

19



OFICINA ESPAÑOLA DE
PATENTES Y MARCAS

ESPAÑA



11 Número de publicación: **2 674 897**

51 Int. Cl.:

H04N 5/45

(2011.01)

12

TRADUCCIÓN DE PATENTE EUROPEA

T3

96 Fecha de presentación y número de la solicitud europea: **05.07.2006** E 17181071 (6)

97 Fecha y número de publicación de la concesión europea: **16.05.2018** EP 3247110

54 Título: **Método y dispositivo para manejar múltiples flujos de vídeo usando metadatos**

30 Prioridad:

18.07.2005 EP 05015576

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

04.07.2018

73 Titular/es:

**THOMSON LICENSING (100.0%)
1-5 Rue Jeanne d'Arc
92130 Issy-les-Moulineaux, FR**

72 Inventor/es:

**WINTER, MARCO;
GANDOLPH, DIRK;
PETERS, HARTMUT;
OSTERMANN, RALF;
HÖRENTROP, JOBST;
SCHEWZOW, ANDREJ y
JANSSEN, UWE**

74 Agente/Representante:

ELZABURU, S.L.P

ES 2 674 897 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 dispositivo para manejar múltiples flujos de vídeo usando metadatos

Campo de la invención

5 Esta invención describe un método y un dispositivo para manejar múltiples flujos de vídeo. En particular, el método y el dispositivo son para sobreponer dinámicamente dos o más flujos de vídeo usando metadatos especiales.

Antecedentes de la invención

Los dispositivos de reproducción para datos de audio/vídeo (AV), usualmente basados en discos duros (HDD) o discos ópticos, pueden tener diferentes posibilidades para acceder a datos adicionales, p. ej. acceso a internet, Bus Serie Universal (USB) interfaz, etc. para descargar contenidos adicionales.

10 Algunos equipos de TV finales proporcionan la posibilidad de ver dos o más programas simultáneamente usando la técnica de imagen sobre imagen (PIP), que sin embargo usa diferentes soluciones técnicas en propiedad. La manera con la que dicho aparato de TV presenta diferentes secuencias de vídeo en la misma pantalla no es flexible, sino inmediatamente fija tras fabricar el aparato de TV.

Las técnicas PIP disponibles funcionan sin ayuda de información extraída del contenido AV.

15 El documento US 6.046.778 describe cómo sobreponer una imagen de subtítulo sobre imágenes en movimiento de un flujo de vídeo. La posición en pantalla de la imagen de subtítulo es determinada según un atributo de pantalla. El atributo de pantalla se almacena en un almacenamiento de atributos, así el atributo de pantalla no se asocia con el flujo de vídeo.

20 El documento US 2002/075407 A1 describe que una imagen de pantalla secundaria se sobrepone sobre una imagen de pantalla primaria. La posición para sobreponer la imagen de pantalla secundaria sobre la imagen de pantalla primaria subyacente se determina analizando un primer flujo de datos de vídeo que transmite la imagen de pantalla primaria para una señal o señales. Así, la señal o señales se extraen de los datos de imagen del primer flujo de datos de vídeo para la imagen de pantalla primaria subyacente, no con un flujo de datos de vídeo para la imagen secundaria suprayacente.

25 Compendio de la invención

La tecnología disponible descrita para sobreposición de vídeo no es flexible. Por lo tanto la presente invención aborda el problema de proporcionar un método y un dispositivo para sobreponer una secuencia de vídeo a otra secuencia de vídeo (o una imagen de vídeo estática) de una manera más flexible, y en particular proporciona posibilidades de control para dicha sobreposición de vídeo para alguien interesado en una apariencia visual particular del programa, p. ej. el proveedor de programa, el autor de contenidos AV pregrabados, el usuario.

30 La presente invención proporciona una manera para cambiar la decisión de cómo sobreponer diferentes secuencias de vídeo en el momento de presentación. Por lo tanto la invención cambia la sobreposición estática de vídeo para un sistema de procesamiento flexible de sobreposición dinámica de vídeo. Con este cambio, p. ej. el proveedor de contenido llega a poder controlar la sobreposición. Esta es una etapa principal para combinar dinámicamente secuencias de vídeo.

35 El quid de la invención es añadir metadatos a una secuencia de vídeo, dichos metadatos contienen parámetros de sobreposición y se pueden usar para controlar una sobreposición de esta secuencia de vídeo a otra secuencia de vídeo o a una imagen estática. En principio es suficiente si una secuencia de vídeo que se sobrepone contiene dichos metadatos. En este caso, la sobreposición puede ser independiente de la otra secuencia de vídeo o imagen a la que se sobrepone. Sin embargo, dentro de la secuencia de vídeo es posible identificar otra cierta secuencia de vídeo o grupo de secuencias de vídeo, con las que están relacionados los parámetros de sobreposición. Además, es posible insertar metadatos de sobreposición a una secuencia de vídeo que definen cómo puede ser sobrepuesta otra secuencia de vídeo, p. ej., para una escena particular puede estar prohibida para que se sobrepongan ciertas áreas dentro de la pantalla.

40 Para el caso de colisiones, se pueden definir prioridades para uno, varios o todos los parámetros. P. ej. una primera secuencia de vídeo puede contener metadatos que definen una escena particular, p. ej. en términos de marcas de tiempo, y que además definen para esta escena un área rectangular de p. ej. 20-80 % de la altura de pantalla y 0-100 % de la anchura de pantalla que no se puede sobreponer. Una segunda secuencia de vídeo puede contener metadatos que definen que, si esta secuencia se usa como sobreposición, se debe producir en un fotograma visible, p. ej., en el 0-10 % de la anchura de pantalla y preferiblemente en el 50 % de la altura de pantalla, es decir, altura central en el canto izquierdo. En este caso, es posible desplazar el fotograma arriba o abajo, de modo que sea expuesto fuera del área prohibida, porque el parámetro "50 % de altura" tiene menor prioridad que los otros. Si el área prohibida en la primera secuencia también tiene menor prioridad, es decir, no "debe" ser sobrepuesta, la segunda secuencia se puede posicionar de manera que se cumplan ambas condiciones. Si la segunda secuencia de

vídeo se expone sola, es decir, no como una sobreposición, se ignoran estos metadatos.

Una ventaja de la invención es que puede proporcionar diversas características útiles, p. ej., una o más secuencias de vídeo se pueden mover, escalar, combinar en alfa y/o superponer sobre una secuencia de vídeo de fondo. Combinar en alfa es una técnica común para transiciones suaves entre escenas. Estas características no necesitan estar estadísticamente disponibles, es decir, pueden ser descargadas dinámicamente de internet o de una memoria rápida o de discos adicionales (p. ej. actualizaciones de películas grabadas) etc.

Según la invención, se proporcionan metadatos especiales con la secuencia de vídeo, que describen cómo "pegar" entre sí diferentes (pero determinadas) secuencias de vídeo. El formato de metadatos debe ser estandarizado, de modo que diferentes dispositivos de reproducción puedan manejar/interpretar este tipo de metadatos. Actualmente, no hay disponibles dichos metadatos.

Un dispositivo de reproducción según la invención puede interpretar metadatos apropiados, y tiene el hardware/software (p. ej. acelerador gráfico 3D) para realizar procesamiento acorde para la superposición, etc. Para metadatos procedentes de internet, el dispositivo puede tener la capacidad de descargar dichos datos. Para metadatos procedentes de un lector de tarjetas rápidas o USB, el dispositivo puede tener conectores apropiados. Además, el dispositivo tiene la capacidad de almacenar dichos metadatos en un área de memoria accesible, p. ej. DRAM local.

