

19



OFICINA ESPAÑOLA DE  
PATENTES Y MARCAS

ESPAÑA



11 Número de publicación: **2 983 243**

51 Int. Cl.:

**G06F 16/48** (2009.01)

**G06F 21/16** (2013.01)

**G06F 9/445** (2008.01)

12

TRADUCCIÓN DE PATENTE EUROPEA

T3

96 Fecha de presentación y número de la solicitud europea: **07.11.2014 E 21160532 (4)**

97 Fecha y número de publicación de la concesión europea: **15.05.2024 EP 3913500**

54 Título: **Identificación de componentes multimedia**

30 Prioridad:

**08.11.2013 GB 201319783**

**23.01.2014 GB 201401130**

**05.06.2014 GB 201410032**

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

**22.10.2024**

73 Titular/es:

**FRIEND FOR MEDIA LIMITED (100.0%)**

**11 Brindley Place 2 Brunswick Place**

**Birmingham B1 2LP, GB**

72 Inventor/es:

**ADAMS, RICHARD WILLIAM;**

**GLADDEN, LAURENCE KIT;**

**DEL STROTHER, WILLIAM;**

**FRIEND, JONATHAN NORMAN;**

**DEMETRIADES, GEORGE DOUGLAS y**

**TROOP, CHRIS**

74 Agente/Representante:

**DEL VALLE VALIENTE, Sonia**

**ES 2 983 243 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

## Identificación de componentes multimedia

5 Las invenciones de la presente memoria se refieren a la identificación de componentes multimedia y sistemas relacionados.

10 En algunas realizaciones, la invención se refiere a la derivación de un conjunto de huellas digitales a partir de un componente multimedia. En particular, pero no exclusivamente, las huellas digitales pueden generarse para datos multimedia. Aún más, pero nuevamente no exclusivamente, las huellas digitales podrían generarse para datos de video que podrían ser datos transmitidos. Las realizaciones pueden usar las huellas digitales para identificar datos multimedia dentro de un almacén o librería de datos multimedia.

15 En algunas realizaciones, esta invención se refiere a sistemas y métodos para la identificación de datos contenidos en componentes multimedia. Por tanto, las realizaciones de la invención pueden permitir la identificación de datos multimedia.

20 En algunas realizaciones, esta invención se refiere a sistemas y métodos para la identificación de usuarios y/o dispositivos a los que se envió contenido multimedia. Algunas realizaciones se refieren a la extracción de marcas de agua del contenido multimedia. En particular, pero no exclusivamente, las realizaciones se relacionan con el procesamiento de video, que podría ser video transmitido.

25 El advenimiento de las redes informáticas, como Internet, ha significado que los dispositivos informáticos se estén interconectando masivamente, lo que facilita la copia y redistribución de datos a una escala sin precedentes. Los datos son muy valiosos y se conocen técnicas para intentar protegerlos. Dichas técnicas incluyen la marca de agua de los datos para que los datos tengan una firma específica y también la toma de huellas digitales de los datos para detectar características de los datos sin imponer realmente una característica a los datos.

30 Una de estas técnicas de generación de huellas digitales se describe en el documento "Spatio-temporal transform based video hashing" de Baris Coskun, Bulent Sankur, Nasir Memon publicado en IEEE Transactions on multimedia Vol. 8, no. 6, diciembre de 2006. Sin embargo, hay deficiencias en las enseñanzas de este artículo a la hora de proporcionar una técnica utilizable.

35 A medida que ha aumentado el ancho de banda disponible en dichas redes, la naturaleza de los datos almacenados y vistos en esas redes ha cambiado de información basada en texto a la actualidad, donde ahora es común tener datos basados en video. No obstante, sea cual sea el tipo de datos, existe un deseo continuo de identificar copias ilícitas de los datos, generar datos de uso para los datos, proporcionar mecanismos para entregar contenido a los usuarios y similares. Dadas las enormes cantidades de datos disponibles en redes como Internet y WWW, el monitoreo del uso de los datos en esas redes se ha vuelto difícil y ahora es virtualmente imposible de realizar manualmente.

40 Puede ser un problema identificar la fuente desde la que se originan los datos, ya que ahora se están escribiendo páginas WWW que extraen contenido de una ubicación de red remota o no relacionada. Así, por ejemplo, una página WWW puede extraer datos de video en una trama presente en esa página. Como tal, es posible que el Localizador Uniforme de Recursos (URL) de la página WWW no identifique la última fuente de los datos del video (u otros). Puede ser un problema importante identificar esa última fuente.

45 Además, como las redes se utilizan como mecanismo de entrega de contenido multimedia de eventos tales como eventos deportivos, películas y similares, se puede entregar una cantidad significativa de contenido, a menudo contenido de video. Como tal, es un problema encontrar técnicas que sean lo suficientemente rápidas para procesar cantidades significativas de información.

Según un primer aspecto de la invención, se proporciona un sistema informático dispuesto para determinar si un conjunto de datos candidato coincide con un conjunto de datos de referencia.

55 El documento US 2011/299721 describe que se identifican múltiples componentes característicos candidatos del contenido multimedia o las matrices de proyección (u otras funciones de hash, por ejemplo, proyecciones no lineales). Cada una de las matrices de proyección candidatas (u otras funciones hash) incluye una matriz de coeficientes que se relacionan con las características candidatas. Se selecciona un subgrupo de las características candidatas o las matrices de proyección (u otras funciones hash) basándose, al menos parcialmente, en una combinación optimizada de al menos dos características de las características candidatas o matrices de proyección (u otras funciones hash). Las huellas digitales multimedia que identifican de forma exclusiva el contenido multimedia se derivan del subgrupo optimizado seleccionado. Se pueden diseñar matrices de proyección óptimas (u otras funciones de hash). Las características de rendimiento o sensibilidad (por ejemplo, el tiempo de búsqueda) de las huellas digitales se equilibran así con las características de robustez de las mismas.

65

El documento WO 2009/026564 A1 usa huellas digitales de vídeo para copiar/recuperar duplicados. Las sub-huellas digitales correspondientes a un segmento de video comprenden una secuencia de valores (claves). La secuencia se divide en 25 bandas LSH (hash sensible a la localidad). Cada banda LSH tiene cuatro valores de bytes y se denomina clave LSH/banda. Una tabla LSH de índice inverso asigna las claves a los identificadores de sub-huellas digitales que comprenden una clave determinada.

Según un primer aspecto de la invención, se proporciona un procedimiento implementado por ordenador para derivar una secuencia de huellas digitales a partir de uno o más componentes multimedia, tal como se reivindica en la reivindicación 1,

Según un segundo aspecto de la invención, se proporciona un sistema informático dispuesto para generar una secuencia de huellas digitales a partir de uno o más componentes multimedia, como se reivindica en la reivindicación 8.

Según un tercer aspecto de la invención, se proporciona un medio legible por ordenador no transitorio que almacena un código de programa informático ejecutable para derivar una secuencia de huellas digitales a partir de uno o más componentes multimedia, como se reivindica en la reivindicación 13.

Las realizaciones preferidas del método, sistema y medio se describen en las reivindicaciones dependientes.

En una realización adicional que es útil para comprender el contexto más amplio de la invención reivindicada, se describe un sistema informático dispuesto para determinar si un conjunto de datos candidato coincide con un conjunto de datos de referencia. Convenientemente, se proporciona un dispositivo de procesamiento que está dispuesto para recibir el conjunto de datos candidato. Se puede proporcionar y disponer un generador de ventanas para seleccionar una parte del conjunto de datos. Se puede proporcionar y disponer un generador de huellas digitales para generar una huella digital a partir de la parte del conjunto de datos seleccionada por el generador de ventanas. Normalmente, el dispositivo de procesamiento está dispuesto para hacer que el generador de ventanas se desplace a través del conjunto de datos y seleccione una pluralidad de porciones del conjunto de datos a partir de al menos algunas de las cuales el generador de huellas digitales está dispuesto para generar una huella digital candidata, creando así una secuencia de huellas digitales candidatas. Se puede proporcionar y disponer un comparador para comparar al menos algunas de las huellas digitales candidatas dentro de la secuencia de huellas digitales candidatas con un almacén de huellas digitales de referencia para determinar si el almacén contiene huellas digitales de referencia que coinciden con al menos una de las huellas digitales candidatas. El dispositivo de procesamiento puede disponerse para determinar que la secuencia de huellas digitales candidatas coincide con un conjunto de huellas digitales de referencia en caso de que el comparador determine que se debe mantener un número mínimo de huellas digitales candidatas dentro de la secuencia de huellas digitales candidatas para hacer coincidir las huellas digitales respectivas con la secuencia de huellas digitales de referencia.

En una segunda realización que es útil para comprender el contexto más amplio de la invención reivindicada, también se describe un método implementado por ordenador para determinar si un conjunto de datos candidato coincide con un conjunto de datos de referencia. Normalmente, el método incluye recibir el conjunto de datos candidato en un dispositivo de procesamiento y hacer que el dispositivo de procesamiento mueva una ventana a través del conjunto de datos candidato a una pluralidad de posiciones para seleccionar una parte del conjunto de datos. El método puede generar una huella digital candidata a partir de la parte del conjunto de datos cubierta por la ventana para al menos algunas de las posiciones de la ventana, creando así una secuencia de huellas digitales candidatas para el conjunto de datos candidato. Normalmente, se hace una comparación de al menos algunas de las huellas digitales candidatas dentro de la secuencia de huellas digitales candidatas con un almacén de huellas digitales de referencia para determinar si el almacén contiene huellas digitales de referencia que se considera que coinciden con al menos una de las huellas digitales candidatas. El método puede determinar que la secuencia de huellas digitales candidatas coincide con una secuencia de huellas digitales de referencia si se mantiene al menos un número mínimo de huellas digitales candidatas dentro de la secuencia de huellas digitales candidatas para que coincidan con las huellas digitales respectivas dentro de la secuencia de huellas digitales de referencia.

En una tercera realización que es útil para comprender el contexto más amplio de la invención reivindicada, se describe un medio legible por ordenador no transitorio que almacena un código de programa informático ejecutable para determinar si un conjunto de datos candidato coincide con un conjunto de datos de referencia, siendo el código de programa ejecutable para realizar etapas que comprenden:

i) recibir el conjunto de datos del candidato en un dispositivo de procesamiento y hacer que el dispositivo de procesamiento mueva una ventana a través del conjunto de datos del candidato a una pluralidad de posiciones para seleccionar una parte del conjunto de datos y generar una huella digital del candidato a partir de la parte del conjunto de datos cubierta por la ventana para al menos algunas de las posiciones de la ventana, creando así una secuencia de huellas digitales del candidato para el conjunto de datos del candidato;

ii) comparar al menos algunas de las huellas digitales candidatas dentro de la secuencia de huellas digitales candidatas con un almacén de huellas digitales de referencia para determinar si el almacén contiene huellas digitales de referencia que se considera que coinciden con al menos una de las huellas digitales candidatas; en donde

5 iii) se considera que la secuencia de huellas digitales candidatas coincide con una secuencia de huellas digitales de referencia si se mantiene al menos un número mínimo de huellas digitales candidatas dentro de la secuencia de huellas digitales candidatas para que coincidan con las huellas digitales respectivas dentro de la secuencia de huellas digitales de referencia.

10 Según el primer aspecto de la invención, se proporciona un método implementado por ordenador para derivar una secuencia de huellas digitales a partir de uno o más componentes multimedia. El método comprende:

a) recibir el o cada componente multimedia en un dispositivo de procesamiento y hacer que ese dispositivo de procesamiento:

15 i) mueva una ventana a través del componente multimedia o de cada uno de ellos a una pluralidad de posiciones;

20 ii) seleccione una parte del componente multimedia o de cada uno de ellos;

iii) procese la parte del o cada componente multimedia para generar una huella digital indicativa de la parte del o de cada componente multimedia;

25 iv) almacene la huella digital;

v) procese la parte para generar una medida de calidad indicativa de la singularidad de la huella digital, en donde las medidas de calidad se determinan según un cambio en una propiedad entre tramas contiguos de uno o más componentes multimedia; y

30 vi) almacenar la medida de calidad en asociación con la primera huella digital.

