De hecho, aumentar la velocidad de fotogramas a al menos 75 Hz proporciona mejoras significativas mientras que exige una tasa de bits inferior que pasar directamente a 100 Hz. Además, 75 Hz es la tasa de refresco nativa de muchas pantallas (principalmente en el mundo informático). Por lo tanto, se ha de preferir una solución 50/60 Hz compatible con mejora de 75/90 Hz, pero no puede realizarse ni con MPEG2 ni con SVC.

25 Esta compatibilidad puede realizarse eficazmente basándose en la presente invención.

Breve descripción de los dibujos

30 Estos y otros aspectos ventajosos de la invención se describirán en más detalle usando las siguientes figuras.

Las Figuras 1a a 1 muestran el flujo de procesamiento de un método de post-procesamiento, que incluye un método de codificación (Figura 1a) y de decodificación (Figura 1b) de acuerdo con una realización de la invención, en el que la Figura 1c ilustra una realización más compleja.

35 Las Figuras 2 a 5 ilustran adicionalmente la realización de la invención como se muestra en las Figuras 1a a 1c.

La Figura 6 ilustra una segunda realización de la invención.

La Figura 7 ilustra una realización adicional de la invención en la que el flujo de base y el flujo de mejora tienen diferentes velocidades de fotogramas.

40 La Figura 8 muestra el principio básico de un aparato de compresión de vídeo que tiene memorias de predicción (típicamente para predicción temporal) en el que al menos una de las menos usables se usa para almacenar una predicción de un medio de predicción alternativo.

Las figuras no están dibujadas a escala. En general, se indican componentes idénticos mediante los mismos números de referencia en las figuras.

45 Descripción detallada de realizaciones preferidas

En resumen la invención puede describirse como sigue:

50 Algunas normas de compresión de vídeo tienen una posibilidad de usar una multitud de predicciones para, por ejemplo, estimación de movimiento/compensación de movimiento. Los inventores se han dado cuenta que algunas de estas predicciones son más redundantes o menos útiles, por ejemplo, cuando se usa AVC se usan múltiples predicciones temporales y el IC está prediseñado para hacer frente a esto. En la invención se omite el dogma de predicción temporal sobrescribiendo una última de estas predicciones (es decir, normalmente la menos precisa, especialmente para movimiento descontrolado) para codificar otra predicción. El decodificador se comporta de manera similar.

60 La Figura 1 muestra un flujo de procesamiento de una realización de nuestra invención usada, un método de codificación y de decodificación. Esto se ilustra a continuación:

Sitio de codificador:

La Figura 1a ilustra una realización del método y sistema de acuerdo con la invención.

65 Las abreviaturas usadas en las figuras indican:

DCT: Transformada del Coseno Discreta  
 Q: Cuantificación  
 DCT<sup>-1</sup>: Transformada del Coseno Discreta inversa  
 Q<sup>-1</sup>: cuantificación inversa

5 VLC: Codificación de Longitud Variable

Base str: flujo de base

Enh str: flujo de mejora

ME/MC: estimación de movimiento/compensación de movimiento

R<sub>0</sub>, R<sub>1</sub> etc.: fotogramas en la memoria usados en la ME/MC

10 MB<sub>i</sub>: MacroBloque i

PTS<sub>i</sub>: Indicación de Tiempo de Presentación i

AVC: codificación de Vídeo Avanzada, la codificación de vídeo avanzada es por ejemplo las normas de codificación H.264 y MPEG4.10.

SEI: Información de Mejora Complementaria.

15 La Figura 1a ilustra un método en el que se hace uso de un flujo de base y un flujo de mejora. Las cantidades masivas de datos intrínsecas en muchos métodos de transmisión de imagen suponen problemas significativos. Más particularmente, cada fotograma de imagen digital es una imagen fija formada de una matriz de píxeles de acuerdo con la resolución de visualización de un sistema particular. Como resultado, las cantidades de información digital en  
 20 bruto incluidas son a menudo masivas. Para reducir la cantidad de datos que deben enviarse, se usan esquemas de compresión para comprimir los datos. Se han establecido diversas normas o procesos de compresión de vídeo, que incluyen MPEG-2, MPEG-4 y H.263.

25 Se posibilitan muchas aplicaciones cuando los datos de imagen están disponibles a diversas resoluciones y/o calidades en un flujo. Los métodos para conseguir esto se denominan generalmente como técnicas de escalabilidad. Hay tres ejes en los que se puede desplegar escalabilidad. El primero es escalabilidad en el eje de tiempo, a menudo denominado como escalabilidad temporal. En segundo lugar, hay escalabilidad en el eje de calidad, a menudo denominado como escalabilidad de señal a ruido o escalabilidad de grano fino. El tercer eje es el eje de  
 30 resolución (número de píxeles en imagen) a menudo denominado como escalabilidad espacial o codificación en capas. En la codificación en capas, el flujo de bits se divide en dos o más flujos de bits o capas. Cada capa puede combinarse para formar una única señal de alta calidad. Por ejemplo, la capa de base puede proporcionar una señal de vídeo de calidad inferior, mientras que la capa de mejora proporciona información adicional que puede mejorar la imagen de capa base. La Figura 1a ilustra un método y sistema para proporcionar una señal en capas, en este caso  
 35 tiene un flujo de base (base str) y un flujo de mejora (base str), en ocasiones denominado también una capa de mejora.

La señal se divide y se envía a un codificador de base 2 después de haber pasado un filtro paso bajo 1, por ejemplo un filtro nyquist. La señal experimenta en el codificador de base una transformación de coseno discreta (DCT), o cualquier otra transformación similar, tal como por ejemplo usando ondículas, y se cuantifica (Q; en el flujo de datos  
 40 resultante se realiza codificación de longitud variable que proporciona que se envíe y/o almacene el flujo de base. En procesamiento de señal de vídeo la señal a menudo se comprime o intracodifica de fotogramas intercodificados, por ejemplo fotogramas I, fotogramas P y fotogramas B. Los fotogramas I están intra-codificados. Los fotogramas P y B se denominan como fotogramas intercodificados. Los fotogramas intra-codificados pueden reconstruirse sin ninguna referencia a otros fotogramas; los fotogramas intercodificados se reconstruyen usando datos de otros fotogramas  
 45 (predicción hacia delante o hacia atrás). Los fotogramas P y B únicamente contienen información o cambios entre los fotogramas I, a menudo expresado en vectores de movimiento para macrobloques. Para encontrar las diferencias la señal original tiene que reconstruirse dentro del codificador. Esto se hace mediante una cuantificación inversa (Q<sup>-1</sup>) y una Transformada de Coseno Discreta inversa (DCT<sup>-1</sup>). Los fotogramas reconstruidos resultantes se usan dentro del codificador para estimar vectores de movimiento. En una disposición sencilla únicamente se usa un  
 50 fotograma reconstruido para estimación de movimiento y compensación de movimiento. Sin embargo, en métodos y sistemas más complejos se usa un número de fotogramas reconstruidos. Esto puede ser útil puesto que la reconstrucción introduce errores que pueden ser mayores para algunos fotogramas reconstruidos que para otros y puede ser también que el fotograma anterior no sea necesariamente el mejor fotograma para tomar como una posición de inicio para estimación de vectores de movimiento o diferencias entre fotogramas. El codificador  
 55 comprende un registro de desplazamiento en el que para comparación, los datos de un número de fotogramas reconstruidos se almacenan en un registro de desplazamiento para uso en la estimación de movimiento y/o estimación de movimiento, es decir en predicción. Algunas normas de compresión de vídeo más complejas como la norma de compresión AVC tienen una posibilidad de muchas múltiples predicciones. Se realiza un número relativamente grande de predicciones temporales (es decir hacia delante o hacia atrás). No solamente se consideran  
 60 los fotogramas más cercanos en tiempo para hacer las predicciones, sino también fotogramas más alejados en el tiempo. En una memoria intermedia se almacenan varios fotogramas en una memoria intermedia para usarse para predicción temporal. A medida que el tiempo progresa los fotogramas se desplazan a través de la memoria intermedia y los más antiguos acaban fuera de la memoria intermedia a medida que se almacena un nuevo fotograma nuevo 'fresco'.