Existen diversas posibilidades para mezclar las secuencias de vídeo y los metadatos: Los metadatos puede ser entrelazados con el flujo AV, los metadatos pueden ser almacenados en un archivo separado o los metadatos pueden ser entregados en paralelo junto con el flujo de vídeo (p. ej. la cabeza hablando del director de la película) para una película especial. En el último caso, la entrega contiene todos datos adicionales necesarios para una reproducción superpuesta. Además, puede haber diferentes maneras de superponer dos o más flujos de vídeo particulares. Así, puede haber disponible más de un conjunto de metadatos para las mismas secuencias de vídeo.

Los metadatos usados por la invención describen cómo superponer dos o más secuencias de vídeo en la misma pantalla de vídeo. Esta superposición puede ser posible en un gran intervalo, es decir, muy flexible. Esto significa que los metadatos contienen, p. ej., información acerca de apariencia geométrica de la pantalla para las secuencias superpuestas de vídeo, la temporización se superposición, preferiblemente para los flujos de vídeo relativamente entre sí, lo que es particularmente útil para sincronizar flujos de audio, metadatos de efectos especiales, p. ej. fundido de entrada/salida, cambios de color, fotogramas visibles para PIP, etc., y metadatos de flujo.

Descripción detallada de la invención

La figura 1 muestra un dispositivo para superposición de fotogramas según la invención, para ilustrar cómo y qué tipo de metadatos se usa para qué etapa de decodificación para realizar la superposición dinámica de vídeo. Las flechas discontinuas indican qué tipo de metadatos se usan para realizar una etapa de decodificación específica. Las flechas continuas indican la corriente de los datos de secuencia de vídeo. El formato de los datos de flujo de vídeo puede cambiar de una etapa a otra. P. ej. inicialmente los datos de secuencia de vídeo están en un formato comprimido (p. ej. MPEG-2, o MPEG-4 o VC-1, etc.), y en la siguiente etapa han sido decodificados hasta búferes de fotogramas (p. ej. formato tipo mapa de bits). El vídeo superpuesto final podría estar en formato YUV 4:2:0.

Además, la figura 1 indica que se da una jerarquía dentro de los propios metadatos. Los metadatos de temporización de superposición controlan los metadatos de efectos especiales y geométricos. Esto es importante, porque p. ej. la posición de una secuencia de subvídeo en una pantalla puede cambiar durante la reproducción, pero no la temporización. P. ej. la secuencia de subvídeo puede moverse suavemente en horizontal, o los fotogramas de vídeo global se pueden escalar en aumento y/o en disminución. Ejemplarmente, la pantalla expone varios vídeos en miniatura de diferentes flujos, con un vídeo en miniatura por flujo, y el usuario selecciona una de estas miniaturas usando un menú. Como efecto especial, la resolución de este vídeo en miniatura se escala en aumento hasta que cubre la pantalla completa. Eso proporcionaría la impresión de que la selección está relacionada con la reproducción inmediatamente siguiente del vídeo a tamaño de pantalla completa. Esto se representa en las figuras 4-6.

Un nivel adicional incluso superior de la jerarquía puede controlar los metadatos de temporización. Esta puede ser la interacción de usuario. El usuario puede cambiar entre diferentes programas de reproducción, como diferentes cadenas de programas (PGC) como usan los DVD. Es como una lista de reproducción, es decir, el usuario puede cambiar a otro programa de superposición.

Debido a la jerarquía de metadatos, los niveles superiores pueden evitar configuraciones de los niveles inferiores. P. ej. se puede recortar un subvídeo y se puede cambiar la relación de aspecto de píxeles. Así, los metadatos de nivel inferior serán manejados como datos predeterminados.

Ventajosamente, si dichos metadatos se almacenan como archivo separado, entonces se deben almacenar usando formatos de documentos abiertos como XML. Si los metadatos se incrustan en el propio flujo de datos, entonces el formato usado para los metadatos incrustados debe seguir las restricciones especiales del formato de flujo de vídeo, p. ej. evitar códigos de inicio en campos de user_data() dentro de flujos de vídeo de MPEG-2 o MPEG-4.

Las figuras 2 y 3 muestran cómo un usuario puede pedir por separado un disco de película, material de bonus extra (p. ej. por medio de internet) y metadatos relacionados. Como ejemplo, un usuario compra una película como disco óptico, abre la cubierta y ve un cupón para una secuencia de vídeo de comentarios del director. El usuario debe conectarse a la página web del proveedor de contenido (p. ej. Disney) y registrarse. Tras registrarse en la página web del proveedor de contenido, el usuario obtiene acceso a la secuencia adicional de vídeo de comentarios del director. El usuario descarga la secuencia de vídeo por medio de su conexión DSL. Después de hacer la descarga, el usuario desea echar un vistazo a su incentivo descargado más reciente. El usuario pone el incentivo descargado en un HDD del dispositivo de reproducción. Entonces, el usuario inicia la reproducción de la nueva secuencia de vídeo descargada. El dispositivo de reproducción reconoce los metadatos insertados. Así, el dispositivo de reproducción sabe que la reproducción de esta secuencia de vídeo está pensada conjuntamente con la película especial. Si la película todavía no ha sido insertada, entonces el dispositivo de reproducción informa al usuario (p. ej. por medio de IU) que el usuario debe insertar el disco apropiado. Cuando se inserta el disco, entonces el dispositivo de reproducción inicia la reproducción de ambas secuencias de vídeo, como definen los metadatos entregados junto con el contenido de incentivo (p. ej. incorporados en el mismo flujo). En la pantalla se muestra la película, con la cara del director superpuesta. El director habla acerca de la escena mostrada en ese momento. La secuencia de vídeo de incentivo contiene una lista de reproducción de metadatos, es decir, un conjunto de partes comentadas. Así, el usuario puede seleccionar una de varias escenas comentadas, p. ej. una lista de reproducción como las siguientes

1. El director comenta la película entera
2. El director comenta las escenas peligrosas
3. El director comenta las escenas de efectos especiales
4. El actor 1 comenta la escena erótica
5. El actor 2 comenta algunas situaciones divertidas durante la grabación

En este ejemplo, hay disponibles dos películas (película original y secuencia de vídeo de incentivo), y cinco conjuntos de metadatos. Cada conjunto combina diferentes partes del vídeo principal y el vídeo de incentivo, y los conjuntos controlan dónde se superponen los vídeos (información geométrica), cuándo se superponen (sincronización), qué efectos se usan (p. ej. combinación alfa), y cómo mezclar los audio de ambos flujos. Incluso se puede añadir información para interacción de usuario puede, p. ej. se pueden ofrecer diferentes maneras de continuar (p. ej. "finalizar reproducción" y "reproducir de nuevo" y "ir al siguiente conjunto de metadatos").

En este ejemplo, la película principal se acaba tras producir los másteres para el disco óptico. Pero, el proveedor de contenido puede añadir posteriormente información adicional para esta película usando los metadatos según la invención, combinándolos con secuencias de vídeo apropiadas y haciendo que estén disponibles en su portal de internet. Ventajosamente, el proveedor de contenido puede añadir información interesante acerca de la película por medio de esta característica en cualquier momento. Los clientes pueden ser informados, p. ej., por medio de un boletín informativo o un correo electrónico directo. Así, se puede lograr una retención de clientes más fuerte.