En una cuarta realización que es útil para comprender el contexto más amplio de la invención reivindicada, también se describe un método implementado por ordenador para determinar si un conjunto de datos candidato coincide con un conjunto de datos de referencia. El método normalmente comprende al menos algunas de las siguientes etapas;

35 i) recibir el conjunto de datos candidato en un dispositivo de procesamiento y hacer que el dispositivo de procesamiento mueva una ventana a través del conjunto de datos candidato a una pluralidad de posiciones para seleccionar una parte del conjunto de datos. El dispositivo de procesamiento puede estar dispuesto para generar una huella digital candidata a partir de la parte del conjunto de datos cubierta por la ventana para al menos algunas de las posiciones de la ventana, creando así una secuencia de huellas digitales candidatas para el conjunto de datos candidato;

40 ii) procesar la parte del conjunto de datos cubierta por la ventana para generar una medida de calidad indicativa de la singularidad de la huella digital. El dispositivo de procesamiento puede estar dispuesto para almacenar la medida de calidad en asociación con la huella digital; y

45 iii) comparar al menos algunas de las huellas digitales candidatas dentro de la secuencia de huellas digitales candidatas con un almacén de huellas digitales de referencia para determinar si el almacén contiene huellas digitales de referencia que se considera que coinciden con al menos una de las huellas digitales candidatas. La comparación puede organizarse de manera que la huella digital candidata coincida con la huella digital de referencia si la medida de calidad es mayor que un umbral predeterminado.

50 En una quinta realización que es útil para comprender el contexto más amplio de la invención reivindicada, se describe además un método implementado por ordenador para derivar un conjunto de huellas digitales a partir de un conjunto de datos. El método puede comprender al menos algunas de las siguientes etapas;

i) recibir el conjunto de datos de referencia en un dispositivo de procesamiento y hacer que el dispositivo de procesamiento mueva lo que normalmente es una ventana de longitud fija a través del conjunto de datos y:

55 a) seleccionar los datos cubiertos por una primera posición de la ventana en relación con el conjunto de datos, generar una primera huella digital desde la primera posición y puede almacenar la primera huella digital;

60 b) posteriormente, el dispositivo de procesamiento puede disponerse para mover la ventana, en relación con el conjunto de datos, a una posición adicional diferente de la primera. En la posición adicional, el dispositivo de procesamiento puede estar dispuesto para generar una huella digital adicional desde la posición adicional y puede almacenar la huella digital adicional; y

65

c) se puede repetir la etapa b) para generar y almacenar más huellas digitales.

En una sexta realización que es útil para comprender el contexto más amplio de la invención reivindicada, se describe un sistema dispuesto para determinar si un conjunto de datos candidato coincide con un conjunto de datos de referencia. El sistema comprende al menos algunos de los siguientes elementos:

i) al menos un generador de huellas digitales de referencia dispuesto para recibir un conjunto de datos de referencia y generar una o más huellas digitales de referencia a partir del mismo;

ii) un almacén de huellas digitales de referencia al que se añade la o cada huella digital de referencia generada por el generador de huellas digitales a través de una conexión entre estas; y

iii) al menos un generador de huellas digitales de candidatos, dispuesto para recibir un conjunto de datos de candidatos y generar una o más huellas digitales de candidatos a partir del mismo; y

en donde el sistema puede disponerse de modo que las huellas digitales candidatas generadas por el generador de huellas digitales candidatas se comparen con las huellas digitales de referencia almacenadas en el almacén de huellas digitales para determinar si una huella digital de referencia almacenada en el almacén es sustancialmente la misma que la huella digital candidata, a fin de identificar si el conjunto de datos candidato que generó la huella digital candidata es el mismo, o al menos sustancialmente el mismo, que un conjunto de datos de referencia.

Las realizaciones que proporcionan un sistema de este tipo se consideran ventajosas, ya que proporcionan lo que es potencialmente un método automatizado para identificar un conjunto de datos candidatos que es robusto y puede procesar conjuntos de datos candidatos en tiempo real, o al menos pseudoreal. Además, el sistema puede escalarse fácilmente con el fin de proporcionar lo que debería permitirle proporcionar servicios de monitorización a gran escala.

Otro método que es útil para comprender el contexto más amplio de la invención reivindicada puede identificar un archivo multimedia candidato. El método puede comprender al menos algunas de las siguientes etapas:

generar una o más huellas digitales del candidato representativas del archivo multimedia del candidato;

comparar una o más huellas digitales candidatas con un almacén de huellas digitales de referencia en donde cada una de las huellas digitales de referencia es representativa de un archivo multimedia de referencia;

utilizar la comparación para determinar si el archivo multimedia candidato es el mismo, o al menos sustancialmente el mismo, que un archivo multimedia de referencia para el que se dispone de una huella digital en la biblioteca; y

si la comparación determina que el archivo multimedia candidato es el mismo, o al menos sustancialmente el mismo, que el archivo multimedia de referencia identificando el archivo multimedia candidato como el archivo multimedia de referencia.

También se describe un método para capturar contenido multimedia transmitido accediendo a un flujo multimedia a través de una conexión de red que es útil para comprender el contexto más amplio de la invención reivindicada. El método comprende al menos algunas de las siguientes etapas:

identificar un flujo multimedia disponible en una red accediendo a través de una conexión a esa red;

hacer que un primer mecanismo de procesamiento inicie el flujo multimedia, a través de la conexión a la red, y actúe como un servidor multimedia para retransmitir el flujo multimedia;

recibir el flujo retransmitido en un segundo mecanismo de procesamiento; y

capturar al menos una de la información de cabecera y los medios del flujo retransmitido.

Las realizaciones que proporcionan un método de este tipo se consideran ventajosas, ya que proporcionan un método para obtener metadatos y/o los medios subyacentes a partir de un flujo de medios. Los metadatos se pueden usar para identificar la fuente principal del flujo multimedia, lo que puede ser útil para tomar medidas coercitivas contra una parte que está generando el flujo multimedia. También es útil poder reconstruir los archivos que contienen los medios transmitidos en una transmisión para analizar esos medios. Se apreciará que los medios transmitidos en flujo continuo a menudo se entregan en fragmentos y puede ser un problema obtener los medios subyacentes para compararlos.

En una realización adicional que es útil para comprender el contexto más amplio de la invención reivindicada, se describe un método para capturar contenido multimedia transmitido accediendo a un flujo multimedia a través de una conexión de red. El método puede comprender al menos algunas de las siguientes etapas:

identificar un flujo multimedia a través de una conexión de red;

hacer que un primer mecanismo de procesamiento inicie el flujo multimedia, a través de la conexión de red, y actúe como un servidor multimedia para retransmitir el flujo multimedia;

5

recibir el flujo retransmitido en un segundo mecanismo de procesamiento; y

capturar al menos una de la información de cabecera y el contenido multimedia del flujo retransmitido.

10 Se describe un método para identificar el contenido multimedia por ser útil para comprender el contexto más amplio de la invención reivindicada, comprendiendo el método al menos algunas de las siguientes etapas:

utilizar un mecanismo de procesamiento para:

15 i) recopilar metadatos de una red, asociándose los metadatos a un conjunto de datos candidato;

ii) procesar los metadatos para obtener datos del cronograma relacionados con el conjunto de datos del candidato y usar los datos del cronograma para determinar un cronograma para capturar el conjunto de datos del candidato; y

20

iii) capturar el conjunto de datos del candidato según el cronograma determinado a partir de los datos del cronograma.

25 Las realizaciones que proporcionan un método de este tipo se consideran ventajosas porque facilitan la monitorización automática de los conjuntos de datos en una red. La recopilación automatizada de metadatos se puede utilizar para ayudar a procesar grandes volúmenes de datos que de otro modo no serían posibles.

30 En otra realización que es útil para comprender el contexto más amplio de la invención reivindicada, se describe un método para entregar una versión mejorada de un archivo multimedia a un usuario después de que ese usuario reproduzca una versión inicial del archivo multimedia. El método comprende al menos algunas de las siguientes etapas: generar una huella digital de referencia a partir de una versión fuente del archivo multimedia y almacenar esa huella digital de referencia en un almacén; permitir a un usuario reproducir la versión inicial del archivo multimedia; identificar la versión inicial del archivo multimedia mediante la generación de una huella digital candidata e identificar esa huella digital candidata mediante una búsqueda en el almacén de huellas digitales de referencia; y entregar al usuario la versión mejorada del archivo multimedia, en comparación con el archivo multimedia inicial, que se ha identificado en la búsqueda en el almacén de huellas digitales de referencia.

35

40 En otra realización que es útil para comprender el contexto más amplio de la invención reivindicada, se describe un método computarizado para identificar un archivo multimedia candidato. El método comprende al menos algunas de las siguientes etapas: hacer que uno o más dispositivos de procesamiento: capturen, desde una red a la que el dispositivo de procesamiento tiene acceso, el archivo multimedia candidato y los metadatos asociados al archivo multimedia candidato, en donde los metadatos especifican al menos la ubicación de red desde la que se originó el archivo multimedia candidato; determinar si se conoce el archivo multimedia candidato; y si se conoce el archivo multimedia candidato, determinar si es apropiado enviar una comunicación a la ubicación de red especificada en los metadatos.

45

50 En otra realización que es útil para comprender el contexto más amplio de la invención reivindicada, se describe un sistema dispuesto para identificar un archivo multimedia candidato, cuyo sistema comprende al menos algunos de los siguientes elementos: un elemento de captura multimedia, dispuesto para tener acceso a una red, y dispuesto además para capturar, desde la red, el archivo multimedia candidato y los metadatos asociados al archivo multimedia candidato, en donde los metadatos especifican al menos la ubicación de red desde la que se originó el archivo multimedia candidato; un mecanismo de entrada dispuesto para permitir a un usuario identificar un archivo multimedia candidato; y un almacén de datos de cumplimiento dispuesto para almacenar metadatos asociados a los archivos multimedia de los candidatos identificados por un usuario.

55

En otra realización más que es útil para comprender el contexto más amplio de la invención reivindicada, se describe un método para determinar la identidad de un usuario al que se entregó el contenido multimedia, comprendiendo el método al menos algunas de las siguientes etapas:

60 i) monitorizar una red en busca de contenido multimedia que tenga una huella digital que coincida con una huella digital de referencia mantenida dentro de un almacén de huellas digitales de referencia para generar contenido multimedia candidato;

ii) procesar el contenido multimedia candidato para extraer una o más marcas de agua codificadas en el mismo; y

65

iii) decodificar la o cada marca de agua para obtener un código de identidad que identifique al usuario al que se envió el contenido multimedia.

5 En otra realización que es útil para comprender el contexto más amplio de la invención reivindicada, se describe un método para proteger el contenido multimedia, que comprende al menos algunas de las siguientes etapas: disponer una plataforma de distribución multimedia para que tenga una pluralidad de dispositivos objetivo conectados a la misma y disponer además la plataforma de distribución multimedia para entregar el contenido multimedia a cada uno de los dispositivos objetivo; registrar una huella digital de referencia que representa el contenido multimedia; disponer la plataforma para generar una versión con marca de agua del contenido multimedia para cada dispositivo objetivo al que se entrega el contenido multimedia, versión con marca de agua dispuesta para identificar el dispositivo objetivo al que se entregó el contenido multimedia; monitorizar una red para detectar contenido multimedia que tenga una huella digital que sea la misma, o al menos sustancialmente la misma, que la huella digital de referencia, donde cada contenido multimedia así detectado proporciona un contenido multimedia candidato a decodificar; y extraer la marca de agua del contenido multimedia candidato o de cada uno de ellos para identificar el dispositivo objetivo al que se entregó el contenido multimedia.

La extracción de la marca de agua puede incluir la alineación temporal del contenido multimedia candidato con respecto a una referencia.