65

El flujo de base resultante puede difundirse, recibirse y mediante un decodificador, visualizarse como está, aunque el flujo de base no proporciona una resolución que se considerara como alta definición.

En la Figura 1a el sistema comprende también un codificador de mejora 3. En el codificador de mejora 3 se realiza un método similar como en el codificador de flujo de base 2, con únicamente esta diferencia de que se trata el flujo de mejora (es decir la diferencia entre la señal original y el flujo de base). El decodificador de mejora comprende un medio para Transformación de Coseno Discreta y cuantificación de la señal de capa mejorada, y para reconstrucción ( $DCT^{-1}$ ,  $Q^{-1}$ ). Los datos reconstruidos del flujo de mejora para los fotogramas  $R_0$ ,  $R_1$  etc., se almacenan y usan para estimación de movimiento y compensación de movimiento para el flujo de mejora. Un experto en la materia estará familiarizado con estos métodos.

Los inventores se han dado cuenta de que los fotogramas almacenados no tienen todos la misma relevancia, en particular, pero no exclusivamente los fotogramas restantes menos prolongados pueden ser también menos importantes. Cuanto más antiguo es un fotograma en el almacenamiento, menor de media es la importancia de la predicción temporal basándose en ese fotograma. En la invención se omite el dogma de predicción temporal usando el espacio reservado para la última de las predicciones temporales (es decir los 'fotogramas más antiguos') para una predicción separada. Esto reducirá ligeramente la precisión de la predicción temporal. Sin embargo, la información sobrescrita es relativamente redundante/menos útil. Sobrecribir la información muy a menudo por lo tanto no reduce gravemente la calidad de la imagen. La información que puede escribirse en los intervalos sobrescritos, sin embargo, abre posibilidades que no son conseguibles con las normas actuales, o únicamente con cambios muy complejos a los procedimientos existentes de modo que ya no son compatibles con normas existentes y/o necesitan un gran aumento en bits en la señal de vídeo. La invención podría usarse sobrescribiendo uno o más de los fotogramas más antiguos.

La Figura 1a ilustra una realización del método y sistema de la invención. Los datos más antiguos en la memoria en el codificador de mejora se sobrescriben mediante un fotograma sobreescalado/entrelazado producido en el codificador de base. Por lo tanto antes de codificar el último (más prolongado en la memoria 5 del codificador de mejora) fotograma de referencia  $R_{n-1}$  en el codificador de mejora 3 se sobrescribe con un fotograma sobreescalado/des-entrelazado 4 desde el codificador/decodificador de base 2 con también PTSi (Indicación de Tiempo de Presentación i). El fotograma más prolongado en memoria en la memoria 5 del codificador de mejora 3 se sobrescribe por lo tanto mediante un fotograma producido por separado, en concreto un fotograma sobreescalado/des-intercalado 4 desde el codificador de flujo de base 2. Aunque esto puede reducir ligeramente la precisión de la estimación de movimiento/compensación de movimiento realizada dentro del codificador de mejora esto permite, sin ningún cambio de sintaxis, usar AVC para compresión escalable. Esto es relevante, por ejemplo, para introducir HDTV 1080p a velocidad de fotograma completa de una manera compatible hacia atrás (a HDTV 720p o 1080i). El principio también funciona para compresión no escalable. La Figura 1a muestra un codificador escalable. Puesto que los recursos están limitados a 180i, el esquema de acuerdo con la invención posibilita con únicamente menores modificaciones un hardware 1080p. La última predicción  $R_{n-1}$  ahora por ejemplo contiene un candidato de sobreescalado adaptativo de contenido inteligente. Como una realización más elaborada, por ejemplo las dos últimas predicciones temporales 'normales' ( $R_{n-1}$ ,  $R_{n-2}$ ) podrían sobrescribirse, por ejemplo, para ser las predicciones (fotogramas) producidas usando dos algoritmos de des-intercalación diferentes.

En las realizaciones únicamente se sobrescribe el fotograma más antiguo en memoria en el codificador de mejora. Dentro del alcance de la invención, dependiendo del número de fotogramas en memoria podría sobrescribirse más de uno de los fotogramas más antiguos. También, en esta realización se hace una elección sencilla, el fotograma más antiguo se sobrescribe. En las realizaciones, uno de los últimos n fotogramas podría sobrescribirse en el que se realiza una elección de cuál es el que se sobrescribe basándose en una estimación de la importancia (clasificación) del conjunto de fotogramas más antiguos a sobrescribir, los fotogramas sobrescritos que necesariamente no necesitan ser el fotograma más antiguo, pero podrían ser el siguiente al fotograma más antiguo.

La Figura 1b muestra el lado del decodificador. VLD indica decodificación de longitud variable. Un fotograma 4 producido en el decodificador de flujo de base 2 se usa para sobrescribir el fotograma de predicción de fotograma  $R_{n-1}$  en la memoria 5' del decodificador de mejora.

La Figura 1c ilustra una realización más compleja del diseño mostrado en la Figura 1a. La diferencia es que se usa un filtro y adicionalmente que los datos relacionados con el filtro y el sobreescalado se insertan en la SEI (Información de Mejora Complementaria). Estos datos se usan en un decodificador para establecer en el extremo del decodificador los mismos parámetros para filtrado y sobreescalado, de modo que los parámetros usados en el método para codificar en el codificador puedan usarse en la decodificación también.

En esta realización, cuando un fotograma "desde fuera" se pone en la memoria intermedia de instantáneas decodificadas, se proporciona una cts (indicación de tiempo de composición) antigua de modo que nunca se visualizará, y se marca también como "referencia a largo plazo" de modo que no acaba fuera de la memoria.

La Figura 2 ilustra en un diagrama de flujo las diferentes etapas en la mejora para asegurar que el fotograma externo (base sobreescalada con PTSi) no sobrescribe el fotograma  $R_{n-1}$  en la memoria.

La idea básica de la invención, en concreto que el fotograma de predicción más prolongado en memoria se sobrescribe con un fotograma de predicción producido externamente no es solo útil en la realización ilustrada en las Figuras 1a a 1c.

5 Se proporciona otra opción en la Figura 2b; se proporciona un ejemplo cuando es seguro sobrescribir un fotograma de referencia en memoria. En este punto es importante que un fotograma se sobrescriba únicamente cuando ya se haya enviado a la salida (pantalla). A medida que el fotograma entra en la memoria se llena con el fotograma (memoria 0, memoria 1 etc.). En esta configuración el número de fotogramas de referencia es tres. Esto significa que los primeros tres fotogramas es imposible que sobrescriban fotogramas, puesto que se necesitan tres fotogramas.  
 10 Cuando se ha visualizado un fotograma, ese fotograma puede sobrescribirse en la memoria. En este ejemplo por lo tanto los fotogramas que se han colocado dentro de los círculos pueden sobrescribirse. Se ha de observar que estos fotogramas son los fotogramas más antiguos, pero no necesariamente el fotograma más antiguo en la memoria. En la columna 4 por ejemplo el fotograma ya visualizado  $I_0$  fotograma puede sobrescribirse, pero no como el fotograma aún no visualizado  $B_1$ . El rectángulo indica fotogramas usados para referencia para  $B_3$  y  $P_6$  (primer rectángulo) y  $B_5$  y  $P_7$  (último rectángulo).

La Figura 3 ilustra una realización en la que un fotograma de base  $I_0$  decodificado sobreescalado sobrescribe en la memoria (memoria intermedia) del codificador/decodificador de mejora el último fotograma de predicción de referencia ( $R_2$ ) en el codificador/decodificador de mejora, que en este ejemplo es  $P_6$ .

La Figura 4 ilustra una realización en la que un  $P_2$  decodificado sobreescalado, con la indicación de tiempo PTS2; el fotograma de flujo de base sobrescribe en la memoria (memoria intermedia) del codificador/decodificador de mejora el último fotograma de predicción de referencia ( $R_2$ ) en el codificador/decodificador de mejora, que en este ejemplo es  $P_4$ .