Los creadores de películas usan grabación cada vez más ayudada por ordenador, es decir, la cantidad de escenas artificiales generadas en las películas actuales crece fuertemente. Así, continuamente se hace más fácil generar diferentes secuencias de vídeo. Como otro ejemplo, un usuario compra un disco óptico con una película erótica. Sin embargo, la mayoría de personas mostradas en la película lleva ropa. Tras pagar una tarifa, el usuario puede descargar del portal de internet del proveedor de contenido información especial de secuencia de vídeo, que proporciona la vista deseada. Esta información especial de secuencia de vídeo (SVSI) contiene una secuencia de vídeo diferente, que cambia la apariencia de la película erótica, modificada píxel a píxel. Para la reproducción, el dispositivo de reproducción superpone ambas secuencias de vídeo según los metadatos inventivos también entregados. La secuencia de vídeo diferente añadida cambia la apariencia de algunos píxeles. En este ejemplo, retira prendas apropiadas de algunos actores en algunas escenas. Así, el cliente obtiene una visión libre de piel más desnuda al pagar una tarifa adicional al proveedor de contenido y descargar la secuencia de vídeo diferente de internet. La cantidad de piel visible puede depender de la cantidad de tarifas pagadas. En un ejemplo similar, la ropa que llevan los actores en una película puede cambiarse según la forma actual aplicando esta característica, o se puede cambiar el fondo (es decir, la ubicación virtual) de una escena de película.

En un ejemplo adicional, un usuario compra un disco óptico con una película. Varios actores representan papeles en esta película. El proveedor de contenido separa la película en las escenas de fondo y las cabezas de los actores (como secuencia de vídeo). La superposición de la escena de fondo y las cabezas de los actores realizará la reproducción real de la película. Los metadatos inventivos describen cuando qué cabezas se superponen con el vídeo de fondo en la posición derecha dentro de fotograma. Se usan para controlar el proceso de superposición. Además, los metadatos inventivos contienen información acerca de la cabeza expuesta, es decir, su posición tridimensional (3D) y posición relativa a la cámara, el nivel de iluminación y direcciones para la cabeza, el ambiente físico (p. ej. ventoso, bajo agua, fuerzas centrífugas), etc.

El cliente puede querer intercambiar los actores. Así, el usuario da al motor de reproducción información apropiada acerca de la cabeza tridimensional del actor deseado (p. ej. el propio cliente). Esto se puede hacer escaneando la cabeza, p. ej. algunas fotos electrónicas de la cabeza y software en el dispositivo de reproducción, que genera un modelo 3D apropiado de la cabeza deseada. Con algunos motores 3D típicos (como los usados en aceleradores gráficos actuales) el dispositivo de reproducción puede generar en tiempo real las cabezas apropiadas, que se superpondrán al vídeo de fondo. Motores 3D actuales pueden incluso manejar estructuras complejas tales como p. ej. pelo largo en combinación con viento. Para primeras versiones más simples, sería posible que el usuario escanee la cabeza deseada y envíe esta información al proveedor de contenido. El proveedor de contenido genera el vídeo de superposición y envía esta de nuevo al usuario. El usuario aplica esta secuencia recibida como se ha descrito anteriormente.

Sin embargo, el último ejemplo describe un método, en donde el proveedor de contenido produce en primer lugar el vídeo, separado en vídeo de fondo y cabezas. Adicionalmente, el proveedor de contenido proporciona metadatos que informan qué cabeza se expondrá en qué posición. El proveedor de contenido también puede proporcionar un modelo 3D del actor y metadatos que informan al dispositivo de reproducción cómo generar la proyección apropiada de esta cabeza.

El disco con este contenido se pone sobre el máster, es decir, es contenido estático como la propia película. Se pueden añadir caras de otros actores al disco (información estática). Entonces el usuario puede intercambiar los actores usando el contenido del disco adquirido. Además, el proveedor de contenido puede ofrecer otros actores para esta película como material de bonus descargando de su portal de internet. Además, el proveedor de contenido puede ofrecer poner modelos 3D en la película al recibir fotos apropiadas de un cliente, que son suficientes para generar el modelo 3D requerido. Esto sería otro modelo de negocio. Es incluso posible que los clientes puedan intercambiar modelos 3D con otros usuarios.

Así, el dispositivo de reproducción de este ejemplo soporta dos maneras de superposición:

- 1) superponer secuencias de vídeo ya proyectadas con el vídeo de fondo (p. ej. las cabezas de actores originales)
- 2) superponer proyecciones que se calculan en tiempo real durante la reproducción sobre la base de un modelo tridimensional de una cabeza. Los metadatos adicionales (p. ej. posición de la cabeza e información de iluminación, e información acerca del ambiente físico, etc.) son entregados por el proveedor de contenido (pueden estar ya en el disco o pueden ser entregados como material de bonus por medio de internet).

En un ejemplo adicional, el usuario genera un propio menú para obtener una visión general sobre vídeos disponibles. El usuario produce metadatos según la invención, que describen cuántos vídeos en miniatura se superponen con una película de fondo. El usuario también puede decidir el comportamiento de una miniatura seleccionada (para navegación) y la apariencia de la miniatura tras la activación. La figura 4 describe el menú desarrollado para este ejemplo. El usuario genera un menú de vídeo en miniatura 3x3 con un vídeo de fondo. La selección de una miniatura (p. ej. por medio de botones de navegación Arriba, Abajo, Izquierda, Derecha) escala en aumento un poco la secuencia de vídeo seleccionada como se muestra en la figura 5. Si el usuario activa esta miniatura (es decir, una selección final), entonces la miniatura crece hasta que cubre la pantalla entera. Entonces puede empezar la reproducción convencional.

En este ejemplo, el propio usuario produce los metadatos que definen dónde se ubicará un vídeo. En este caso, no es necesario que sea fijo el conocimiento acerca del vídeo presentado específico y la estructura de menú, es decir, durante la generación del menú puede no estar claro qué vídeos específicos se presentan en una posición especial en el menú, es decir, la superposición se hace interpretando en tiempo real la información de flujo del propio flujo de vídeo (para obtener el formato y la resolución, etc.) y la información geométrica e información de la estructura de menú del usuario. Así, estos metadatos son necesarios, pero usualmente están separados. El proveedor de contenido puede proporcionar software y/o plantillas para generar dicho menú. El dispositivo de reproducción puede soportar la generación de dichos menús.

Los metadatos según la invención se pueden ubicar de manera diferente.

Tales metadatos se pueden ubicar en el mismo medio (p. ej. HDD o disco óptico, p. ej. DVD, HD-DVD, o Blu-ray) que las secuencias de vídeo superpuestas. Generalmente, se pueden ubicar en el mismo medio que al menos una de las secuencias de vídeo superpuestas.

Los metadatos también se pueden ubicar un medio distinto que las secuencias de vídeo superpuestas.

Además, los metadatos pueden ser recibidos por separado de otras fuentes (p. ej. USB, memoria rápida, internet, etc.).