20 Convenientemente, el proceso de comparación de huellas digitales empleado se utiliza para proporcionar una alineación aproximada del contenido multimedia candidato y la referencia. Esto se considera ventajoso porque facilitan la extracción de la marca de agua al tiempo que utilizan parte del mecanismo de identificación (es decir, la coincidencia de las huellas digitales) para aumentar la robustez de esa extracción.

25 Las características pueden extraerse de los medios candidatos y usarse para proporcionar una alineación precisa de los medios candidatos con la referencia. Se cree que esto aumenta la precisión de la extracción de la marca de agua al identificar con mayor precisión la posición inicial de cualquier marca de agua incrustada.

30 El contenido multimedia puede dividirse en una pluralidad de segmentos de duración limitada y puede mantener al menos una primera y una segunda copia de cada segmento de duración limitada. Esto permite que el método se utilice con protocolos de transmisión adaptativos. Normalmente, la versión con marca de agua se genera mediante la combinación de segmentos de duración limitada de la primera y la segunda copia de los segmentos de duración limitada.

35 Convenientemente, se puede aplicar una marca de agua diferente a una primera copia de cada segmento de duración limitada en comparación con la segunda copia de ese segmento de duración limitada. Se puede aplicar un código de identidad al contenido multimedia a través de los segmentos de duración limitada.

40 Los segmentos de duración limitada tienen una longitud predeterminada, que convenientemente está sustancialmente entre 0,1 segundos y 15 segundos, o más preferiblemente sustancialmente en el intervalo de 5 segundos y 10 segundos.

45 Cada una de las longitudes predeterminadas de la marca de agua puede disponerse para proporcionar un dígito del código de identidad que puede proporcionar un poco de número binario.

El contenido multimedia puede transmitirse a través de una red de área amplia (WAN). Convenientemente, la red utiliza los protocolos TCP/IP. La red puede ser Internet.

50 La entrega del contenido multimedia se puede terminar al o a cada usuario identificado y/o dispositivo objetivo identificado.

55 El contenido multimedia puede ser contenido de vídeo. Se puede aplicar una marca de agua sustancialmente a cada trama del contenido de vídeo. El proceso de extracción de la marca de agua se vuelve más robusto, ya que existen múltiples posibilidades de extraer la marca de agua correcta, ya que hay una pluralidad de tramas por segmento de duración limitada. Las marcas de agua extraídas de cada trama de un segmento de duración limitada pueden puntuarse o analizarse estadísticamente de otro modo para determinar la marca de agua codificada en ese segmento de duración limitada.

60 En otra realización que es útil para comprender el contexto más amplio de la invención reivindicada, se describe un método para detectar una marca de agua en el contenido multimedia, que comprende al menos algunas de las siguientes etapas: usar contenido multimedia de referencia para generar contenido multimedia entregado en el que una marca de agua esté incrustada en al menos un segmento de duración limitada del mismo; registrar una huella digital del contenido multimedia de referencia utilizado para generar el contenido multimedia entregado; identificar uno o más contenidos multimedia de los candidatos; procesar el contenido multimedia candidato o cada uno de ellos para determinar si tienen la misma huella digital, o al menos sustancialmente la misma, que el contenido multimedia de referencia; usar la determinación de la huella digital del contenido multimedia candidato para obtener información de

temporización del contenido multimedia candidato para permitir identificar el segmento o cada segmento de duración limitada dentro del contenido multimedia candidato; y extraer una marca de agua del segmento de duración limitada o de cada uno de ellos.

5 En otra realización que es útil para comprender el contexto más amplio de la invención reivindicada, también se describe un método adicional para detectar una marca de agua dentro del contenido multimedia, comprendiendo el método adicional al menos una de las siguientes etapas: usar contenido multimedia de referencia para generar contenido multimedia entregado en el que una marca de agua esté incrustada en al menos un segmento de duración limitada del mismo; grabar información de referencia para el contenido multimedia de referencia utilizado para generar el flujo multimedia entregado; identificar uno o más contenidos multimedia de los candidatos; procesar el contenido multimedia candidato o cada uno de ellos y extraer al menos una característica del contenido multimedia candidato; comparar la característica extraída del contenido multimedia candidato con la información de referencia; utilizar la comparación para determinar el inicio del segmento de duración limitada; y extraer una marca de agua del segmento de duración limitada o de cada uno de ellos.

15 También se describe un sistema informático dispuesto para proporcionar cualquiera de los métodos descritos anteriormente y un medio legible por ordenador no transitorio que contiene instrucciones para hacer que un ordenador utilice cualquiera de los métodos descritos anteriormente.

20 El experto en la técnica apreciará que el término conjunto de datos es más amplio e incluye datos tales como componentes multimedia. Además, los componentes multimedia pueden incluir medios transmitidos por streaming, incluido un protocolo de transmisión adaptativo (tal como transmisión en vivo HTTP (HLS)); transmisión HTTP adaptativa 3GPP; transmisión fluida de Microsoft™; transmisión dinámica adaptativa a través de HTTP (DASH)), archivos de vídeo como MP4, AVI, archivos de vídeo de Windows Media, contenido de audio como contenido de audio transmitido o archivos de audio discretos (como archivos MP3).

El experto en la técnica apreciará que una característica de cualquier aspecto de la invención se puede aplicar, mutatis mutandis, a cualquier otro aspecto de la invención.

30 Además, el experto en la técnica apreciará que los elementos de los aspectos pueden proporcionarse en software. Sin embargo, los expertos también apreciarán que cualquier elemento de software puede proporcionarse en firmware y/o en hardware, o viceversa.

35 El medio legible por máquina al que se hace referencia en cualquiera de los aspectos anteriores de la invención puede ser cualquiera de los siguientes: un CDROM; un DVD ROM/RAM (incluidos -R/-RW o +R/+RW); un disco duro; una memoria (que incluye una unidad USB); una tarjeta SD; una tarjeta Compact Flash o similar); una señal transmitida (que incluye una descarga por Internet, una transferencia de archivos FTP, etc.); un cable; etc.

40 A continuación sigue, a modo de ejemplo únicamente, una descripción detallada de las realizaciones de la invención con referencia a los dibujos adjuntos de los cuales:

**la Figura 1** muestra esquemáticamente un sistema de procesamiento dispuesto para llevar a cabo una realización de la invención;

45 **la Figura 2** muestra esquemáticamente un sistema dispuesto para generar huellas digitales;

**la Figura 3** muestra un diagrama que describe los procesos y almacenes de datos llevados a cabo por una realización de un sistema de monitoreo;

50 **la Figura 4** muestra un detalle adicional de una porción del sistema de monitoreo descrito con referencia a la Figura 3;

**la Figura 5** muestra un detalle adicional de una porción del sistema de monitoreo descrito con referencia a la Figura 4;

55 **la Figura 6 (técnica anterior)** muestra los paquetes de red que se pasan a través de una red;

**la Figura 7** muestra esquemáticamente cómo se generan las huellas digitales;

60 **la Figura 8** muestra un diagrama de flujo que describe un proceso de generación de huellas digitales;

**la Figura 9** muestra un diagrama de flujo que describe un proceso para generar una medición de calidad asociada con una huella digital;

65 **la Figura 10** muestra un diagrama de flujo que describe un proceso para determinar si una huella digital de referencia coincide con una huella digital candidata;

**la Figura 11a** ejemplifica un proceso para determinar si una huella digital candidata coincide con una huella digital de referencia;

5 **la Figura 11b** ejemplifica una serie de huellas digitales coincidentes dentro de una serie de huellas digitales candidatas;

**la Figura 12** ejemplifica la naturaleza superpuesta de las huellas digitales creadas por el sistema;

10 **la Figura 13** muestra un diagrama que destaca un flujo de proceso de una realización adicional del sistema descrito en la Figura 2;

**la Figura 14** muestra esquemáticamente un sistema dispuesto para entregar flujos multimedia a los usuarios;

15 **la Figura 15** ejemplifica cómo se alinean dos flujos multimedia entre sí; y

**la Figura 16** muestra un diagrama de flujo del procesamiento de contenido multimedia por las realizaciones.

20 La siguiente descripción proporciona una descripción de varias realizaciones y el experto en la técnica apreciará fácilmente que una característica descrita en relación con una realización dada puede aplicarse, cambiando lo que deba cambiarse, a cualquiera de las otras realizaciones.

25 El sistema informático 100 de la Figura 1 ejemplifica un sistema informático que puede utilizarse para proporcionar los métodos implementados por ordenador descritos en la presente memoria o como un sistema informático descrito en la presente memoria. El sistema informático 100 comprende una pantalla 102, un circuito de procesamiento 104, un teclado 106 y un ratón 108. El circuito de procesamiento 104 comprende una unidad de procesamiento 112, un sistema de gráficos 113, un disco duro 114, una memoria 116, un subsistema de E/S 118 y un bus de sistema 120. La unidad de procesamiento 112, el sistema de gráficos 113, el disco duro 114, la memoria 116 y el subsistema de E/S 118 se comunican entre sí a través del bus de sistema 120, que en esta realización es un bus PCI, de una manera bien conocida en la técnica.

30 El sistema de gráficos 113 podría comprender un procesador de gráficos dedicado dispuesto para realizar parte del procesamiento de los datos que se desea mostrar en la pantalla 102. Dichos sistemas gráficos 113 son bien conocidos y aumentan el rendimiento del sistema informático eliminando parte del procesamiento requerido para generar una pantalla de la unidad de procesamiento 112.

35 Se apreciará que, aunque se hace referencia a una memoria 116, es posible que la memoria pueda ser proporcionada por una variedad de dispositivos. Por ejemplo, la memoria puede ser proporcionada por una memoria caché, una memoria RAM, un dispositivo de almacenamiento masivo local tal como el disco duro 114, cualquiera de estos conectados al circuito de procesamiento 104 a través de una conexión de red. Sin embargo, la unidad de procesamiento 112 puede acceder a la memoria a través del bus de sistema 120 para acceder al código de programa para indicarle qué pasos realizar y también para acceder a los datos a procesar. La unidad de procesamiento 112 se dispone para procesar los datos como se indica en el código de programa.

40 De hecho, en algunas realizaciones es totalmente posible que varios sistemas informáticos 100, circuitos de procesamiento 104 y/o unidades de procesamiento 112 puedan conectarse en paralelo y/o distribuirse a través de una red, para proporcionar el método y/o los sistemas informáticos descritos en la presente memoria.

45 En la Figura 1 se muestra un diagrama esquemático de la memoria 114, 116 del sistema informático. Puede verse que la memoria comprende una porción de almacenamiento de programas 122 dedicada a almacenar programas y una porción de almacenamiento de datos 124 dedicada a almacenar datos, pero no es necesario que la memoria esté estructurada de esta manera.

50 En la realización que se describe, la porción de almacenamiento de programas 122 comprende al menos algo de lo siguiente: un generador de ventana 126, un comparador 128, un generador de huellas digitales 130, un evaluador de calidad 132 y un reductor de resolución 134 como se describe a continuación. Resultará evidente a partir de lo siguiente que algunos de los circuitos de procesamiento descritos pueden comprender solo algunos de los elementos mostrados en relación con la Figura 1.

55 Pasando a la Figura 2, se muestra una descripción general de un sistema de generación de huellas digitales 200. Es conveniente describir el sistema 200 en términos de monitoreo, identificación y/o procesamiento de componentes multimedia, tales como multimedia, que incluye video, que puede o no proporcionarse como un flujo y, en particular, puede proporcionarse como un flujo de Internet. Sin embargo, las realizaciones no necesitan ser tan limitadas.

60 El sistema 200 comprende al menos uno, y típicamente varios generadores de huellas digitales de referencia 202a, 202b, 202c, 202d. Cada uno de los generadores de huellas digitales de referencia 202a-202d puede comprender un

circuito de procesamiento y tener una arquitectura como se describe en relación con la Figura 1. Es concebible que los generadores de huellas digitales de referencia 202a-d se proporcionen como máquinas virtuales pero típicamente las realizaciones, incluida la realización que se describe, proporcionarán los generadores de huellas digitales de referencia como dispositivos de hardware que están asociados con los flujos multimedia de referencia 204a-d.