La Figura 5 ilustra una realización en la que el flujo de base tiene una salida de 720p y el flujo de mejora una salida de 1080p. Un flujo de base sobreescalado/desintercalado con la indicación de tiempo PTS1 sobrescribe el fotograma  $R_{n-1}$  en la memoria del codificador de mejora AVC. Esto permite introducir HDTV 1080p a velocidad de fotograma completa de una manera compatible hacia atrás (a HDTV 720p o 1080i), que no es posible o a costes mucho mayores con métodos conocidos.

También podemos aplicar en este principio de sobreescritura uno de más fotogramas de referencia para el caso multivista, de modo que no tenemos que cambiar tampoco ninguna sintaxis, y básicamente podemos usar AVC normal también para esta aplicación.

Ha habido intentos para compresión eficaz de casos de multivista. En mpeg ya se ha reconocido que la AVC normal que hace uso de fotogramas B jerárquicos puede usarse para compresión eficaz de vídeo multivista. En tales métodos los fotogramas de las diferentes vistas están intercalados y se asegura que los fotogramas de vista no de referencia (no-C) están siendo del tipo B jerárquico (B almacenado, almacenados). Esto permite para cada MB tanto predicción espacial (vecina) como temporal (desde la misma vista).

Especialmente es muy interesante el caso con una vista central (C) y una vista izquierda (L) y una vista derecha (R) para vídeo estereoscópico. En este punto la vista central C podría ser la vista de referencia y las vistas L y R serían los fotogramas B jerárquicos.

Aunque la mejor solución actual (AVC + fotogramas B jerárquicos) es una gran mejora, aún la tasa de bits del flujo de AVC de 3 vistas es alta para muchas aplicaciones con respecto a solo un flujo C de AVC  $(1+2*0,4/0,5)= 1,8..2^*$ .

En MPEG se está intentando ampliar AVC con nuevos modos (nueva sintaxis), para incluir IBR (presentación de imagen fuera de 2D + profundidad) como unos nuevos modos de predicción para mejorar la eficacia del codificador AVC multivista. Pero esto presentará dificultades con respecto a la implementación, debido a los muchos cambios que esto significa para AVC.

Los inventores, sin embargo, se han dado cuenta que puede usarse el concepto básico de la invención, es decir el principio de sobrescribir uno o más fotogramas de referencia para el caso multivista, de modo que no tenemos que cambiar tampoco ninguna sintaxis, y básicamente se puede usar AVC normal también para esta aplicación.

Esto se ilustra en la Figura 6, en la que el número mínimo de fotogramas de referencia sería normalmente 4, véase a), en la que en la memoria del codificador AVC se usan cuatro fotogramas de referencia, en concreto Cfw, Cbw, Bsr, Bsl.

Este número de fotogramas de referencia por supuesto podría ampliarse (véase la Figura 6b Cfw, Cbw, Bsr, Bsl, Bsr-1, Bsl-1) y esto proporcionaría alguna ganancia (limitada). Esta situación se presenta en b).

Para una aplicación en 3D real requeriría también 3 vistas de profundidad. Con estas vistas de profundidad y la vista C, podemos aplicar IBR (presentación basada en imagen) para crear aproximaciones (predicciones) de L' y R'. Las

predicciones pueden sobrescribir los fotogramas de referencia menos importantes cada vez antes de empezar a codificar un nuevo fotograma (véase la Figura 6c).

La ganancia ahora será mucho mayor como se ilustra en la parte c) en comparación con b).

5 Una tercera aplicación posibilitada mediante esta invención es crear un flujo de mejora para una velocidad de fotogramas superior con una relación diferente de potencia de 2.

10 De hecho, un fotograma decodificado de la secuencia de 50 Hz se pone en la lista de referencia de la secuencia de 75 Hz, de acuerdo con el esquema general de esta invención. Este fotograma de predicción desde la secuencia de 50 Hz es un buen predictor y ayuda a alcanzar una mejor eficacia de codificación puesto que está temporalmente muy cercano al fotograma para codificar en la versión de 75 Hz. Mirando la Figura 7, donde el *fotograma i* pertenece a la secuencia de 75 Hz mejorada, y el *fotograma i'* pertenece a la secuencia de 50 Hz de base:

- 15 - el fotograma 1' tiene la misma localización temporal que el fotograma 1  
 - el fotograma 2' está temporalmente más cercano al fotograma 2 que cualquier fotograma en la secuencia de 75 Hz  
 - el fotograma 2' está temporalmente más cercano al fotograma 3 que cualquier fotograma en la secuencia de 75 Hz  
 20 - el fotograma 3' tiene la misma localización temporal que el fotograma 4  
 - etc.

En resumen la invención puede describirse mediante:

25 Algunas normas de compresión de vídeo usan múltiples predicciones temporales. Una o más de las predicciones temporales más antiguas se sobrescriben con otra predicción. Una predicción usada en un codificador de mejora se sobrescribe en la realización mediante una predicción producida en un codificador de flujo de base. En otra realización una predicción temporal se sobrescribe mediante una vista en 3D.

30 La invención se refiere a un método y sistema de codificación, así como a un método y sistema de decodificación, como se ha descrito anteriormente a modo de ejemplo.

35 La invención se realiza también en una señal de vídeo que comprende señales de vídeo codificadas e información de control que comprende, por ejemplo, parámetros funcionales para uso en el método de decodificación. La información de control puede comprender datos de acuerdo con cualquiera, o cualquier combinación, de las realizaciones anteriormente descritas.

La información de control encerrada en la señal de vídeo de acuerdo con la presente invención puede comprender una o más de la siguiente información:

40 A: información general, es decir aplicable para toda la señal de vídeo

1. Que el método se use para codificar la señal de vídeo, esto es útil si existiera, en el futuro hay únicamente un método convencional usado  
 45 2. Que una realización particular del método se use para codificar la señal de vídeo  
 3. Parámetros usados en el método, tales parámetros pueden ser por ejemplo

- La predicción particular a sobrescribir (la más prolongada en memoria, o las dos más prolongadas en memoria,  
 50 - Parámetros usados a lo largo de toda la codificación de la señal de vídeo, tales como parámetros para filtros, algoritmos de sobreescalado o algoritmos para determinar cuáles de las predicciones se han de sobrescribir.

B: información dinámica particular.

55 Si la señal de vídeo se genera dinámicamente, es decir ciertas realizaciones realizadas durante la codificación son dependientes del contenido de vídeo; las elecciones realizadas en el codificador pueden incluirse en la señal de vídeo codificada. Por ejemplo, si el método de codificación comprende un algoritmo para realizar una decisión en cuanto a cuál de la predicción se ha de sobrescribir por ejemplo estimando la importancia de la predicción a sobrescribir posiblemente en comparación a la predicción que se sobrescribirá, existe una elección de incluir los  
 60 detalles de dicho algoritmo o algoritmos de selección en la señal de vídeo (y por lo tanto incluir la información general en la señal de vídeo). Sin embargo, esto funciona únicamente si el decodificador comprende dicho algoritmo. Si el decodificador no lo comprende, la información dinámica particular puede enviarse, es decir se especifica para una parte particular de la señal de vídeo (por ejemplo mediante una bandera) que una predicción particular se ha de sobrescribir.

65

La señal de vídeo puede comprender información general así como información dinámica. Todos los tipos de información anteriores, así como cualquier otro tipo de información relacionada con el uso del método de acuerdo con la invención se denominan dentro del marco de aplicación un parámetro. Los parámetros por lo tanto pueden ser parámetros sencillos sí/no (punto A1 anterior), parámetros que indican una elección dentro de un conjunto de posibilidades (punto A2 anterior por ejemplo), o un parámetro para control de una etapa dentro del método (punto A3 anterior) o parámetros dinámicos (punto B anterior) en cualquier forma o tipo.