Existen diversas posibilidades para mezclar las secuencias de vídeo y los metadatos:

- Los metadatos pueden ser entrelazados con el flujo AV.
 - Esto significa que el proveedor de contenido decide durante la creación algo o todo sobre la superposición de los vídeos.
 - Técnicamente, este se puede hacer almacenando los metadatos p. ej. dentro de user_data() de Vídeo MPEG-2 o Vídeo MPEG-4 o flujos VC-1.
 - Otra alternativa sería almacenar los metadatos en paquetes de transporte separados del flujo de transporte (p. ej. en los archivos .m2ts del contenido BD).
 - Sin embargo, este tipo de metadatos necesita un estándar especial para volverse legibles por diferentes dispositivos de reproducción.
- Los metadatos se pueden almacenar en un archivo separado.
 - Esto significa que los metadatos se pueden vincular al archivo tras crear las secuencias de vídeo. Dicho tipo de metadatos puede ser material de bonus adicional descargado de la página web del proveedor de contenido, como incentivo para que el usuario visite la página web del proveedor de contenido o para un registro voluntario (para lograr fidelidad del cliente y seguimiento del cliente), etc.
 - Técnicamente, el almacenamiento de los metadatos en archivos especiales puede hacer esto.
 - Una alternativa podría ser almacenar estos metadatos en otros archivos de información, p. ej. en ".clpi" (ClipInfo) de Blu-ray o archivos ".mpls" (PlayList), o en los archivos ".IFO" de (HD-)DVD.
 - Sin embargo, este tipo de metadatos necesita un estándar especial para volverse legibles por diferentes dispositivos de reproducción.
- Los metadatos pueden ser entregados junto con el flujo de vídeo (p. ej. "cabeza hablando" del director de la película) para una película especial, p. ej. en un disco especial. Así, la entrega contiene todos datos adicionales necesarios para una reproducción superpuesta.
- Puede haber diferentes maneras para superponer los flujos de vídeo idénticos. Así, puede haber disponible más de un conjunto de metadatos para las mismas secuencias de vídeo. Pueden estar priorizadas implícita o explícitamente, o el usuario puede seleccionar el conjunto de metadatos preferido.

Ubicaciones de las secuencias de vídeo que serán superpuestas:

- Dichas secuencias de vídeo se pueden ubicar en el mismo medio (p. ej. HDD o disco óptico).
- Dichas secuencias de vídeo se pueden ubicar en medios diferentes. Puede realizarse una reproducción simultánea por medio de una red.
- Se pueden recibir secuencias de vídeo por separado de otras fuentes (p. ej. USB, memoria rápida, internet, etc.), p. ej. como vídeo transmitido por flujo. P. ej. la secuencia del "director de la película" viene directamente del servidor de internet del proveedor de contenido (transmisión por flujo) mientras la película real es leída de un disco óptico. El dispositivo de reproducción superpone ambas secuencias de vídeo simultáneamente interpretando los metadatos según la invención.

35 Contenido de los metadatos:

Los metadatos describen cómo superponer dos o más secuencias de vídeo en la misma pantalla de vídeo. Ventajosamente, esta superposición puede ser posible en un gran intervalo, es decir, muy flexible. Esto significa que los metadatos contienen información acerca de p. ej.

- metadatos de exposición geométrica
 - posición dentro de la pantalla
 - escalado de los fotogramas presentados
 - Los posibles niveles de escalado pueden limitarse a un conjunto fijo de valores para simplificar la realización de un escalado de este tipo, es decir, con el fin de simplificar los requisitos de hardware.
 - rotación antes de la exposición
 - Los posibles niveles de rotación pueden limitarse a un conjunto fijo de valores para simplificar la

realización de una rotación de este tipo, es decir, con el fin de reducir los requisitos de hardware.

- distorsión (p. ej. relación de aspecto), si la hay
 - evitar relaciones de aspecto, p. ej. los metadatos restringen una relación de aspecto distinta a la que indica el flujo de vídeo
- 5
 - volteo del vídeo entero horizontal y/o verticalmente
- información de recorte de fotograma, si lo hay
- evitar relaciones de aspecto de píxel, p. ej. los metadatos restringen una relación de aspecto distinta a la que indica el flujo de vídeo
 - metadatos de temporización de superposición
- 10
 - se usa información de temporización y puntos de acceso de flujo para sincronizar el proceso de superposición completo. Esta puede ser campo de vídeo, fotograma de vídeo o fotograma de audio exactos, o sincronizados aproximadamente
 - si no hay disponible información de temporización específica, entonces se pueden usar valores predefinidos en su lugar,
- 15
 - en general, la información de temporización puede describir un momento inicial, momento final, una duración o una mezcla de cualquiera de estos.
 - información para navegar a través de una secuencia de vídeo, como Cadenas de Programas (PGC) en DVD publicadas en internet en
<http://www.dvd-replica.com/DVD/pgca.php>
- 20
 - bucles (sin fin), o parada tras una presentación completada
 - reordenar partes de una secuencia de vídeo para la reproducción
 - indicar posibles interacciones de usuario (como PGC)
 - información para cambiar entre diferentes metadatos geométricos y/o metadatos de efectos especiales. Puede ser campo de vídeo o fotograma de vídeo o sincronizados aproximadamente.
- 25
 - se puede proporcionar información en diferentes niveles de precisión
 - sincronizada aproximadamente (es decir, usualmente del orden de periodos de pocos fotogramas)
 - sincronizada estrictamente (es decir, preciso por fotograma o preciso por campo)
 - información para el audio
 - cambio a un flujo de audio apropiado o
- 30
 - indicar flujos de audio de superposición (indicar si se superpondrá o no un flujo de audio)
 - Nivel de superposición (p. ej. indicar qué flujo de audio se pondrá en el primer plano o en el fondo).
 - ubicación 2D o 3D (indicar la ubicación geométrica desde donde proviene o es audible el audio de un audio específico)
 - indicar nivel de volumen de cada flujo de audio antes de mezclar
- 35
 - efectos de audio (p. ej. añadir efecto de reverberación al audio)
 - si estos comportamientos de audio están disponibles, entonces evitan los efectos especiales para audio de los otros tipos de metadatos.
- Metadatos de efectos especiales
 - fundido de entrada/salida
- 40
 - cambios de color
 - fotograma para la frontera de fotogramas

- indicar método de superposición usado
 - información de inserción croma para transparencia (principalmente serán necesarios gamas de color porque la compresión de vídeo puede entregar diversos valores dentro de la misma área de transparencia)
- 5 ▪ valor de transparencia global para
 - una única secuencia de vídeo
 - varias secuencias de vídeo
 - añadir subvídeo al vídeo de fondo
 - exclusivo o combinación de subvídeo con vídeo de fondo
- 10 ○ Diferentes efectos para diferentes tipos de pantalla, p. ej. los valores de transparencia dependen del tipo de pantalla usada.
- Son posibles muchos otros efectos
- Efectos de audio
 - efectos de audio (p. ej. añadir efecto de reverberación al audio)
- 15 • Metadatos de flujo
 - información técnica detallada acerca de una secuencia de vídeo
 - formato de compresión (p. ej. secuencia de bitmaps, MPEG-2, MPEG-4, VC-1, etc.)
 - resolución
 - relación de aspecto (para un fotograma y para un píxel)
 - 20 ▪ espacio cromático usado (p. ej. ITU 709 para HDTV o ITU 624 para SDTV, etc.)
 - tasa de fotogramas
 - entrelazado/progresivo
 - Valores de inserción croma, si la hay
 - valores promedio o medios para el brillo
 - 25 ○ información detallada de contenido acerca de una secuencia de vídeo
 - género
 - actores
 - etc.
 - propiedades de audio
 - 30 ▪ indicar nivel de volumen del flujo de audio
 - ubicación 2D o 3D (indicar la ubicación geométrica desde donde proviene o es audible el audio de un audio específico)
 - indicar nivel de volumen de cada flujo de audio antes de mezclar
- 35 Si los metadatos están en el segundo subflujo, es decir, para sobreponerse a otros, se debe considerar lo siguiente. Se pueden ubicar
 - a) en el flujo elemental de vídeo (ES) (user_data(), vídeo MPEG-2, MPEG-4, VC-1)
 - b) en el flujo de transporte (m2ts) en el que está contenido el vídeo ES; probablemente también el audio secundario está contenido en el mismo flujo de transporte (TS)
 - c) en el archivo ClipInfo (clpi) del TS correspondiente