Además, cada uno de los generadores de huellas digitales de referencia 202a-202d ha introducido en el mismo al menos un flujo multimedia de referencia 204a, 204b, 204c, 204d. Este flujo multimedia de referencia 204a-d puede considerarse como un componente multimedia de referencia. Sin embargo, las realizaciones del generador de huellas digitales de referencia pueden no tener pantallas 102 o, de hecho, pueden no contener sistemas de gráficos 113.

Normalmente, los generadores de huellas digitales de referencia 202a-d están situados, en el caso del monitoreo de flujos multimedia, dentro de una organización generadora de contenido multimedia, como una emisora o similar. El flujo multimedia de referencia 204a-d se dispone típicamente de modo que sea una T fuera de un material de transmisión para no interferir con la transmisión.

Cada uno de los generadores de huellas digitales de referencia 202a-d funciona como se describe a continuación y genera huellas digitales del flujo multimedia de referencia 204a-204d para generar una secuencia de huellas digitales de referencia RFPn (donde se pretende que n sea un número entero). El generador de huellas digitales de referencia 202a-d se dispone entonces para enviar las huellas digitales de referencia generadas (es decir, la secuencia de huellas digitales de referencia) a un circuito de procesamiento adicional 206 que aloja un sistema para procesar las huellas digitales de referencia RFPn recibidas. Por tanto, los generadores de huellas digitales de referencia 202a-202d están conectados, típicamente a través de una Red de Área Amplia, tal como Internet, al circuito de procesamiento 206 a través de las conexiones 208a, 208b, 208c, 208d. El circuito de procesamiento 206 también puede tener una arquitectura similar a la que se muestra en la Figura 1 (tal como una arquitectura X86 o similar), un servidor virtual, una pluralidad de tales máquinas/máquinas virtuales, o similares.

Como se describe con más detalle a continuación, el circuito de procesamiento 206 procesa las huellas digitales de referencia RFPn recibidas dentro de la secuencia y las almacena dentro de un almacén 210 que proporciona así un almacén de huellas digitales de referencia. El almacén 210 puede considerarse como un almacén de activos. En la realización que se describe, el almacén 210 está conectado al circuito de procesamiento 206 a través de una conexión de red 212, que puede ser una conexión de Red de Área Amplia tal como Internet. Por tanto, en la realización que se describe, el circuito de procesamiento 206 y el almacén 210 se proporcionan en dos máquinas separadas. Sin embargo, el experto en la técnica apreciará que en otras realizaciones, el circuito de procesamiento 206 y el almacén 210 pueden proporcionarse dentro de una sola máquina o más de dos máquinas.

Conectado al almacén 210, a través de las conexiones de red 214a, 214b, 214c, 214d, que de nuevo pueden ser conexiones de Red de Área Amplia como Internet, hay al menos uno y típicamente una pluralidad de mecanismos de procesamiento de identificación 216a, 216b, 216c, 216d. En la realización que se describe, hay cuatro de tales mecanismos de procesamiento de identificación 216a-d, pero en otras realizaciones pueden ser cualquier otro número. Es una ventaja de algunas realizaciones que es altamente escalable ya que se pueden añadir más generadores de huellas digitales 202 y/o mecanismos de procesamiento de identificación 216 al sistema 200 para aumentar su rendimiento. Los mecanismos de procesamiento de identificación 216 pueden ser proporcionados por máquinas que tengan una arquitectura similar a la de la Figura 1 o pueden proporcionarse dentro de un entorno virtual tal como un servidor virtual o similar en un entorno de nube. Al igual que con los generadores de huellas digitales 202 de referencia, los mecanismos de procesamiento de identificación 216 pueden no estar provistos de una pantalla 102 y/o sistemas de gráficos 113.

Un sistema de monitoreo 218 se dispone para monitorear y/o identificar los componentes multimedia existentes en las fuentes de datos accesibles por el sistema. En la realización que se describe, los componentes multimedia son flujos multimedia, tales como videos y pueden ser flujos de Internet. En la realización que se describe, estos se encuentran a través de enlaces en la World Wide Web (WWW) y, por lo tanto, las fuentes de datos son proporcionadas por sitios web, foros de usuarios, salas de chat o similares que son accesibles a través de WWW. En otras realizaciones, los datos que se monitorizan pueden ser distintos de los flujos multimedia y, en particular, pueden ser medios no transmitidos, como un archivo de audio (por ejemplo, mp3; wav; o similar) o un archivo de vídeo (tal como MP4; WMV; o similares). Por tanto, el sistema de monitoreo 218 se dispone para recibir un componente multimedia candidato que se dispone para ser procesado como se describe en la presente memoria.

En realizaciones alternativas o adicionales, el sistema puede disponerse para monitorear fuentes de componentes multimedia a través de una red distinta a la WWW. Por ejemplo, la red 251 puede ser Internet (y no WWW) y las fuentes de datos pueden ser proporcionadas por sitios FTP (Protocolo de transferencia de archivos) o similares.

El sistema de monitoreo es proporcionado típicamente por una pluralidad de circuitos de procesamiento, que pueden ser máquinas físicas, máquinas virtuales o una combinación de ambas.

En la realización que se describe, el sistema de monitoreo 218 también se dispone para generar secuencias de huellas digitales candidatas CFPn (donde se pretende que la n sea un número entero) para lo que puede denominarse uno o

más componentes multimedia candidatos en los que se dispone el sistema 200 a monitorear. Típicamente, se genera una secuencia de huellas digitales candidatas CFPn para cada uno de los componentes multimedia candidatos. En otras realizaciones, el sistema puede no generar las huellas digitales candidatas CFPn.

5 En la realización que se describe, el componente multimedia candidato es un flujo multimedia y, en particular, puede ser un video y, en particular, puede ser un flujo de Internet. Por tanto, en esta realización, un mecanismo de procesamiento de identificación 216a-c se dispone para procesar una secuencia de huellas digitales candidatas CFPn de un flujo multimedia de entrada 219.

10 La, o cada, secuencia de huellas digitales candidatas CFPn generadas por el sistema de monitoreo 218 se colocan en cola en una cola 220 para ser analizadas por el sistema 200. A continuación, el sistema 200 se dispone de modo que una de las secuencias de huellas digitales candidatas CFPn desde dentro de la cola 220 se asigne a un mecanismo de procesamiento de identificación 216a-d que carga la secuencia para su procesamiento adicional. En la realización que se describe, las secuencias de huellas digitales cargadas en un mecanismo de procesamiento de identificación 216a-d en un momento dado son típicamente representativas de un minuto de flujo multimedia. Una transmisión de un minuto, que proporciona un período relativamente corto, es particularmente conveniente para las realizaciones que procesan multimedia transmitidos, quizás los denominados multimedia transmitidos en vivo.

15 Algunas realizaciones pueden disponerse para procesar un componente multimedia completo, o al menos sustancialmente todo un componente multimedia (por ejemplo, un archivo multimedia) y, por lo tanto, generar huellas digitales candidatas.

20 Ese mecanismo de procesamiento de identificación se dispone entonces para realizar una búsqueda de cada una de las huellas digitales candidatas CFPn dentro de esa secuencia para ver si esa huella digital CFPn corresponde (es decir, coincide) con alguna de las huellas digitales de referencia RFPn almacenadas dentro del sistema 200. Este proceso de búsqueda se describe más adelante.

25 Adicional, o alternativamente, a las realizaciones de los mecanismos de procesamiento de identificación se les puede asignar una única huella digital CFPn de la cola 220.

30 La Figura 3 muestra una ampliación del sistema de monitoreo 218 de la Figura 2 y ahora se explica con más detalle.

35 De manera similar a los mecanismos de procesamiento de identificación 216a-d mostrados en la Figura 2, el sistema de monitorización 218 comprende una pluralidad de mecanismos de procesamiento adaptables 250a-d. En esta realización, solo se muestran cuatro mecanismos de procesamiento personalizables, pero en realidad podría haber un número significativo más. Cada uno de los mecanismos de procesamiento adaptables está conectado a una red 251.

40 Cada uno de los mecanismos de procesamiento adaptables podría ser una máquina física, como se ejemplifica en relación con la Figura 1. En caso de que los mecanismos de procesamiento adaptables se proporcionen como máquinas físicas, es posible entonces que estos no comprendan una pantalla o controlador de pantalla, etc. En realizaciones alternativas o adicionales, incluida la realización que se describe, los mecanismos de procesamiento adaptables pueden proporcionarse como una máquina virtual.

45 Cada uno de los mecanismos de procesamiento adaptables está controlado por una interfaz de programación de aplicaciones (API) 252 del sistema y, por lo tanto, el sistema se dispone para controlar la funcionalidad de los mecanismos de procesamiento adaptables 250. También se verá que la API se dispone para generar la cola 220 descrita anteriormente.

50 Un portal de usuario 254 permite que la entrada de usuario 256 se realice en la API 252, lo que permite controlar la funcionalidad de los mecanismos de procesamiento adaptables 250.

55 La Figura 4 amplía el flujo de proceso del sistema descrito con referencia a las Figuras 2 y 3. Cabe señalar que la Figura 4 es un diagrama de flujo en lugar de un diagrama estructural, pero la descripción a continuación relaciona aspectos de la Figura 4 con los diagramas de sistema de las Figuras 2 y 3.

60 El sistema 218 mantiene datos de términos que pueden usarse para localizar uno o más componentes multimedia candidatos para su análisis por parte del sistema. Como tal, se mantiene un almacén de datos de términos 400 para su procesamiento como se describe más adelante. Los datos de términos almacenados dentro del almacén 400 pueden comprender URL (Localizador de Recursos Universal), términos de búsqueda o similares, cualquiera de los cuales puede usarse para localizar componentes multimedia en la red 251.

Los términos, URL u otros datos de términos pueden introducirse en el almacén de datos de términos 400 a través del portal de usuario 254.

65 Los mecanismos de procesamiento adaptables 250 mostrados en la Figura 3 se disponen para actuar como arañas 402 para recopilar información de la red 251 de fuentes de datos dadas en la red, que en la realización que se describe

son típicamente páginas web. En este proceso, la API 252 asigna un mecanismo de procesamiento adaptable 250 dado a una fuente de datos para procesar, dicha fuente de datos puede ser proporcionada por los datos de términos proporcionados al mecanismo de procesamiento adaptable 250.

5 Por ejemplo, si los datos de términos fueran un término de búsqueda, entonces el mecanismo de procesamiento adaptable 250 se dispone para buscar en la red, o al menos un sitio web dado, o una página de un sitio web para ese término de búsqueda. Si los datos de términos fueran una URL, entonces el mecanismo de procesamiento adaptable 250 se dispone para procesar esa URL. Los datos de términos pueden incluir especificar qué porciones de la página web, sitio o similares deben buscarse. En particular, algunas realizaciones pueden especificar etiquetas, parámetros, 10 tales como etiquetas XML o parámetros que deberían buscarse.

La función de rastreo en la que los mecanismos de procesamiento adaptables 250 se disponen para proporcionar búsquedas en la fuente de datos asignada al mecanismo de procesamiento adaptable 250 por la API 252 y devuelve 15 datos que podrían coincidir con criterios especificados dados por los datos de términos.

En la realización que se describe, los criterios especificados serían un flujo multimedia que podría coincidir potencialmente con un flujo multimedia para el que el almacén 210 tiene una huella digital de referencia RFPn representativa de ese flujo multimedia. Sin embargo, adicional o alternativamente, el mecanismo de procesamiento adaptable 250 puede disponerse para capturar otras formas de datos, incluidos los componentes multimedia, como 20 cualquier forma de archivo multimedia, la fuente de datos sospechosa de contener un componente multimedia, o similares.