La Figura 8 especifica de nuevo el aparato más genérico (que corresponde al método más genérico) de la invención. En una compresión de vídeo que permite que una unidad de predicción 801 aplique un criterio SEL\_F() tras varias predicciones generadas mediante unidades de predicción similares T, T2, que hacen por ejemplo una transformación y compensación de movimiento con pérdidas (por ejemplo, en AVC, la predicción temporal puede resultar de una combinación de varias imágenes previas), uno de estos modos "convencionales" puede "re-usarse" almacenando datos desde una predicción separada desde otra unidad de predicción PR\_UN (es decir en la submemoria MEM\_P2, se sobreescribe una primera predicción cuantificada de manera basta, por ejemplo una aproximación de ondícula, mediante otra región de predicción, mediante otro algoritmo desde PR\_UN). Este nuevo modo de predicción puede introducirse fácilmente a continuación en una norma existente (cualquiera que se quiera). El criterio de selección SEL\_F() puede cambiarse también.

La otra unidad de predicción puede usar por ejemplo una síntesis de textura. Si el lado del codificador reconoce que hay, por ejemplo, porción (por ejemplo bloques) de hierba, puede usarse un criterio de comparación o de predicción que solamente use hierba generada sintéticamente (por ejemplo típicamente sin actualización residual). Esto puede seleccionarse a continuación como la mejor calidad/región comprimida en comparación con por ejemplo la predicción temporal. El decodificador aplicará a continuación también la generación de hierba (incluso si no necesitara recibir información predictiva temporal para esa región). Todos los tipos de predicciones mejoradas pueden generarse mediante la unidad de predicción PR\_UN, por ejemplo aplicando alguna transformación de imagen tal como acentuación de contornos o suavizado en algún dato de imagen (o dos algoritmos de desintercalación diferentes en los dos últimos, memorias de predicción temporal menos precisas, que forman entonces predicciones competitivas con, por ejemplo, una predicción temporal), o un modelo en 3D (por ejemplo generado a partir de un par de imágenes secuenciales en tiempo), etc.

Tales datos de sincronización pueden transmitirse también mediante la señal de imagen comprimida normalizada, por ejemplo, en las SEI.

La invención se realiza también en cualquier producto de programa informático para un método o dispositivo de acuerdo con la invención. Bajo producto de programa informático debería entenderse cualquier realización física de una colección de comandos que posibilitan a un procesador genérico o de fin especial, después de una serie de etapas de carga (que puede incluir etapas de conversión intermedias, como traducción a un lenguaje intermedio y un lenguaje de procesador final) obtener los comandos en el procesador, para ejecutar cualquiera de las funciones características de una invención. En particular, el producto de programa informático puede realizarse como datos en un soporte tal como por ejemplo un disco o cinta, datos presentes en una memoria, datos de viajan a través de una conexión de red -cableada o inalámbrica- , o código de programa en papel. Además del código de programa, los datos característicos requeridos para el programa pueden realizarse también como un producto de programa informático.

Debería indicarse que las realizaciones anteriormente mencionadas ilustran en lugar de limitar la invención, y que los expertos en la materia podrán diseñar muchas realizaciones alternativas sin alejarse del alcance de las reivindicaciones adjuntas.

En las reivindicaciones, cualquier signo de referencia colocado entre paréntesis no deberá interpretarse como que limita la reivindicación.

Será evidente que dentro del marco de la invención son posibles muchas variaciones. Se apreciará por los expertos en la materia que la presente invención no está limitada por lo que se ha mostrado y descrito particularmente anteriormente en el presente documento. La invención reside en cada uno y todo rasgo característico novedoso y cada una y toda combinación de rasgos característicos. Los números de referencia en las reivindicaciones no limitan su alcance de protección.

Por ejemplo, el método puede usarse para únicamente una parte de la imagen, o diferentes realizaciones del método de la invención pueden usarse para diferentes partes de la imagen, por ejemplo usar una realización para el centro de la imagen, mientras se usa otra para los bordes de la imagen.

El uso del verbo "comprender" y sus conjugaciones no excluye la presencia de elementos distintos a aquellos establecidos en las reivindicaciones. El uso del artículo "un" o "una" precediendo a un elemento no excluye la presencia de una pluralidad de tales elementos.



La invención se describe con respecto a fotogramas. Dentro del marco de la invención "fotograma" podría ser una parte de un fotograma si por ejemplo el método de acuerdo con la invención se hace en una parte del fotograma, por ejemplo únicamente en la parte media del vídeo o únicamente las partes de primer plano de un vídeo.

## REIVINDICACIONES

1. Método para codificar una señal de vídeo en el que se realiza compresión de imagen en el que para la señal de vídeo se codifica un flujo de mejora en un codificador de flujo de mejora (3) y se codifica un flujo de base en un codificador de flujo de base (2), y las velocidades de fotogramas para el flujo de mejora y el flujo de base se diferencian por un factor diferente de una potencia de 2, y se usan múltiples predicciones temporales en el codificador de flujo de mejora (3) en el que se almacenan múltiples fotogramas de predicción en una memoria (5) del codificador de flujo de mejora (3) en el que un fotograma de predicción (Rn-1) en dicha memoria (5) del codificador de flujo de mejora (3) se sobrescribe con un fotograma (4) obtenido desde el codificador de flujo de base (2), en el que el fotograma de predicción sobrescrito en dicha memoria (5) es el fotograma de predicción más prolongado en memoria.
2. Método de acuerdo con la reivindicación 1, en el que el fotograma obtenido desde el codificador de flujo de base es un fotograma sobreescalado.
3. Método para decodificar una señal de vídeo que comprende un flujo de mejora y un flujo de base en el que el flujo de mejora se decodifica en un decodificador de flujo de mejora (3') y el flujo de base se decodifica en un decodificador de flujo de base (2'), y las velocidades de fotogramas para el flujo de mejora y el flujo de base se diferencian por un factor diferente de una potencia de 2, y se usan múltiples predicciones temporales en el decodificador de flujo de mejora (3') en el que se almacenan múltiples fotogramas de predicción en una memoria (5') del decodificador de flujo de mejora (3') en el que un fotograma de predicción (Rn-1) en la memoria (5') del decodificador de flujo de mejora (3') se sobrescribe con un fotograma (4') obtenido desde el decodificador de flujo de base (2') en el que el fotograma de predicción sobrescrito en dicha memoria (5') es el fotograma de predicción más prolongado en la memoria.
4. Método para decodificar de acuerdo con la reivindicación 3, en el que el fotograma obtenido desde el decodificador de flujo de base es un fotograma sobreescalado.
5. Codificador para codificar una señal de vídeo en un flujo de mejora y un flujo de base que comprende un codificador de flujo de mejora (3) y un codificador de flujo de base (2), en el que se realiza compresión de imagen, y las velocidades de fotogramas para el flujo de mejora y el flujo de base se diferencian por un factor diferente de una potencia de 2, en el que el codificador de flujo de mejora (3) está dispuesto para generar múltiples predicciones temporales y almacenar las múltiples predicciones temporales en una memoria (5) y en el que el codificador está dispuesto para sobrescribir un fotograma de predicción (Rn-1) en dicha memoria (5) del codificador de flujo de mejora (3) con un fotograma obtenido desde el codificador de flujo de base (2) en el que el fotograma de predicción sobrescrito en dicha memoria (5) es el fotograma de predicción más prolongado en memoria.
6. Codificador de acuerdo con la reivindicación 5, en el que el codificador está dispuesto de manera que el fotograma obtenido desde el codificador de flujo de base sea un fotograma sobreescalado.
7. Decodificador para una señal de vídeo comprimido que comprende un flujo de base y un flujo de mejora, comprendiendo el decodificador un decodificador de flujo de base (2') y un decodificador de flujo de mejora (3'), y las velocidades de fotogramas para el flujo de mejora y el flujo de base se diferencian por un factor diferente de una potencia de 2, en el que el decodificador de flujo de mejora (3') está dispuesto para generar múltiples predicciones temporales y almacenar las múltiples predicciones temporales en una memoria (5') en el que el decodificador está dispuesto para sobrescribir un fotograma de predicción (Rn-1) en la memoria (5') del decodificador de flujo de mejora (3') con un fotograma (4') obtenido desde el decodificador de flujo de base (2') en el que el fotograma de predicción sobrescrito en dicha memoria (5') es el fotograma de predicción más prolongado en memoria.
8. Decodificador de acuerdo con la reivindicación 7, en el que el decodificador está dispuesto de manera que el fotograma obtenido desde el decodificador de flujo de base es un fotograma sobreescalado.
9. Producto de programa informático que comprende medios de código de programa almacenados en un medio legible por ordenador para realizar un método de acuerdo con una cualquiera de las reivindicaciones 1 a 4 cuando dicho programa se ejecuta en un ordenador.