- d) en el archivo PlayList (mpls)
- e) como archivo de metadatos separado

Ejemplos de contenidos/finalidad de metadatos

- a) posición (incluida modificación exacta de fotograma)
- 5 b) escalado (incluida modificación exacta de fotograma o dinámica)
- c) selección de factores de escalado fijos predefinidos
- d) información de inserción croma para transparencia (NB: considérese la gama de color debido a la compresión)
- e) valor de transparencia global
- 10 f) transparencia dependiente de pantalla (tipo de pantalla)
- g) información de gama de color ITU-(T o R) 709 para HDTV; ITU 6?? para SDTV
- h) evitar información de relación de aspecto de píxel
- i) información de volteo
- j) información de rotación
- 15 k) efectos
- l) fotograma para los PIP (imagen en imagen)
- m) información de inicio y parada para PIP asíncronas
- n) brillo base, para lograr múltiples PIP uniformes/igualables en apariencia atractiva
- o) cálculos en ambas imágenes (p. ej., operaciones ráster) (p. ej. sustraer objetos de imágenes)

20 Todos los metadatos según la invención opcionalmente también deben poder modificar el fotograma de subvídeo.

Además, todos estos metadatos únicamente pueden ser para valores predeterminados, es decir, pueden ser sobrescritos por, p. ej., llamadas API BD-Java. Los metadatos mencionados anteriormente degeneran entonces en estos valores predeterminados.

25 El Vídeo Secundario es típicamente contenido que puede ser descargado de internet y entonces se almacena en almacenamiento local, de modo que los metadatos mixtos no son almacenados necesariamente en el disco óptico (puede ser un disco ROM). También es posible que datos de vídeo primario y secundario estén ambos en el mismo soporte de datos, p. ej. disco ROM óptico.

30 Una aplicación ejemplar es para una escena particular de un primer vídeo, p. ej. expresada como marcas de tiempo inicial y final, los metadatos según la invención definen qué área de la pantalla es relativamente oscura, y/o qué área es relativamente brillante. Esta información también se puede combinar con áreas prohibidas o preferibles para sobreposiciones, en las que esta escena sirve como fondo. Si se debe sobreponer un segundo vídeo a esta escena en una ventana separada, el segundo vídeo contiene metadatos, p. ej., que indican que su contenido es más brillante, o que indican que su contenido se debe presentar en un área oscura, la ventana puede ser posicionada automáticamente para que cumpla todos estos requisitos, es decir, es posicionada automáticamente en un área indicada como "oscuro". Además, si cambia el fondo, los metadatos del primer vídeo indican este cambio y el segundo vídeo sobrepuesto puede ser movido automáticamente a otra área, donde se cumplen todas las condiciones. Este movimiento es preferiblemente un desplazamiento suave en lugar de un salto.

40 En otro ejemplo, el primer flujo de vídeo es un juego, p. ej., producido sintéticamente. Se pueden sobreponer flujos adicionales de vídeo que proporcionan, p. ej., diferentes vistas de otros personajes del juego. P. ej. un jugador A tiene en su pantalla áreas en las que se muestran las acciones concurrentes de otros jugadores B y C. Estos otros jugadores pueden ser humanos, p. ej. en un juego en línea, o virtuales. Por lo tanto, se sobreponen otros dos flujos a la escena de fondo que muestran la propia vista de los jugadores. La posición de los fotogramas sobrepuestos es definida por metadatos que se conectan a los flujos respectivos. Estos metadatos también pueden depender, p. ej. de las acciones que realiza el jugador respectivo.

45 En incluso otro ejemplo, el primer flujo de vídeo es una imagen estática o una escena corta que se repite continuamente, tal como un fondo que sirve como menú. Se sobreponen varios flujos de vídeo, correspondientes p. ej. a resúmenes de los vídeos que el usuario tiene en su biblioteca de vídeos. El primer flujo de vídeo tiene

metadatos conectados que proporcionan áreas de pantalla a los flujos de vídeo, de modo que 1. cada vídeo se puede mostrar en un icono, 2. se puede mostrar un número definido de iconos vídeo, y 3. se deja espacio entre los iconos de modo que el fondo se puede ver pero cada icono es tan grande como sea posible. P. ej. los metadatos definen que se exponen doce flujos de vídeo en cuatro columnas y tres filas, donde

- 5 cada columna tiene el 25 % de la anchura, cada fila tiene el 33,3 % de la altura y cada icono usa el 80 % de su anchura y altura es decir, el 81 % de su área, de modo que en el 10 % restante en cada canto se ve el fondo. En un ejemplo similar, el fondo se debe esponer, pero independiente del número de flujos de vídeo disponibles cada icono debe tener un área máxima en la pantalla. P. ej. hay dos vídeos en la biblioteca, cada uno de ellos obtiene la mitad del área de pantalla. Generalmente, cada uno de los vídeos obtiene el mismo reparto de la pantalla (hasta un cierto tamaño mínimo), que es controlado dinámicamente por los metadatos que se conectan al fondo.
- 10

Como ventaja particular, el tipo relativo descrito de definición de área permite la adaptación a cada tamaño o resolución de pantalla respectivamente.

REIVINDICACIONES

1. Método para sobreponer dos o más flujos de vídeo dinámicamente, donde se asocian metadatos a al menos uno de los flujos de vídeo mediante una lista de reproducción, que comprende las etapas de

- exponer un primer flujo de vídeo que es uno de dichos flujos de vídeo;
- 5 - extraer dichos metadatos desde un archivo de lista de reproducción que pertenece a al menos un segundo flujo de vídeo distinto a dicho primer flujo de vídeo;
- detectar que los metadatos extraídos definen cómo sobreponer el segundo flujo de vídeo a dicho primer flujo de vídeo, en donde dichos metadatos contienen datos de posición que definen un área en pantalla para sobreposición y una referencia al primer flujo de vídeo, y donde el contenido de la imagen del primer
10 flujo de vídeo no se analiza; y
- sobreponer al menos el segundo flujo de vídeo al primer flujo de vídeo, según definen dichos metadatos.