El mecanismo de procesamiento adaptable 250 que proporciona la función de rastreo 402 se dispone para cargar los resultados de las tareas que se les ha asignado realizar en un almacén de resultados de rastreo 404. 25

Los resultados de rastreo guardados dentro del almacén de resultados de rastreo 404 son procesados por un motor de verificación 406 que también se dispone para verificar los datos enviados directamente a través del portal de usuario 256 que se envían directamente al motor de verificación en lugar de ser localizados por la función de rastreo 402 de los mecanismos de procesamiento adaptables 250. En particular, el motor de verificación se dispone, al menos en algunas realizaciones, para puntuar metadatos y devolver una creencia (que puede considerarse un factor de 30 importancia) en esos metadatos. Algunas realizaciones del sistema, incluida la que se está describiendo, pueden usar entonces esa creencia en cómo se debe usar cualquier determinación de esos metadatos para determinar una programación basada en esos metadatos. Aquí los metadatos pueden comprender etiquetas de texto, XML o HTML, o similares que han sido sometidas a web scraping, por la araña 402, desde la fuente de datos (por ejemplo, página web) que ha procesado. 35

En la realización que se describe, los resultados verificados del rastreo 404 /entrada de información directamente al sistema se almacenan dentro de un almacén de páginas capturadas 408 para su procesamiento posterior. La referencia aquí a un almacén de páginas capturadas 408 es adecuada para realizaciones en las que la red 251 es la 40 World Wide Web. Sin embargo, en otras realizaciones, el almacén puede denominarse más adecuadamente un almacén de protodatos capturados; es decir, datos que puedan ser pertinentes y que el sistema debería seguir procesando.

Es posible que los datos se ingresen directamente en el almacén de páginas capturadas 408 y esto se representa como la entrada 411 en la Figura. Típicamente, estos datos se introducirían a través del Portal de Usuario 254. 45

De nuevo, es posible que los protodatos puedan introducirse directamente en el almacén de protodatos, que en la realización que se describe pueden considerarse como páginas precargadas 410. Estas páginas precargadas (u otros repositorios de datos) pueden ser páginas que habitualmente alojan datos, como transmisiones de multimedia u otros 50 datos multimedia, que cumplen los criterios especificados. En la realización que se describe, el sistema se dispone para monitorear e identificar los flujos multimedia que son copias de los flujos multimedia de referencia 204 y, como tal, el criterio especificado podría ser que esas páginas precargadas habitualmente alojen flujos multimedia que sean copias de los flujos multimedia de referencia 204. En otras realizaciones, el sistema se dispone para monitorear más que el flujo multimedia y puede disponerse para monitorear o identificar de otra manera cualquier forma de archivo multimedia, o de hecho otras formas de datos. 55

Los protodatos, que en esta realización son las páginas capturadas y/o precargadas, se procesan entonces mediante los elementos de captura multimedia 412 que pueden proporcionarse mediante los mecanismos de procesamiento personalizables 250a-d. Una vez más, estos elementos 412 de captura multimedia están controlados por la API 252, 60 que dirige y programa la actividad de los elementos de captura multimedia 412. Por tanto, en la realización que se está describiendo, el mismo mecanismo de procesamiento adaptable 250 se dispone para funcionar tanto como la araña 402 o como el elemento de captura de multimedia 412. Se cree que una disposición de este tipo en la que el mecanismo de procesamiento adaptable 250 tiene una función dual es ventajosa porque puede conducir a un uso eficiente de los recursos dentro del sistema. Otras realizaciones pueden proporcionar la araña 402 y los elementos de 65 captura de multimedia 412 como entidades separadas.

En la realización que se describe, los elementos de captura de multimedia 412 se disponen para procesar flujos multimedia candidatos 219, ubicados dentro de las fuentes de datos (es decir, páginas web), y generar huellas digitales candidatas CFPn y, por lo tanto, el elemento de captura de multimedia 412 proporciona un generador de huella digital candidata. Estas huellas digitales candidatas CFPn se almacenan en la cola 220 para su posterior análisis por parte del sistema.

Además de la generación de las huellas digitales candidatas CFPn, el elemento de captura de multimedia 412 se dispone para almacenar una copia de al menos una porción del componente multimedia candidato dentro de un almacén de contenido multimedia candidato 413. Por ejemplo, si el componente multimedia candidato fuera un flujo de video, entonces el elemento de captura de multimedia 412 puede disponerse para almacenar uno o más iconos que representan el flujo de video. Aquí, un icono pretende significar una versión en la que la o cada trama tiene una resolución, contenido de color, frecuencia de tramas o similares reducidos, con el fin de reducir la cantidad de espacio necesario para almacenar la miniatura en comparación con el componente multimedia original.

Si el componente multimedia candidato fuera un archivo de audio, entonces el elemento de captura de multimedia puede disponerse para almacenar una versión del audio del archivo de calidad reducida, longitud reducida, etc.

En la realización que se describe, el elemento de captura de multimedia 412 también se dispone para almacenar, en el almacén de contenido multimedia candidato 413, metadatos asociados con el componente multimedia candidato. Los metadatos almacenados en el medio candidato 413 para cada componente multimedia candidato pueden incluir cualquiera de los siguientes elementos: la URL que apuntaba al componente multimedia (por ejemplo, el flujo multimedia); la dirección IP de la que se obtuvo el componente multimedia (por ejemplo, el flujo multimedia); el tiempo en el que se obtuvo el componente multimedia (por ejemplo, el flujo multimedia); o similares. Estos metadatos pueden almacenarse además como datos de ejecución 414, al menos en la realización que se describe, que especifica la fuente a partir de la cual se obtuvo el componente multimedia (por ejemplo, flujo multimedia).

La funcionalidad del elemento de captura de multimedia se describe a continuación en relación con las Figuras 5 y 6.

También es posible que las huellas digitales candidatas CFPn se puedan agregar a la cola desde una fuente externa 224. En algunas realizaciones, puede haber múltiples fuentes externas desde las cuales se pueden añadir huellas digitales candidatas CFPn a la cola 220. El portal de usuarios 254 puede proporcionar un ejemplo de la fuente externa.

Una vez que las huellas digitales candidatas CFPn se han añadido a la cola 220, estas son procesadas por los mecanismos de procesamiento de identificación 216 bajo la dirección de la API 252. Los mecanismos de procesamiento de identificación 216a-d hacen comparaciones con las Huellas digitales de Referencia RFPn mantenidas dentro del almacén 210 como se describe en otra parte.

En la realización que se describe, además del proceso de identificación/monitoreo automatizado que se describe a continuación, los usuarios pueden inspeccionar manualmente los componentes multimedia guardados en el almacén de contenido multimedia candidato 413. Este proceso se realiza a través del portal de usuario 254 y se muestra como proceso 409. En este proceso, se presenta al usuario, en una pantalla de video, una copia del componente multimedia. En la realización que se describe, al usuario se le presenta la miniatura del componente multimedia candidato que ha sido generado por el elemento de captura de multimedia 412.

En la Figura 3 se muestra una unidad de verificación manual 260 que puede proporcionarse en algunas realizaciones y que permite al usuario proporcionar el proceso 409. La unidad de verificación manual 260 proporciona al usuario una entrada al sistema. Esta entrada también puede ser a través del portal de usuarios 254. Sin embargo, en otras realizaciones, puede que no se proporcione la unidad de verificación manual 260 y el sistema puede funcionar automáticamente.

Si un componente multimedia candidato se identifica como uno que necesita más acción, entonces cualquier metadato asociado con el componente multimedia es capturado por el elemento de captura de multimedia 412 y asociado con ese componente multimedia candidato se transfiere al almacén de datos de ejecución 414 para la acción como se describe a continuación. Por tanto, una ventaja de tales realizaciones es que los metadatos se capturan para el flujo multimedia candidato 219 que luego se pueden usar para automatizar procesos, tales como la generación de avisos de eliminación, en relación con ese flujo multimedia 219. Esta captura de metadatos se puede lograr en realizaciones que identifican automáticamente un componente multimedia así como aquellos con una identificación manual.

La araña 402 y el elemento de captura de multimedia 412 típicamente serán proporcionados por máquinas, ya sean virtuales o físicas o una combinación de las mismas, que están dentro, conectadas directamente, o similares, a la red 251. Se apreciará que tales realizaciones son ventajosas ya que la araña 402 y los elementos de captura de multimedia 412 propician un acceso regular a la red 251 y por lo tanto la proximidad a la red 251 es ventajosa porque reducirá el tráfico de la red.

Además, al menos algunas realizaciones, incluida la que se describe, disponen los mecanismos de procesamiento de identificación 216 de manera que estén separados de la red 251. Tal separación puede ser proporcionada por un

## ES 2 983 243 T3

- 5 cortafuegos, una zona desmilitarizada (DMZ) o similar. Se apreciará que en algunas realizaciones, incluyendo las descritas, los mecanismos de procesamiento de identificación 216 están conectados más estrechamente al sistema 200 en comparación con la araña 402 o el elemento de captura de multimedia 412 y, como tal, es ventajoso separarlos de la red 251 para aumentar la seguridad del sistema 200. Por tanto, las realizaciones que estructuran los mecanismos de procesamiento de identificación 216 de modo que estén protegidos de la red 251 abordan la seguridad del sistema.
- 10 Una vez que los mecanismos de procesamiento de identificación 216 han identificado las huellas digitales candidatas CFPn como una coincidencia con una huella digital de referencia RFPn mantenida dentro del almacén 210, en la realización que se describe, estas se disponen para cargar, o transmitir de otro modo, detalles del componente multimedia (por ejemplo, el flujo multimedia en esta realización) a un almacén de datos de ejecución 414, que incluye metadatos asociados con el componente multimedia candidato.
- 15 En al menos algunas realizaciones, es posible añadir datos de ejecución directamente al almacén 414 y esto está representado por el elemento 415 en la Figura. Dichos datos se pueden agregar a través del portal de usuarios 254.
- 20 Pueden disponerse otras realizaciones, alternativas o adicionales, para analizar datos y tal análisis puede proporcionar inteligencia comercial o información en relación con los componentes multimedia que se están monitoreando. En tales realizaciones, los mecanismos de procesamiento de identificación 216 pueden disponerse para almacenar los datos identificados por el sistema para un componente multimedia candidato para un análisis posterior. Típicamente, los datos identificados por el sistema se almacenan con referencia al componente multimedia candidato que dio lugar a esos datos.
- 25 En la realización que se describe, se proporciona un proceso de eliminación 416 adicional que se dispone para procesar los datos almacenados en el almacén de datos de ejecución 414 y enviar avisos de eliminación a los destinatarios especificados por los datos de ejecución. En esta realización, esos avisos se envían por correo electrónico, pero en realizaciones alternativas o adicionales, los avisos de eliminación pueden enviarse por correo postal; fax; Mensajería instantánea; SMS; MMS; o similares.
- 30 Las realizaciones del sistema también pueden disponerse para continuar monitoreando la fuente de datos (que en la realización que se describe pueden ser enlaces o similares) que dio lugar a los datos de ejecución que dieron lugar a un aviso de eliminación. Este monitoreo continuo puede ocurrir de vez en cuando, que puede ser periódicamente a intervalos predeterminados. Por ejemplo, la araña 402 y/o el elemento de captura de multimedia 412 pueden disponerse para monitorear una fuente de datos sustancialmente cada 10 minutos, 15 minutos, 20 minutos, 30 minutos o similar.
- 35 Si el sistema determina que el componente multimedia que provocó la generación del aviso de eliminación todavía está disponible en la fuente del componente multimedia, entonces se puede escalar aún más el proceso de eliminación. Por ejemplo, en lugar de enviar un correo electrónico a un sitio web en particular, se puede enviar un aviso de eliminación directamente a un ISP o similar.
- 40 Las realizaciones, incluida la que se describe, pueden disponerse de modo que la araña 402 y/o los elementos de captura de multimedia 412, de vez en cuando, realicen la función que se les ha asignado; independientemente de si esto se refiere a los datos de ejecución.
- 45 En la realización que se describe, la API 252 puede indicar a la araña 402 que repita el proceso que se le asignó en los datos de términos. La repetición se puede realizar periódicamente a intervalos regulares tales como sustancialmente cada 10 minutos, 15 minutos, 20 minutos, 30 minutos o similares. Pueden disponerse otras realizaciones para repetir el proceso a intervalos irregulares.
- 50 De hecho, como se describió anteriormente, algunas realizaciones pueden organizarse para obtener metadatos de la captura de arañas realizada por la araña 402; es decir, datos obtenidos de la fuente de datos sobre lo que se conserva en la fuente de datos. Tales realizaciones pueden disponerse además para usar los metadatos para establecer un horario en cuanto a cuándo la araña 402 debe repetir el proceso que se le ha asignado.
- 55 Además, el elemento de captura de multimedia 412 también puede disponerse para generar huellas digitales candidatas CFPn a partir de los protodatos (es decir, las páginas capturadas) de vez en cuando. La generación de huellas digitales candidatas CFPn puede lograrse recapturando la página (es decir, la fuente de datos) en la que se encuentra el componente multimedia candidato. La repetición puede realizarse periódicamente, a intervalos regulares, como sustancialmente cada 10 minutos, 15 minutos, 20 minutos, 30 minutos o similares. Pueden disponerse otras realizaciones para repetir el proceso a intervalos irregulares.
- 60 En algunas realizaciones, el sistema puede omitir el almacén de datos de ejecución 414 y el proceso de eliminación 416 y simplemente estar dispuesto para generar datos de uso.
- 65 La Figura 5 muestra más detalles de los procesos que se ejecutan dentro del mecanismo de procesamiento adaptable 250 y, en particular, el elemento de captura de multimedia 412 que puede ser proporcionado por el mecanismo de