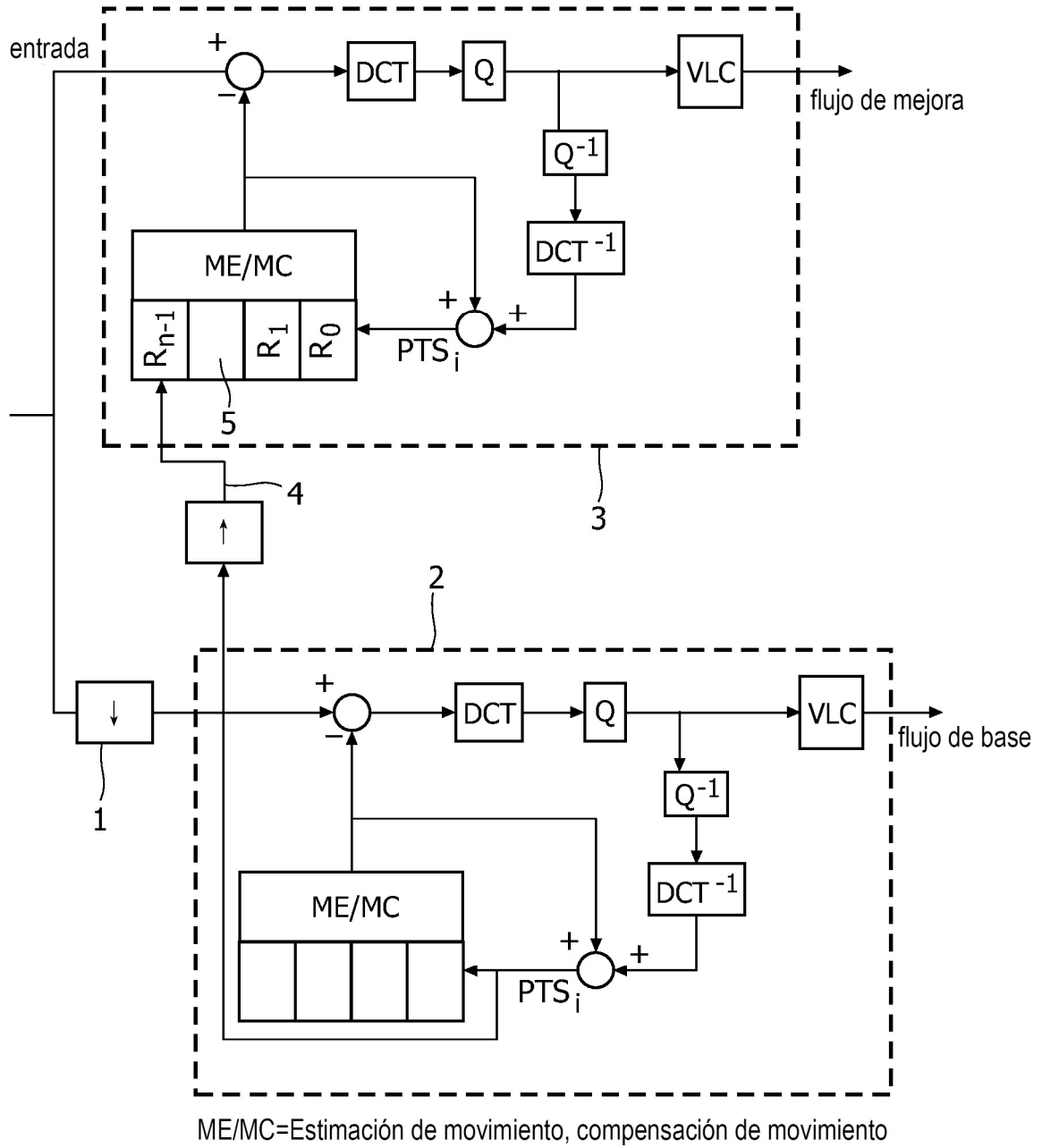


FIG. 1a

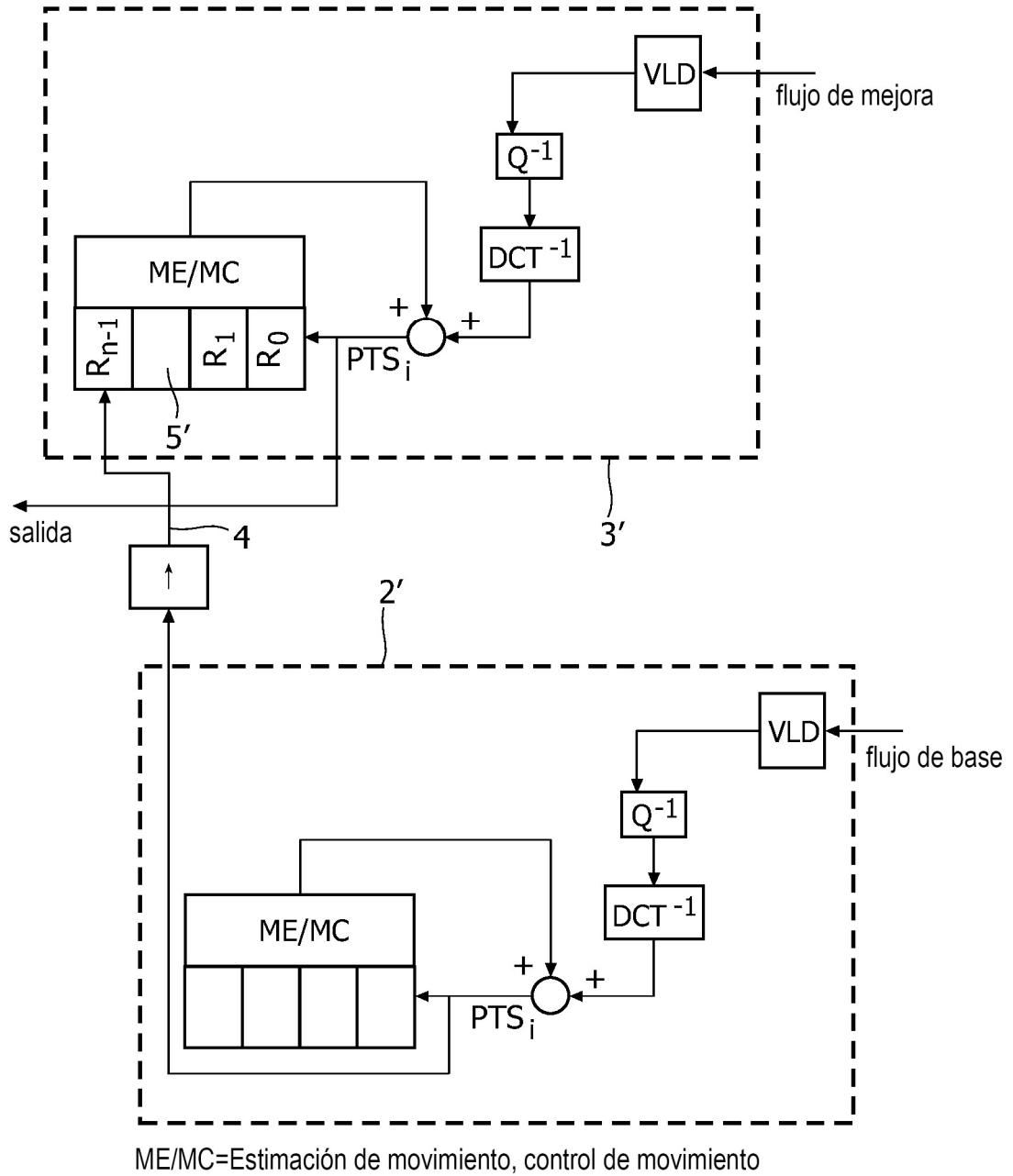


FIG. 1b

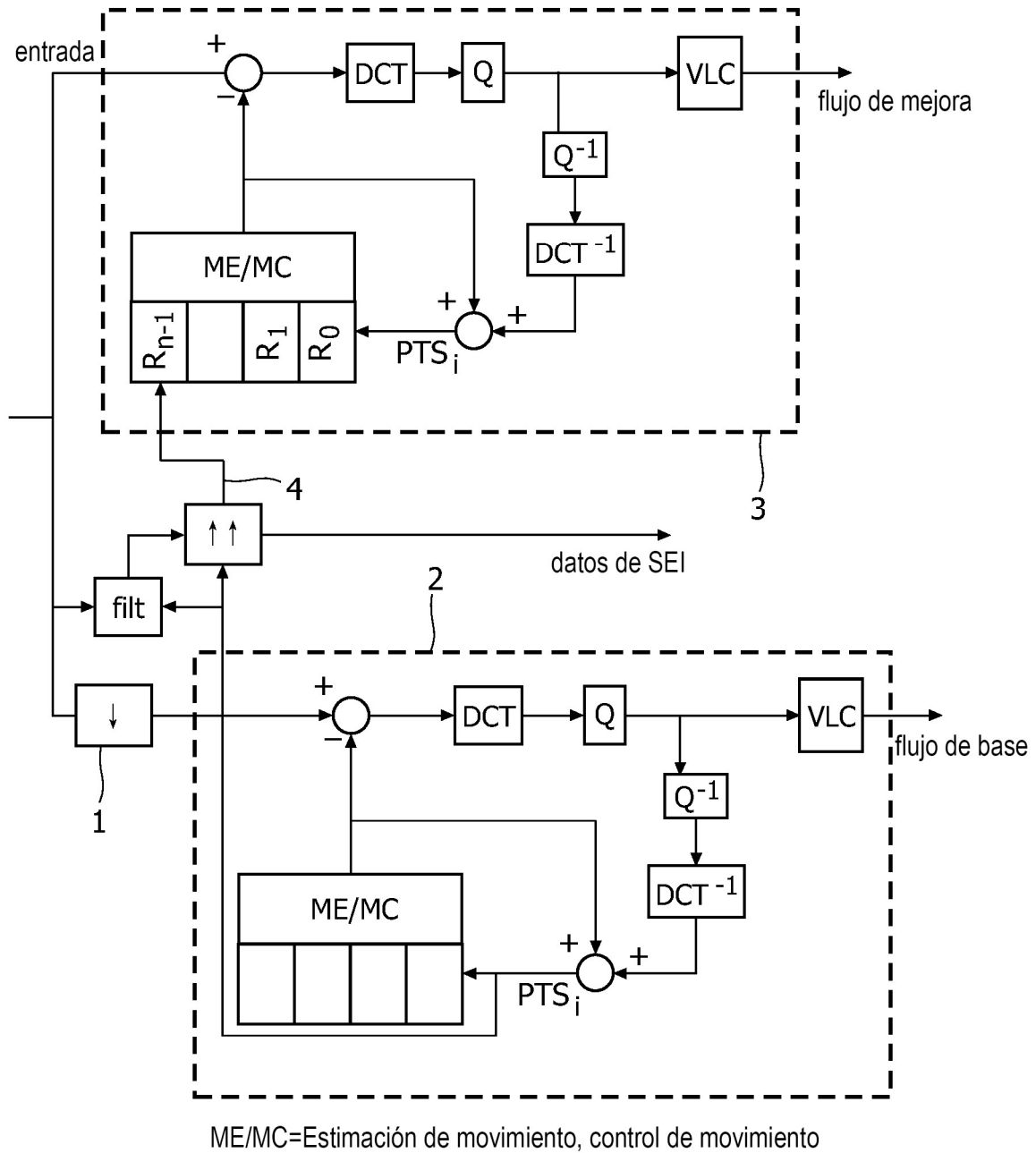


FIG. 1c