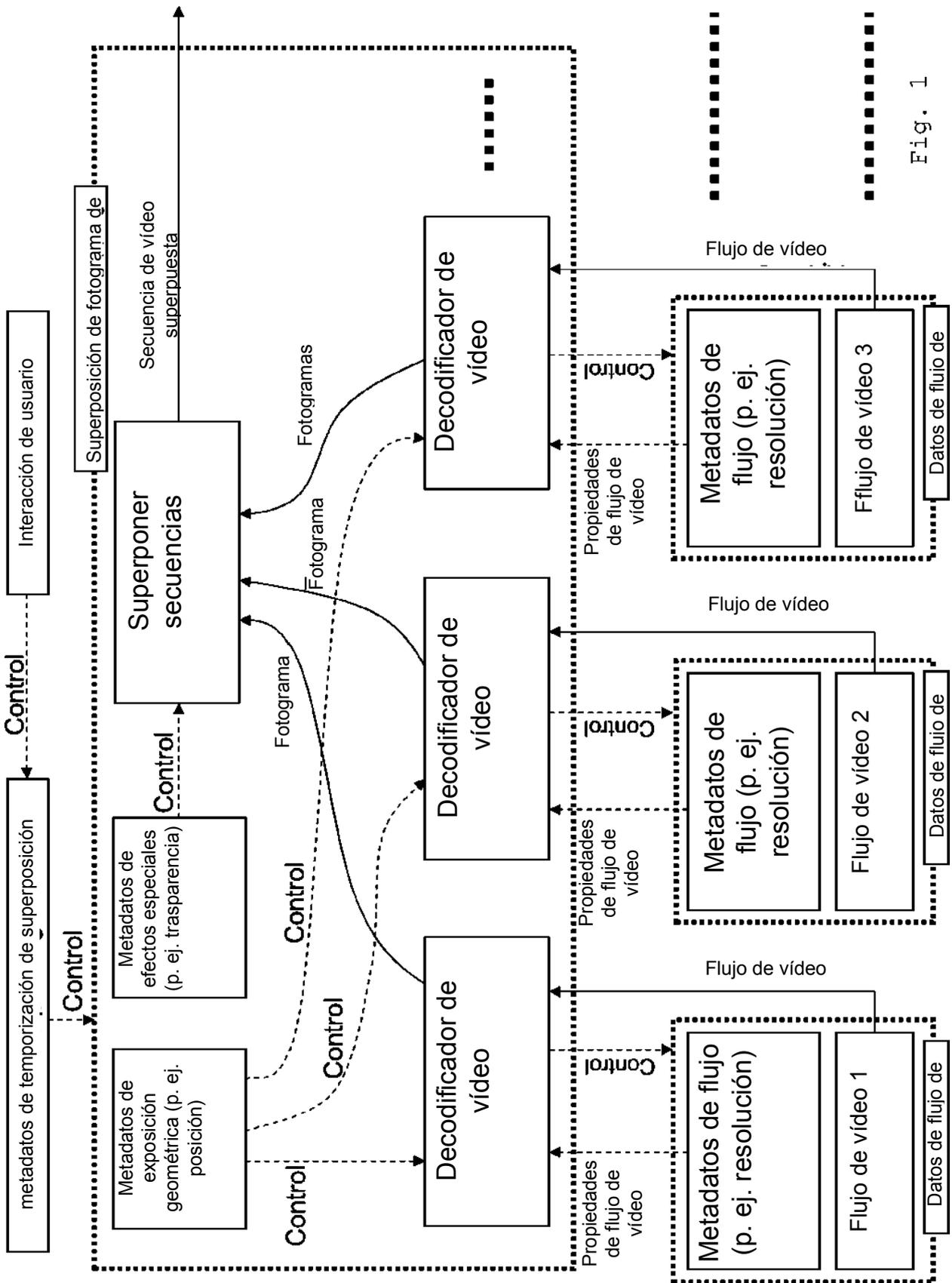


Fig. 1

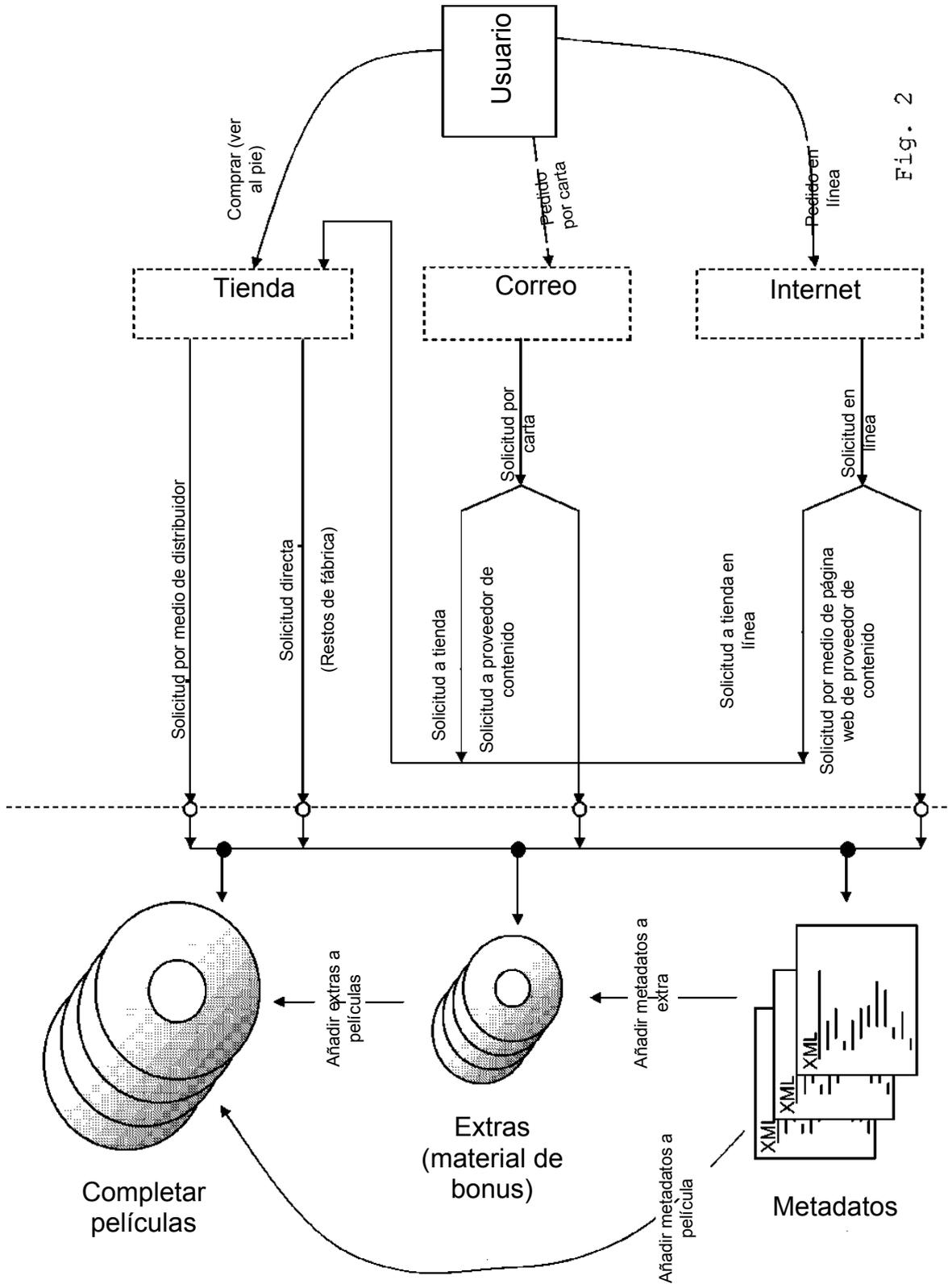


Fig. 2