procesamiento adaptable 250. También se hace referencia a la Figura 6, que se utiliza para ilustrar los paquetes de datos analizados en relación con la Figura 4.

5 Como se describió anteriormente, el elemento de captura de multimedia 412, en la realización que se está describiendo, captura un flujo de multimedia 219 que ha sido identificado por el sistema de monitoreo 218 que necesita investigación adicional. En otras realizaciones, el elemento de captura de multimedia 412 puede disponerse para capturar otros componentes multimedia distintos de los contenidos multimedia transmitidos.

10 Para capturar los componentes multimedia, y en la realización que se describe, el flujo multimedia 219, el elemento de captura de multimedia 412 comprende un primer mecanismo de procesamiento 550. El primer mecanismo de procesamiento funciona como un servidor de multimedia y recibe el componente multimedia entrante (en esta realización, el flujo multimedia 219) y retransmite ese componente multimedia al segundo mecanismo de procesamiento 552.

15 En la realización que se describe, el primer y segundo mecanismos de procesamiento son proporcionados por máquinas virtuales que proporcionan un circuito de procesamiento simulado 104. El experto en la técnica apreciará que el primer y/o segundo mecanismos de procesamiento 550, 552 pueden proporcionarse como máquinas físicas. Sin embargo, en la realización que se describe, proporcionar el primer y segundo procesamiento como máquinas virtuales permite flexibilidad en cuanto a la ubicación del sistema de monitoreo 218 y simplifica la creación de mecanismos de procesamiento 250 adaptables adicionales, que proporcionan los elementos de captura de multimedia 412 para expandir el sistema de monitoreo 218. Por tanto, el primer 550 y el segundo 552 mecanismos de procesamiento, en la realización que se describe, se proporcionan dentro de una única máquina virtual.

25 En la realización que se describe, la red 251 es la World Wide Web (WWW) que se ejecuta por encima de Internet. Por supuesto, otras realizaciones podrían utilizar otras arquitecturas. Sin embargo, la Figura 6 ejemplifica las cuatro capas de una red WWW, como la que se puede utilizar para la red 251. Existe la capa de enlace que es específica del hardware que proporciona la conexión y la información de encabezado de la capa de enlace no se monitorea en la realización que se describe. Sin embargo, es concebible que en otras realizaciones los paquetes de datos de la capa de enlace puedan ser monitoreados y/o capturados.

30 Sin embargo, el elemento de captura de multimedia 412, y en particular su mecanismo de captura de datos 554, se dispone para capturar la información de encabezado de cada una de las capas de Internet, Transporte y Aplicación. Por tanto, se apreciará que los datos, que en la realización que se describe, son uno o más paquetes de datos de multimedia, están rodeados por más información de encabezado en las capas inferiores. El experto en la técnica apreciará que en algunos protocolos la capa de aplicación contendrá información de encabezado y, por lo tanto, el bloque de datos (representado por el bloque sombreado en la Figura), en todas las capas del diagrama, de hecho puede contener tantos datos como información de encabezado de la capa de aplicación.

40 Los datos contenidos en la información del encabezado son útiles para determinar la fuente de un componente multimedia de los datos y se puede considerar que proporcionan metadatos asociados con los contenidos multimedia (u otros datos) que se transmiten a través de la red. El experto en la técnica apreciará que los sitios web, u otras fuentes de datos en redes distintas de la WWW, pueden redirigir a los usuarios a otras fuentes sin que el usuario se dé cuenta. La inspección de la información del encabezado se puede utilizar para determinar si esto ha ocurrido. Previamente, ha sido un problema comprender la última fuente de un componente multimedia, por ejemplo, un flujo multimedia, un archivo multimedia, etc., y la inspección de la información del encabezado como se usa en la realización que se describe puede ser utilizada por el proceso de eliminación 416 para enviar avisos a la última fuente según se encuentra en la información del encabezado, registrar la identidad de la última fuente, etc.

50 En la realización que se describe, el componente multimedia puede transmitirse en un formato de flujo de red que se procesa para determinar las secciones del flujo que están asociadas con el control y la configuración del flujo; y los relacionados con el contenido multimedia. A partir de los datos de configuración de flujo y control, la sintaxis específica del protocolo utilizada en el establecimiento del flujo multimedia se guarda y se usa para agregar múltiples fuentes del mismo contenido multimedia incorporadas en diferentes páginas web. Aquí, el experto en la técnica apreciará que una pluralidad de sitios web pueden proporcionar acceso al mismo flujo, donde se puede pensar que el flujo es la última fuente del componente multimedia. Como tal, puede ser ventajoso agregar las fuentes de los componentes multimedia de manera que sea más evidente cuán ampliamente se está distribuyendo esa fuente de componentes multimedia.

60 El contenido del flujo real se guarda en un formato de archivo y forma el contenido multimedia candidato que se procesa para generar huellas digitales candidatas como se describe en otra parte.

El experto también apreciará que existen tecnologías que entregan contenido multimedia, y en particular contenidos multimedia transmitidos, de manera fragmentada. Los ejemplos de dicha tecnología incluyen HTTP Live Streaming (HLS), Microsoft Smooth Streaming, Adobe™ HDS (Adobe Dynamic Streaming) o mecanismos personalizados. Las realizaciones que vuelven a transmitir los componentes multimedia como se describe en relación con la Figura 5 son ventajosas ya que permiten que tales tecnologías que entregan contenido de una manera fragmentada se procesen y

que el contenido multimedia se guarde en un formato de archivo y se procese posteriormente. Puede ser un problema procesar y tomar huellas digitales de estos flujos fragmentados en vista de la corta duración de los fragmentos.

5 Algunas realizaciones del sistema pueden disponerse para procesar componentes multimedia que no corresponden a un formato conocido. Algunas realizaciones del sistema pueden volver a combinar archivos de audio y video de flujos/archivos separados. Esta funcionalidad surge típicamente del procesamiento de la información de encabezado de las diversas capas del tráfico de la red.

10 En la realización que se describe, el mecanismo de captura de datos 554 se dispone para reconstruir los componentes multimedia entregados de una manera tan fragmentada. En particular, el mecanismo de captura de datos 554 se dispone para analizar la información de encabezado de al menos algunas de las capas y usar esa información de encabezado para reconstruir los paquetes de datos (que en la realización que se describe proporcionan los contenidos multimedia) en un formato no fragmentado; es decir, para guardar el contenido en formato de archivo. Así, en la realización que se describe, el elemento de captura de datos 554 se dispone para procesar la información de encabezado de las diversas capas y para detectar un flujo de video HLS (u otro fragmentado), para capturar los datos, 15 que en este caso serán tramas de video, de los diversos paquetes de red y unir (es decir, agregar) los datos para formar un archivo multimedia que proporcione una secuencia multimedia continua. A continuación, el archivo multimedia se puede procesar, por ejemplo, tomar las huellas digitales, como se describe en otra parte.

20 Para generar una huella digital de referencia RFPn, cada uno de los generadores de huellas digitales de referencia 202a-d procesa el flujo multimedia de referencia 204a-d (es decir, el componente multimedia de referencia) como se describe ahora. Se apreciará que la técnica descrita para generar huellas digitales se proporciona como ejemplo y otras realizaciones pueden utilizar otras técnicas de toma de huellas digitales.

25 El proceso también se ejemplifica en relación con las Figuras 7 y 8. La Figura 7 representa el flujo multimedia entrante como una secuencia de tramas; por ejemplo, 300a, 300b, 300c, etc. (etapa 800). Cada trama del flujo multimedia puede considerarse como un elemento de datos del componente multimedia. Otras realizaciones pueden procesar otras formas de datos.

30 También se realiza un proceso similar dentro del sistema de monitoreo 218 con el fin de generar la secuencia de huellas digitales candidatas CFPn y de nuevo la Figura 7 representa el flujo multimedia candidato 219 como una secuencia de tramas 306a-n.

35 En la realización que se describe, los generadores de huellas digitales de referencia 202a-d se disponen como un primer paso para eliminar información del flujo multimedia. En la realización que se describe, el generador de huellas digitales de referencia se dispone para convertir las tramas a la escala de grises 802 con el fin de reducir el tamaño de los datos a procesar. Se apreciará que para las realizaciones que se disponen para procesar video en tiempo real, puede resultar útil reducir la cantidad de datos para aumentar la velocidad a la que se pueden generar huellas digitales de referencia. Ha sido un problema generar huellas digitales (es decir, huellas digitales de referencia) para datos, 40 como un video o transmisiones de video, que se pueden generar en tiempo real y, sin embargo, permiten una detección sólida de transmisiones de multimedia que dan lugar a las mismas, o al menos a huellas digitales similares, (es decir, huellas digitales candidatas). Se cree que al menos algunas de las realizaciones descritas en la presente memoria resuelven este problema.

45 En realizaciones en las que el componente multimedia (el flujo multimedia en la realización que se describe) es distinto de los datos de video, la calidad puede eliminarse eliminando una o más frecuencias, eliminando el estéreo u otra información multicanal, o similar. Alternativamente, o adicionalmente, otras realizaciones podrían procesar canales de un contenido multimedia dentro de un componente multimedia por separado.

50 A continuación, los generadores de huellas digitales de referencia 202a-d se disponen para cambiar el tamaño de las tramas 300a-n con el fin de que las tramas 300a-n tengan dimensiones uniformes 804. En la realización que se describe, las tramas se redimensionan para que sean de 32x32 píxeles, reduciendo así la calidad del flujo multimedia. Los generadores de huellas digitales de referencia se disponen además para mantener la relación de aspecto de las tramas de manera que se añaden barras negras en la parte superior/inferior o en los bordes de las tramas para 55 mantener la relación de aspecto de las tramas en el estado redimensionado.

En otras realizaciones, los flujos multimedia se pueden redimensionar para que tengan un tamaño distinto de 32x32 píxeles. Se apreciará a partir del análisis que sigue que las realizaciones que usan un tamaño mayor generarán un mayor número de posibles huellas digitales reduciendo así la posibilidad de falsos positivos pero también que aumenta el procesamiento involucrado para generar las huellas digitales de referencia. En consecuencia, se ha descubierto que 60 las realizaciones que utilizan 32x32 píxeles ofrecen un buen equilibrio entre varios factores, incluido el número de huellas digitales potenciales, el procesamiento requerido para procesar los flujos multimedia en tiempo real (o al menos en tiempo pseudo-real) y el almacenamiento necesario para almacenar el flujo multimedia redimensionado.