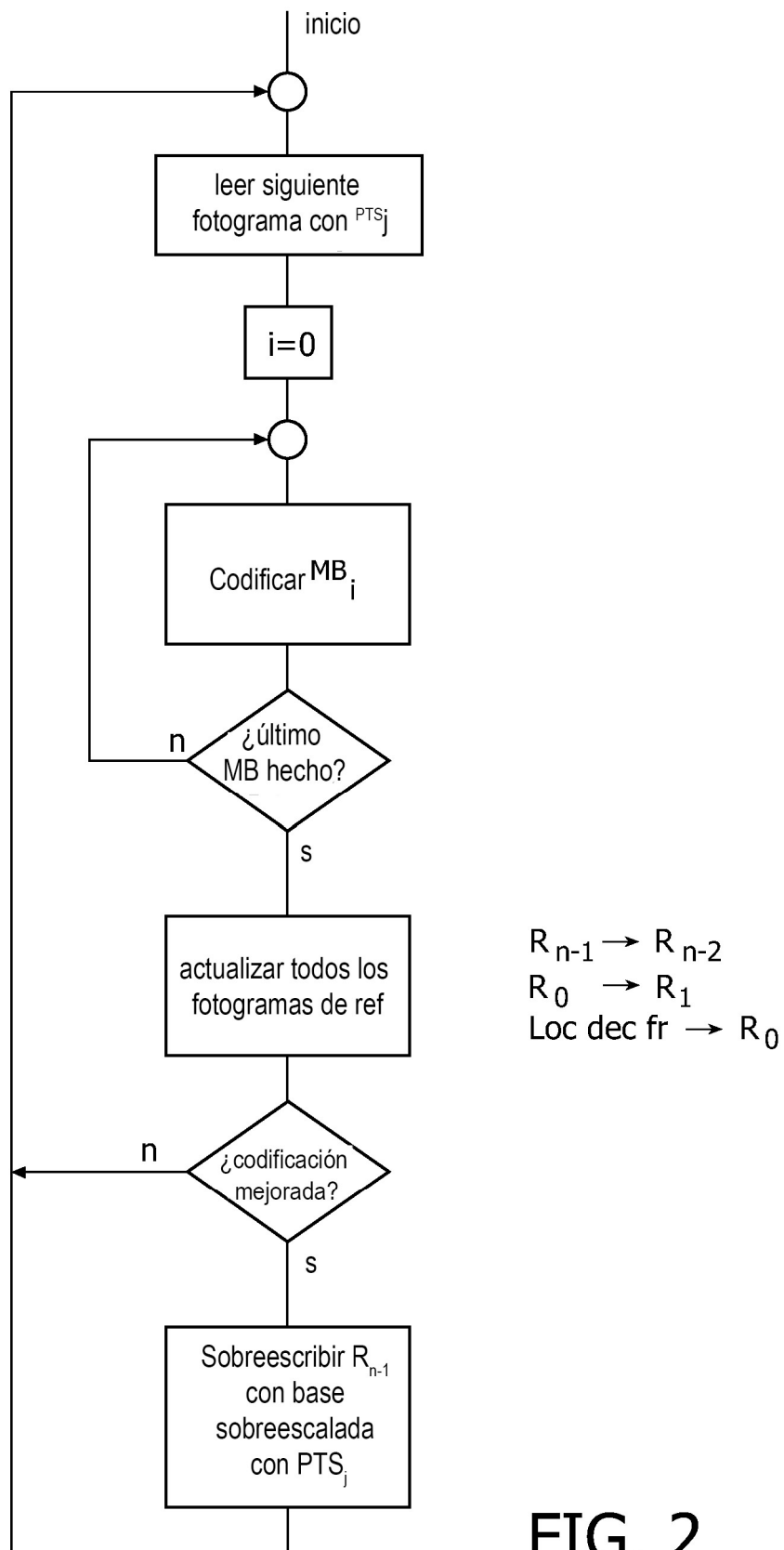


FIG. 2

$\underline{I}_0$   $B_1$   $\underline{P}_2$   $B_3$   $\underline{P}_4$   $B_5$   $\underline{P}_6$  (orden de visualización)