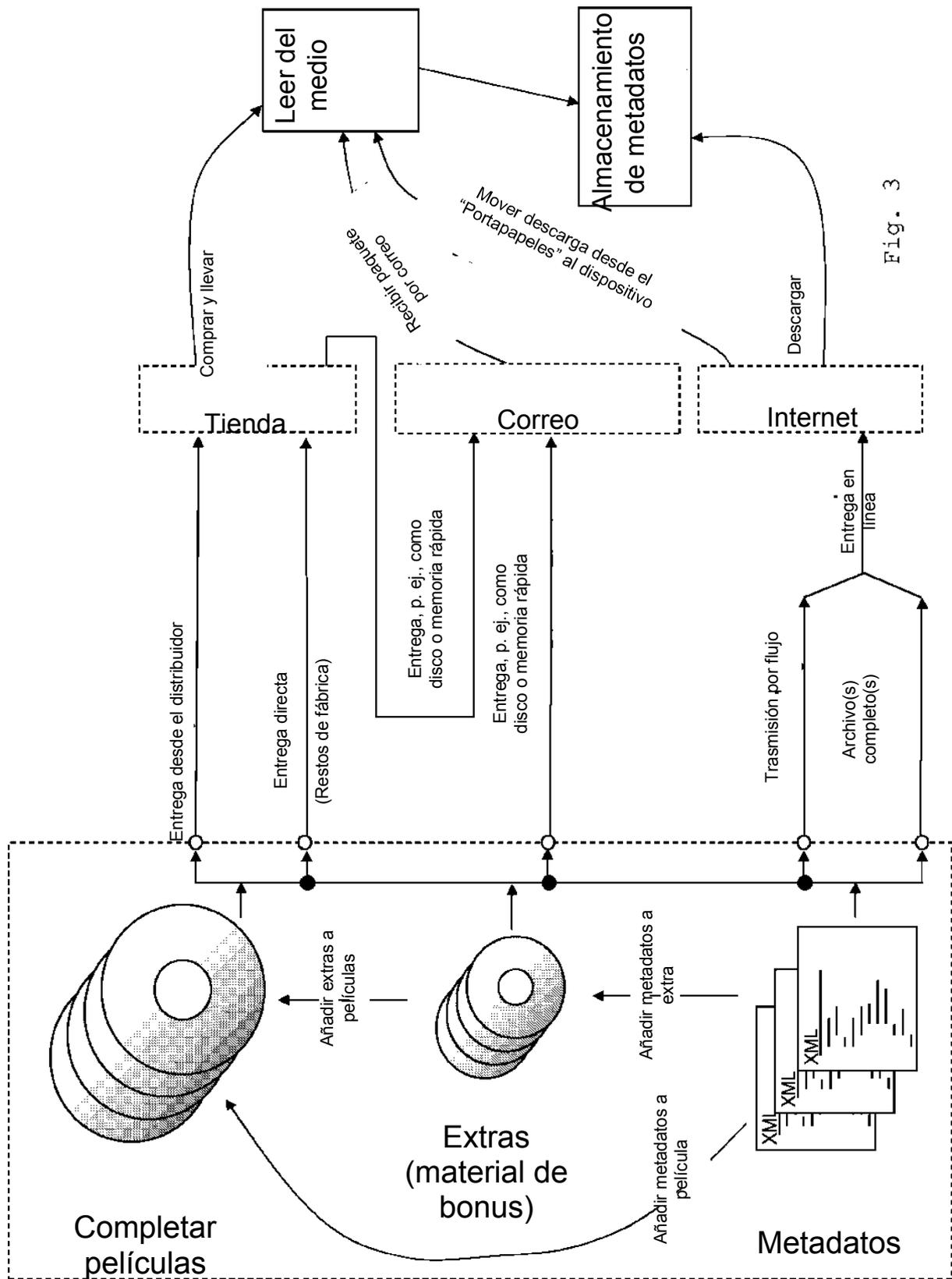


Fig. 3

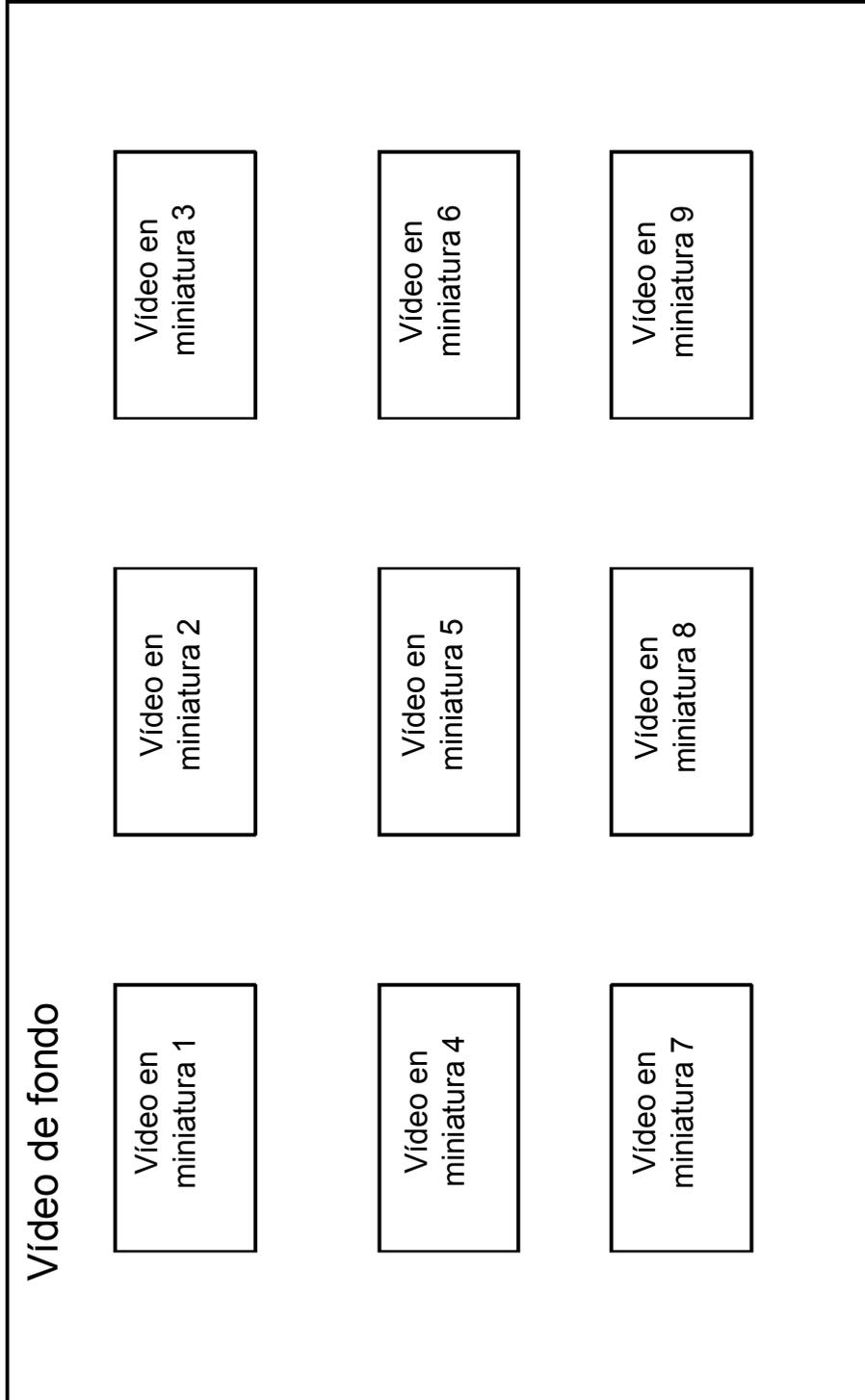


Fig. 4