65 En la realización que se describe, una segunda copia de las tramas 300a-n se redimensiona a 16x16 píxeles y se usa para determinar una cantidad de movimiento (es decir, cambio) entre tramas consecutivas 300a-n.

Las tramas de escala de grises redimensionadas se alimentan luego a dos memorias intermedias separadas, que en la realización que se describe son circulares, cada una dispuesta para contener porciones del componente multimedia (etapa 806). Las tramas de escala de grises redimensionadas a 32x32 píxeles se suministran a una de estas memorias intermedias y se utilizan para generar una huella digital y las tramas de escala de grises redimensionadas a 16x16 píxeles se introducen en una segunda de estas y se utilizan para detectar cambios. En la realización que se describe, cada una de las partes alimentadas a las dos memorias intermedias dura 4 segundos y comprende partes del flujo de medios de referencia 204a-d. Ahora se describe la generación de las huellas digitales utilizando las tramas redimensionadas que se encuentran en la primera de las memorias intermedias.

Una vez que se recibe el flujo multimedia 20a4a-d, no ocurre ninguna acción hasta que se llenan las memorias intermedias, pero luego se activa la generación de las huellas digitales de referencia RFPn y se genera una huella digital de referencia a intervalos de medio segundo. Otras realizaciones pueden estar provistas de memorias intermedias de diferente longitud y/o pueden generar huellas digitales de referencia en intervalos distintos de medio segundo. Otras realizaciones pueden tener, por ejemplo, memorias intermedias que son sustancialmente cualquiera de las siguientes longitudes (en segundos); 2, 3, 5, 6, 7, 8, 10, 20, 30, 60. Pueden disponerse otras realizaciones adicionales para generar huellas digitales de referencia sustancialmente en cualquiera de los siguientes periodos (en segundos): 0,1; 0,25; 1; 2; 3; 5; 10;60.

Por tanto, el flujo de tramas almacenado en memoria intermedia comprende una secuencia de tramas que comprende cada una X por Y píxeles. Por tanto, un píxel en (x, y) dentro de una trama se puede ubicar en cada una de las tramas y esto puede pensarse como un tubo de píxeles que atraviesan las tramas dentro de la memoria intermedia. Como tal, la memoria intermedia contiene X por Y tubos de píxeles y en la realización que se describe hay 32 x 32 tubos; 1024 tubos.

Como primer paso en la generación de huellas digitales de referencia, se suaviza cada uno de los tubos de píxeles. En la realización que se está llevando a cabo, esto se logra con un paso de perfil gaussiano 808.

A continuación, el contenido de la memoria intermedia se muestrea a una frecuencia de trama de 8 tramas por segundo, paso 810. Como tal, una memoria intermedia de cuatro segundos contiene 32 tramas de flujo multimedia muestreados. Por lo tanto, ahora hay 32 tramas cada una de 32 por 32 píxeles; un total de 32.768 píxeles.

Por tanto, se puede pensar que cada uno de los generadores de huellas digitales 202 de referencia contiene un reductor de resolución 134 que se dispone para procesar el flujo multimedia entrante u otros datos antes de que ese flujo multimedia se procese como se describe a continuación. Dicho submuestreador 134 puede estar dispuesto para realizar una o más de las etapas de procesamiento descritas anteriormente: conversión del flujo multimedia a escala de grises; redimensionar el flujo multimedia, que en la realización que se describe era de 32 x 32 píxeles; adición de las barras negras; y una reducción de la velocidad de tramas, que en la realización que se describe era de 8 fps.

A continuación, los generadores de huellas digitales de referencia se disponen para realizar una Transformación de Coseno Discreta (DCT) en los cubos de 32x32x32 (etapa 812); es decir, en esta realización se realiza una DCT tridimensional (3D). Se apreciará que las realizaciones que realizan la DCT en cubos de píxeles de lados iguales son ventajosas en vista de la optimización que puede producirse en la DCT, reduciendo así la cantidad de procesamiento. Como tal, en otras realizaciones, si se alteraran las dimensiones X o Y a las que se redimensionan las tramas 300a-n, entonces también puede ser conveniente alterar el número de tramas a las que se le redujo la resolución del flujo multimedia.

Por tanto, se apreciará que se extraen los componentes de frecuencia del componente multimedia (el flujo multimedia en esta realización). En la realización que se describe, esta extracción del componente de frecuencia la realiza la DCT, pero otras realizaciones pueden usar otros mecanismos para lograr esto. Por ejemplo, otras realizaciones también pueden usar una Transformada de Fourier, una Transformada de Fourier discreta, una Transformada de coseno o similares.

La DCT proporciona una serie de coeficientes, cada uno de los cuales representa un componente de frecuencia del bloque de 4 segundos del flujo multimedia de referencia 204a-d que ha sido procesado. Debido a que la DCT es una DCT 3D y se ha realizado no solo en dos dimensiones (XY de la trama) sino también en la dimensión del tiempo, a través de las tramas, la huella digital de referencia resultante se denomina huella digital espacio-temporal. En la realización que se describe, el 2do y 5to coeficientes, en cada una de las tres dimensiones, se seleccionan para su procesamiento adicional (etapa 814). En otras realizaciones, se puede seleccionar un conjunto diferente de coeficientes, sin embargo, las realizaciones que utilizan el 2do y 5to coeficientes se consideran ventajosas ya que estos coeficientes tienden a capturar el mayor cambio dentro de las tramas del flujo multimedia de referencia 204a-d proporcionando así una identificación sólida del componente multimedia candidato frente al componente multimedia de referencia. Estos componentes de frecuencia extraídos, en la realización que se describe, proporcionan la base de la huella digital que se genera.

A continuación, los generadores de huellas digitales de referencia se disponen para cuantificar los coeficientes seleccionados (en esta realización, estos son el 2do y 5to coeficientes) alrededor del valor medio de los coeficientes seleccionados (paso 816). Por lo tanto, hay coeficientes de 4x4x4, cada uno de los cuales se almacena como un número de un bit. Tal cuantificación da como resultado un número igual de unos y ceros y, en la realización que se describe, se dispone para generar 32 unos y 32 ceros. Las huellas digitales se muestran esquemáticamente en 304 en la Figura 7.

Por tanto, se genera una huella digital de referencia 304 según el contenido de una de las dos memorias intermedias. La otra memoria intermedia se usa para detectar movimiento dentro de la porción de 4 segundos del flujo multimedia como se describe ahora con referencia a la Figura 9. En la realización que se describe, esta reducción se realiza en la frecuencia de trama completa del componente multimedia de entrada, pero en otras realizaciones esto podría realizarse en las tramas de resolución reducida. Para realizaciones que procesan componentes multimedia que no son de video, la segunda memoria intermedia puede disponerse para monitorear la presencia de audio.

Por tanto, se puede pensar que el generador de huellas digitales 202a-d de referencia comprende un generador de huellas digitales 130 que se dispone para realizar los pasos descritos. De forma similar, también se puede pensar que el sistema de monitoreo 218 comprende un generador de huellas digitales 130. Estos generadores de huellas digitales 130 se disponen para generar una secuencia de huellas digitales a partir de una porción de los componentes multimedia candidatos o de referencia; es decir, los flujos multimedia respectivos en la realización que se describe.

Entonces se genera una medición de calidad, que en la realización que se describe puede ser realizada por el evaluador de calidad 132. Como primer paso, se determina un cambio en una propiedad entre al menos algunos de las tramas. En la realización que se describe, el cambio de propiedad es la diferencia de intensidad entre píxeles en tramas vecinas (es decir, se calcula la diferencia de intensidad entre píxeles (x1, y1) en tramas vecinas), paso 900, para generar un cambio de intensidad entre tramas. En la realización que se describe, este cambio de intensidad entre los píxeles vecinos se suma en todos los píxeles de las tramas para cada período entre tramas. En otras realizaciones, se puede sumar un subconjunto de píxeles.

Por lo tanto, se genera una matriz de 31 números, con un número para cada período entre tramas, y se utilizan 32 tramas para generar cada medición de calidad.

A continuación, la matriz se clasifica en orden de magnitud (paso 902) creando así una matriz ordenada que va en un extremo con el valor más alto en un extremo de la matriz al valor más bajo en el otro extremo de la matriz.

A continuación, el generador de huellas digitales de referencia 202 o el sistema de monitoreo 218 se dispone para calcular un percentil predeterminado de la matriz. En la realización que se describe, este percentil predeterminado es el percentil 95 de la matriz (paso 904). Posteriormente, el valor dado por el percentil predeterminado se normaliza por el número de píxeles (paso 906). Esto proporciona una medición de calidad que da una indicación de la cantidad de movimiento entre tramas que se utiliza durante las comparaciones de huellas digitales realizadas por el sistema 200.

Por último, los generadores de huellas digitales de referencia 202a-d se disponen para empaquetar las huellas digitales generadas en el paso 816 junto con las mediciones de calidad y cargadas en el almacén 210 (paso 908). Por tanto, las mediciones de calidad se almacenan en asociación con las huellas digitales. Algunas realizaciones pueden almacenar las huellas digitales y las mediciones de calidad por separado de una manera en la que están vinculadas o asociadas entre sí.

En la realización que se describe, las secuencias de huellas digitales de referencia RFPn se cargan en intervalos de minutos y, como tales, se cargan 57 huellas digitales a la vez, no habiendo huellas digitales generadas durante los primeros 3 segundos del flujo multimedia 204a-d hasta que la memoria intermedia se llena. Se asocia con cada una de las huellas digitales de referencia RFPn una marca de tiempo y la medición de calidad derivada para esa huella digital.

Se considera que la carga en el almacén 210 en lotes de un minuto proporciona un equilibrio entre la obtención rápida de huellas digitales en el almacén 210 y la reducción de la sobrecarga de la red asociada con la transmisión de pequeñas cantidades de información. Como tal, se pueden disponer otras realizaciones para cargar huellas digitales en el almacén 210 a diferentes intervalos. Por ejemplo, se pueden organizar otras realizaciones para cargar las huellas digitales de forma individual; o prácticamente en cualquiera de los siguientes períodos de tiempo: 10 segundos; 30 segundos; 45 segundos; 90 segundos; 2 minutos; 5 minutos; o más.

Cabe señalar que, en la realización que se describe, aunque cada huella digital representa 4 segundos del flujo multimedia 204a, se genera una huella digital mediante el método cada medio segundo. Por lo tanto, la generación de las huellas digitales de referencia 300a-n se puede considerar como si hubiera una ventana 305 movida a través de las tramas del flujo multimedia 204a (es decir, para moverse a través de la referencia o los componentes multimedia candidatos) para tener una pluralidad de posiciones relativas al componente multimedia (es decir, el flujo multimedia en la realización que se describe). Entonces se genera una huella digital para posiciones predeterminadas de la ventana 305. Por tanto, en la realización descrita, la memoria intermedia de cada uno del generador de huellas digitales

de referencia 202 y el sistema de monitoreo 218 proporciona generadores de ventana dispuestos para seleccionar respectivamente una porción del componente multimedia de referencia y los componentes multimedia candidatos (es decir, los flujos multimedia) y generar respectivamente una huella digital de referencia o candidata de la porción del componente multimedia cubierto por la ventana.

Por tanto, las posiciones de ventana generadas en períodos de tiempo consecutivos pueden considerarse como posiciones de ventana vecinas. Por tanto, se puede ver que las posiciones de ventana 305a en un primer período de tiempo que generó RFP1 es una posición de ventana vecina a la posición de ventana 305b usada para generar huellas digitales RFP2.