n.º ref = 3	1	2	3	4	5	6	7	
	$\underline{I}_0$	$\underline{P}_2$	$B_1$	$\underline{P}_4$	$B_3$	$\underline{P}_6$	$B_5$	(orden de codificación)
mem 0	$\underline{I}_0$	$\underline{P}_2$	$\underline{P}_2$	$\underline{P}_4$	$\underline{P}_4$	$\underline{P}_6$	$\underline{P}_6$	
mem 1		$\underline{I}_0$	$\underline{I}_0$	$\underline{P}_2$	$\underline{P}_2$	$\underline{P}_4$	$\underline{P}_4$	
mem 2			$B_1$	$\underline{I}_0$	$\underline{I}_0$	$\underline{P}_2$	$\underline{P}_2$	
mem 3				$B_1$	$B_3$	$B_3$	$B_5$	
Visualizado	-	-	$\underline{I}_0$	$\underline{B}_1$	$\underline{P}_2$	$B_3$	$\underline{P}_4$	

$\underline{I}_0$   $\underline{P}_2$   $\underline{P}_4$   $\underline{P}_6$  son fotogramas referenciados

$B_1$   $B_3$   $B_5$  son fotogramas no referenciados

FIG. 2b

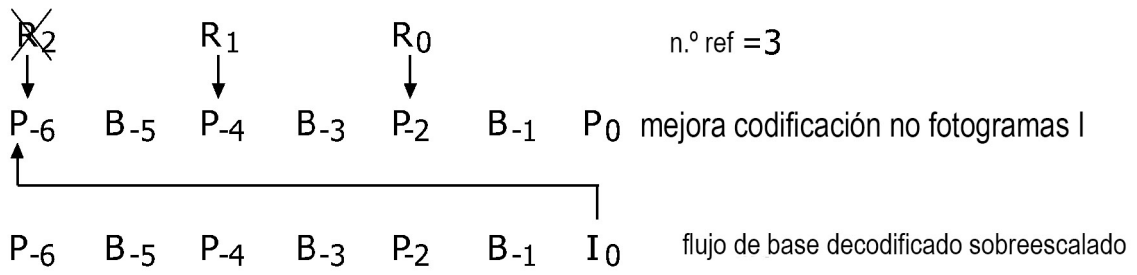


FIG. 3

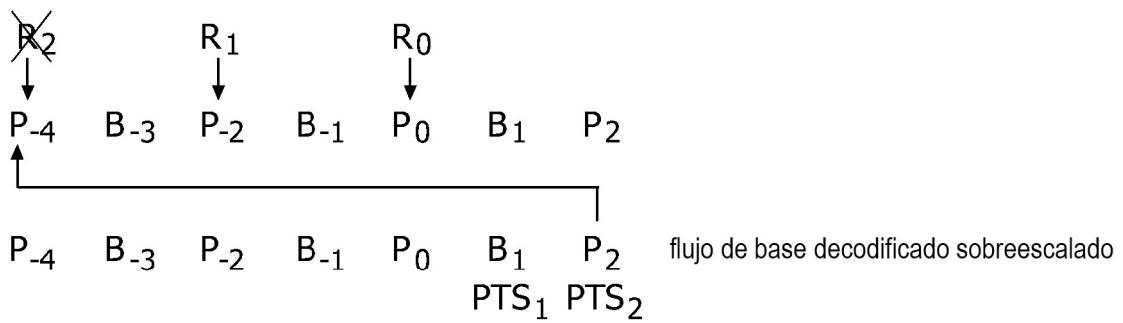


FIG. 4

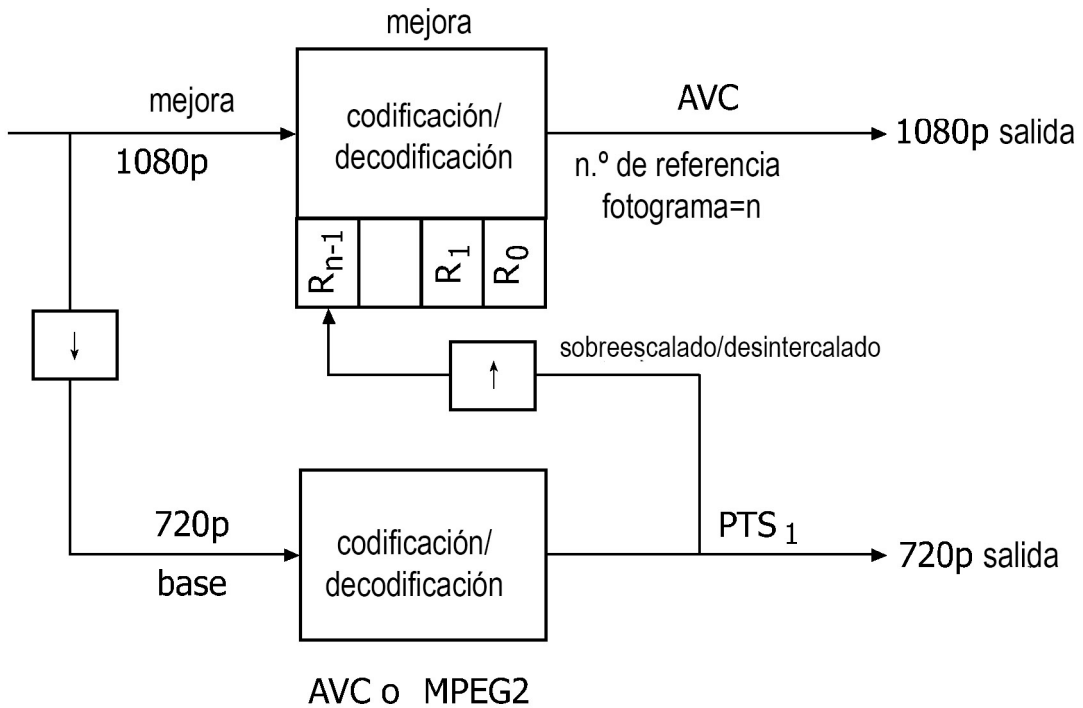


FIG. 5



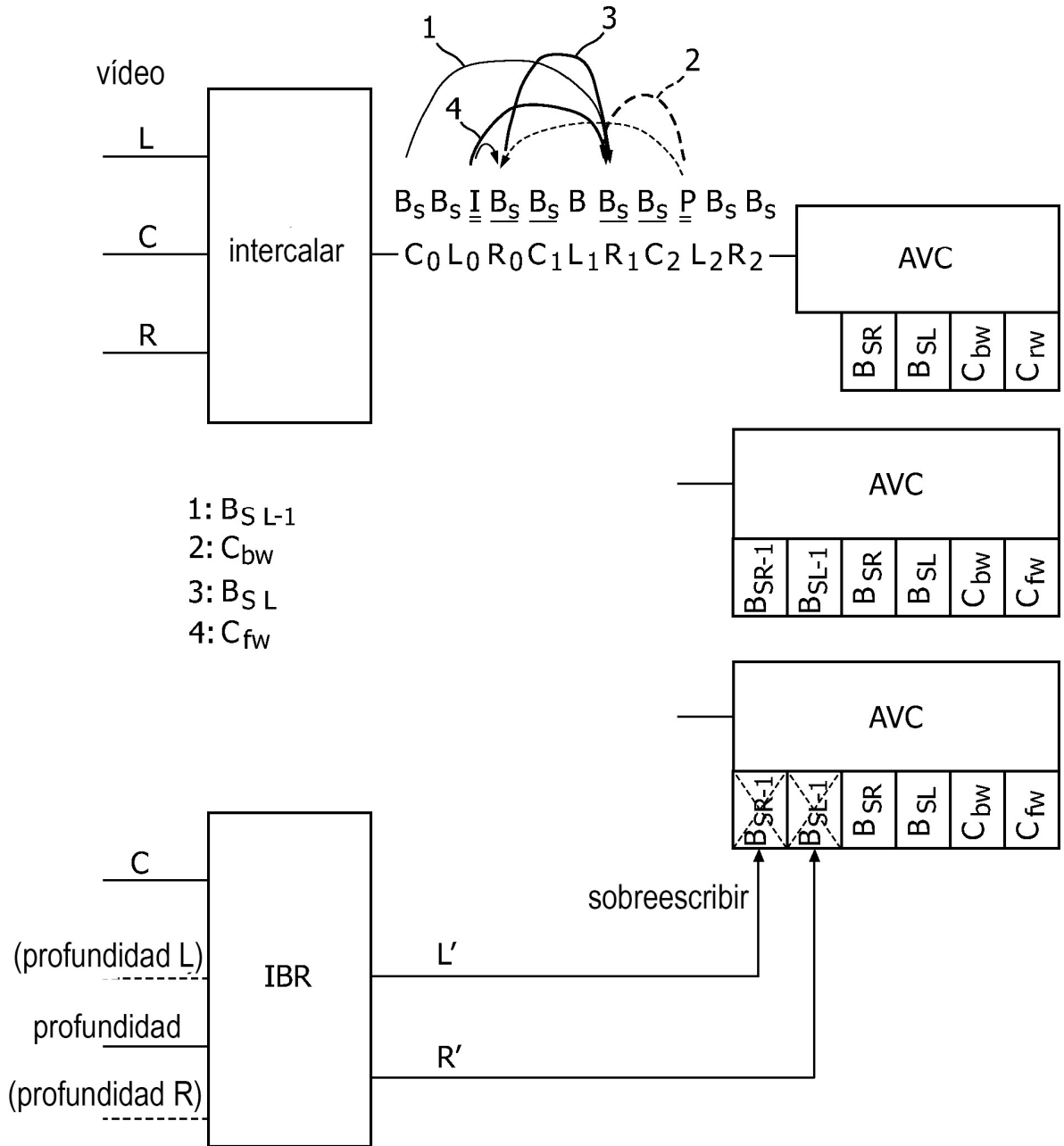


FIG. 6

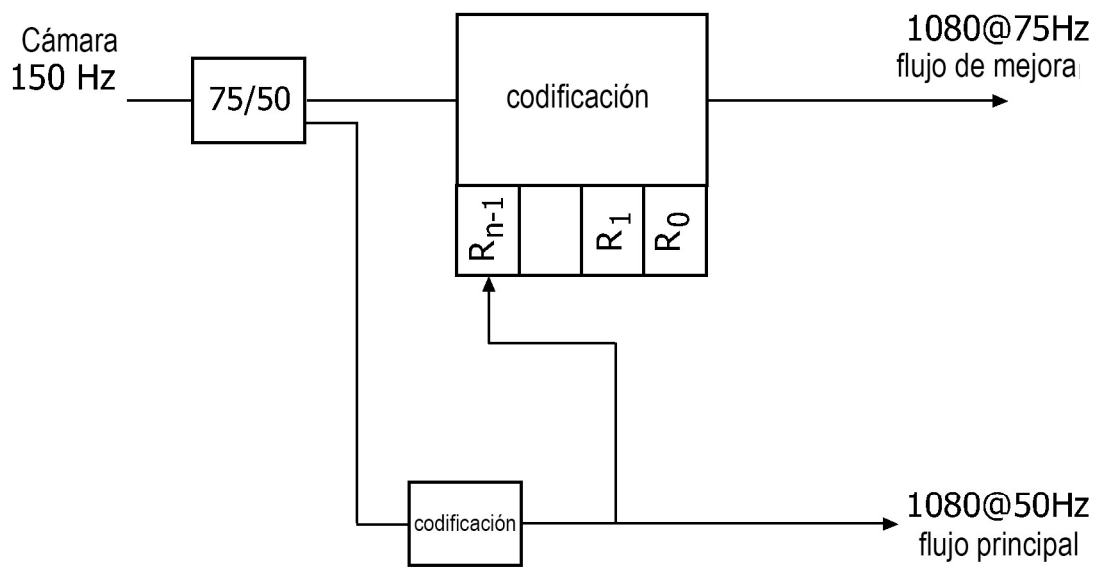
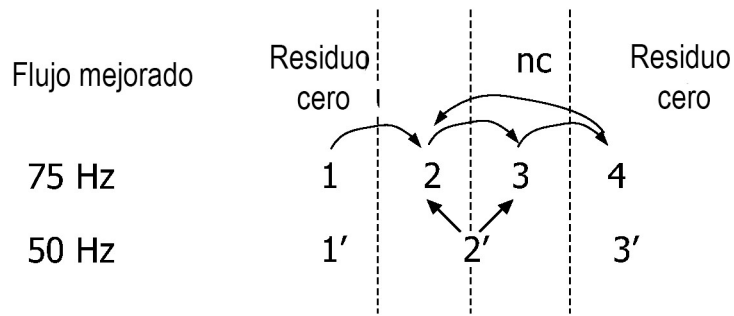


FIG. 7

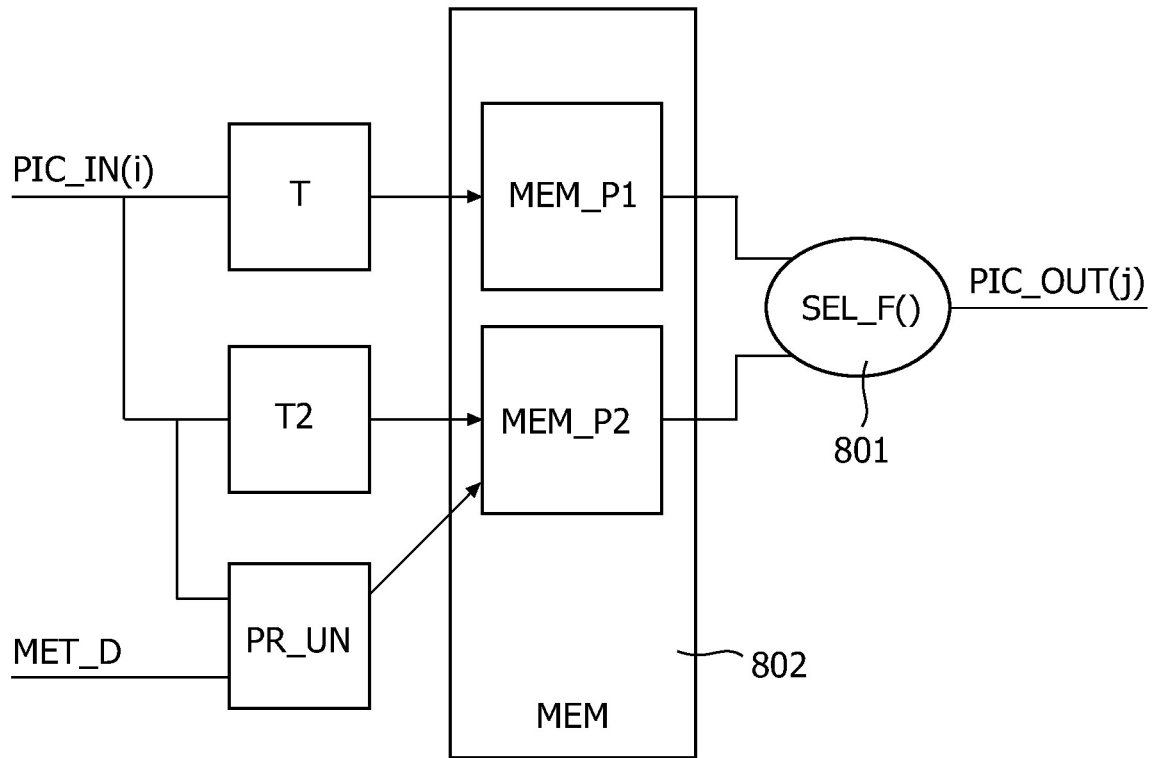


FIG. 8