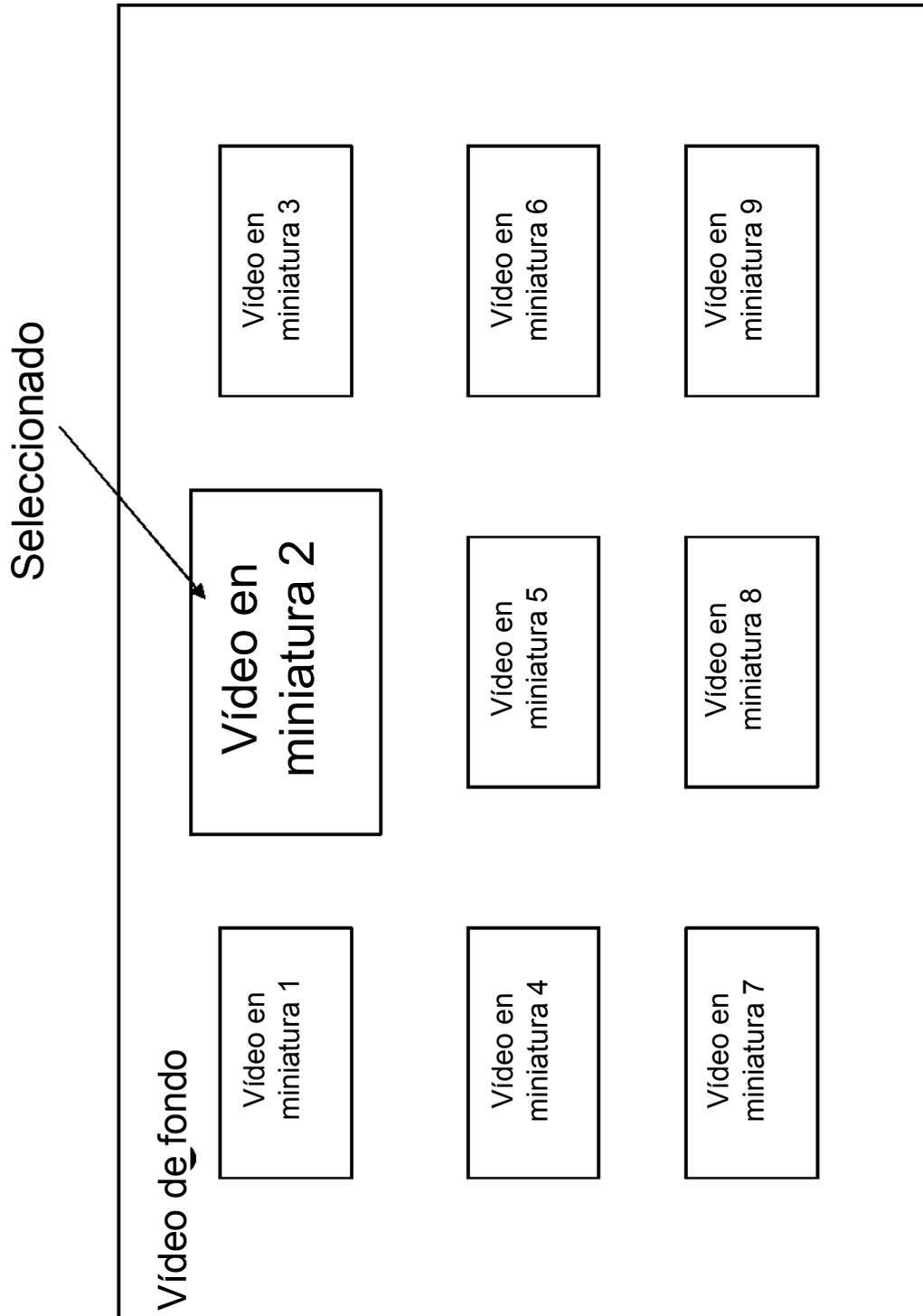


Fig. 5

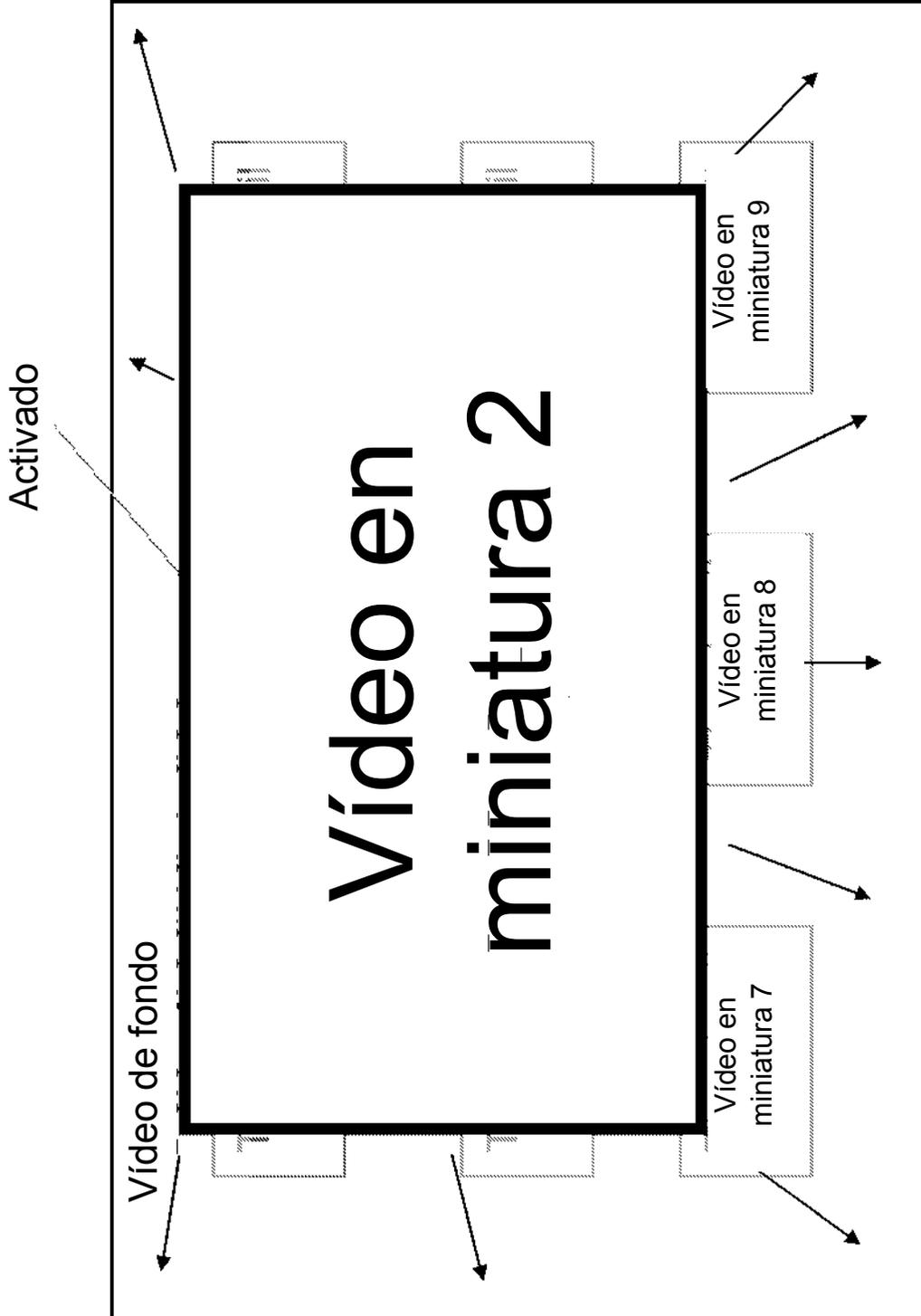


Fig. 6