Por tanto, en la realización que se describe, la información de cualquiera de las tramas de resolución reducida está contenida en varias huellas digitales (por ejemplo, 4 en la realización que se describe). Esto se puede ver en las Figuras 7 y 8, donde se puede ver que la Huella Digital de Referencia 1 (RFP<sub>1</sub>) contiene información relacionada con las tramas 300a, 300b y 300c. La segunda huella digital de referencia (RFP<sub>2</sub>) se inicia desde la trama 300c. El experto apreciará que la Figura 7 es una simplificación y que el número de tramas se ha reducido para aumentar la legibilidad de la Figura y que en la realización que se describe, se genera una huella digital (por ejemplo, RFP<sub>1</sub>) a partir de 32 tramas (después del muestreo descendente).

En la realización que se describe, el almacén 210 contiene dos bases de datos: una primera base de datos que contiene huellas digitales de referencia RFP<sub>n</sub> para flujos multimedia de referencia 204a-d que están actualmente en vivo y una segunda base de datos que contiene huellas digitales de referencia RFP<sub>n</sub> para flujos multimedia que no están en vivo. Otras realizaciones pueden proporcionar diferentes disposiciones de datos con el almacén 210, pero se cree que es ventajoso proporcionar una base de datos para las huellas digitales asociadas con los flujos multimedia de referencia 204a-d en vivo para garantizar que la base de datos se reduzca en tamaño, aumentando así la velocidad de acceso a los flujos multimedia que se están transmitiendo actualmente (es decir, flujos multimedia en vivo).

En la realización que se describe, la o cada base de datos contenida en el almacén 201 (en esta realización dos bases de datos) se replican luego en los mecanismos de procesamiento de identificación 216a-d de manera que cada uno de los mecanismos de procesamiento de identificación mantiene una copia de la, o de cada, base de datos de huellas digitales de referencia RFP<sub>n</sub>. Se cree que tales realizaciones son ventajosas ya que es altamente escalable y el tamaño del sistema puede aumentarse añadiendo mecanismos de procesamiento de identificación adicionales 216a-d.

En la realización que se describe, el almacén 210 y los mecanismos de procesamiento de identificación 216a-d se alojan usando un sistema LINUX y el comando de shell LINUX rsync se usa para replicar la, o cada, base de datos en los mecanismos de procesamiento de identificación.

Por lo tanto, típicamente, una vez que uno de los generadores de huellas digitales de referencia 202a-d ha generado una huella digital de referencia 304 (RFP<sub>n</sub>), se carga en el almacén 201 dentro del orden de segundos después de que haya transcurrido el ciclo de carga del intervalo de minutos. Además, se cree que las realizaciones que utilizan el comando rsync son ventajosas ya que solo replican los cambios en la, o en cada, base de datos del almacén 210 en los mecanismos de procesamiento de identificación 216a-d aumentando así la velocidad y la eficiencia a la que se realiza la carga.

Además de la generación de huellas digitales de referencia a partir de los flujos multimedia de referencia 204a-d, el almacenamiento puede disponerse, en algunas realizaciones, para recibir huellas digitales de referencia de otras fuentes 222. Dichas huellas digitales de referencia pueden incluirse posteriormente en el proceso de análisis descrito a continuación para determinar si los componentes multimedia candidatos corresponden a los componentes multimedia de referencia (por ejemplo, flujos multimedia) que están siendo monitorizados por el sistema 200; los medios u otros datos que tienen una huella digital de referencia almacenada en el almacén 210 pueden considerarse medios (u otros datos) monitoreados por el sistema.

Algunas realizaciones pueden disponerse para monitorear múltiples tipos de contenido multimedia a la vez. Por ejemplo, algunas de las huellas digitales de referencia pueden generarse a partir de uno o más de los siguientes: flujos multimedia; videoclips; transmisiones de audio; clips de audio.

Las huellas digitales de referencia RFP<sub>n</sub> cargadas en el almacén 210 desde la fuente externa 222 pueden ser representativas de cualquier tipo de componente multimedia que esa realización sea capaz de monitorear identificando de otro modo los contenidos multimedia candidatos.

El sistema de monitoreo 218 se dispone para monitorear una o más fuentes de flujos multimedia candidatos 219. Estas fuentes pueden ser típicamente sitios web y el sistema de monitoreo típicamente puede disponerse para monitorear una pluralidad de dichas fuentes en cualquier momento. Cada flujo multimedia candidato 219 puede considerarse como un componente multimedia candidato; se apreciará que algunas realizaciones pueden monitorizar datos distintos de los flujos multimedia.

Convenientemente, algunas realizaciones se disponen para procesar porciones del flujo multimedia candidato de una longitud predeterminada. En particular, pueden disponerse realizaciones para generar huellas digitales para porciones de sustancialmente un minuto de un flujo multimedia candidato. Se apreciará que si el flujo multimedia se relaciona con algo como un partido de fútbol, entonces será significativamente más largo que esto. Sin embargo, se apreciará que si la realización se dispone para monitorear e intentar prevenir la transmisión de flujos multimedia no autorizados, entonces es deseable detectar la transmisión de ese flujo multimedia tan pronto como sea posible y, como tal, no es deseable esperar hasta que la transmisión del flujo multimedia haya terminado antes de generar huellas digitales y, como tal, es ventajoso procesar los flujos multimedia candidatos en porciones de la misma manera que es ventajoso generar huellas digitales para porciones de los flujos multimedia de referencia. Sin embargo, tampoco es deseable generar demasiado tráfico de red enviando cantidades excesivas de huellas digitales y se considera que proporcionar realizaciones que procesen porciones del flujo multimedia candidato en porciones de sustancialmente de 60 segundos proporciona un equilibrio. De nuevo, otras realizaciones pueden utilizar un período de tiempo distinto de 60 segundos.

Durante el uso, el sistema de monitoreo 218 se dispone para generar, usando el mismo método que se describe en las Figuras 8 y 9, una secuencia de huellas digitales candidatas CFPn para el o cada flujo multimedia candidato 219. Es decir, se utiliza el mismo método para generar la huella digital candidata CFPn que se utiliza para generar la huella digital de referencia. Por lo tanto, se generan una o más huellas digitales candidatas CFPn que son iguales, o al menos similares, a una o más huellas digitales de referencia RFPn, siempre que el flujo multimedia candidato subyacente a partir del cual se generan las huellas digitales candidatas CFPn (u otros datos) sea el mismo, o al menos similar, a un flujo multimedia de referencia que generó la huella digital de referencia RFPn. Por lo tanto, las realizaciones del sistema 200 podrían encontrar aplicación para determinar si los flujos multimedia candidatos corresponden a los flujos multimedia de referencia en vivo 204a-d, lo que en sí mismo es útil para determinar e intentar detener la piratería en línea y detectar copias de los flujos multimedia de referencia 204 monitoreados por el sistema.

Otras realizaciones pueden encontrar otra aplicación. Por ejemplo, el almacén 210 puede proporcionar una librería en la que las huellas digitales candidatas CFPn generadas a partir de un componente multimedia candidato pueden identificarse frente a las huellas digitales de referencia RFPn contenidas en el almacén 210. Dichas realizaciones permitirían, por lo tanto, que un componente multimedia (tal como un archivo de vídeo; una imagen, un archivo de audio o similar) se verificara con el almacén 210 en un intento de identificar ese componente multimedia candidato como una coincidencia con los componentes multimedia a los que se hace referencia en el almacén 210.

La generación de las huellas digitales candidatas se representa de nuevo esquemáticamente en la Figura 7, que muestra un flujo multimedia candidato 219 entrante representado como una secuencia de tramas 306a, 306b, 306c, 306d. Como es el caso con la generación de las huellas digitales de referencia, una ventana móvil 308 se mueve a lo largo del flujo multimedia 219 y se genera una Huella digital Candidata CFP<sub>1</sub> - CFP<sub>n</sub> para posiciones predeterminadas de la ventana 308. Al igual que con la generación de huellas digitales de referencia, en la realización que se describe, se genera una huella digital sustancialmente cada medio segundo, aunque cada huella digital cubre 4 segundos de las tramas 306a-n. Como se ha descrito anteriormente, una memoria intermedia circular proporciona un generador de ventanas diseñado para proporcionar la ventana necesaria para recibir y procesar una parte del componente multimedia.

Al igual que con la generación de las huellas digitales de referencia, las huellas digitales candidatas se agrupan antes de cargarse en la cola 220. De nuevo, las huellas digitales candidatas, una marca de tiempo y la medición de calidad asociada se cargan en grupos de 57 huellas digitales candidatas que representan la longitud de un minuto del flujo multimedia candidato 219 que se ha procesado.

Como se describió anteriormente, el sistema genera una cola 220 de secuencias de huellas digitales candidatas CFPn para ser comparadas con las huellas digitales de referencia RFPn dentro del almacén 210. Además de las huellas digitales candidatas CFPn generadas por el sistema de monitoreo 218, se pueden añadir huellas digitales candidatas, en al menos algunas realizaciones, a la cola 220 desde una fuente externa 224 capaz de generar huellas digitales candidatas CFPn.

A continuación, se accede a las secuencias en cola de huellas digitales candidatas y se eliminan de la cola 220 mediante uno de los mecanismos de procesamiento de identificación 216a-dy ese mecanismo de procesamiento de identificación busca luego en las huellas digitales de referencia contenidas en el almacén 210 para la aparición de una huella digital de referencia RFPn correspondiente a una o más huellas digitales candidatas CFPn de la secuencia de huellas digitales candidatas CFPn. Esta búsqueda de huellas digitales candidatas CFPn intenta localizar huellas digitales de referencia RFPn coincidentes dentro del almacén 210 identificando así el componente multimedia (por ejemplo, flujo multimedia) que generó la huella digital candidata CFPn como el mismo componente multimedia (por ejemplo, flujo multimedia) que generó la huella digital de referencia RFPn.

Cada uno de los mecanismos de procesamiento de identificación 216a-d mantiene una copia completa de la, o de cada base de datos del almacén 210. En la realización que se describe, esta copia se mantiene dentro de la memoria de acceso aleatorio (RAM) del mecanismo de procesamiento de identificación 216a-d. Las realizaciones con una disposición de este tipo son ventajosas debido a la velocidad de procesamiento que esto permite. Sin embargo, otras realizaciones pueden contener la copia de la, o de cada, base de datos dentro de cualquier forma de almacenamiento

al que pueda acceder el circuito de procesamiento del mecanismo de procesamiento de identificación 216a-d, ya sea en red o local al circuito de procesamiento.

5 La copia de la o de cada base de datos dentro del mecanismo de procesamiento de identificación 216a-d se ve obligada de vez en cuando a actualizarse para reflejar cualquier cambio dentro de la(s) base(s) de datos en el almacén 210. La actualización puede, por ejemplo, obligarse a que se produzca sustancialmente en el orden de minutos, tal vez sustancialmente en el orden de cualquiera de los 30 segundos siguientes; cada minuto; cada 5 minutos; cada 10 minutos; en cualquier momento entre estos tiempos.

10 Una vez que el mecanismo de procesamiento de identificación 216a-d ha obtenido una secuencia de huellas digitales candidatas CFPn de la cola 220, este busca al menos algunas, y típicamente cada una de las bases de datos a las que tiene acceso para determinar si las huellas digitales candidatas CFPn de la cola coinciden con alguna de las huellas digitales de referencia RFPn que se encuentran dentro de las bases de datos que se buscan. Aunque las huellas digitales se cargan en la cola 220 en grupos relacionados con porciones del flujo multimedia candidato (como una porción de un minuto de duración), cada una de las huellas digitales de 64 bits relacionadas con 4 segundos del flujo multimedia candidato se busca en la o en cada base de datos.

15 Con el fin de aumentar la seguridad de una coincidencia y reducir las posibilidades de un falso positivo y/o un falso negativo, se pueden agregar huellas digitales como se describe a continuación. La agregación puede tomarse como una característica que se evalúa para determinar si una huella digital candidata dentro de una secuencia coincide con una huella digital de referencia dentro de una secuencia.

20 La distancia de Hamming entre dos secuencias de bits es igual al número de bits que difieren entre las dos secuencias. En el caso de una huella digital de 64 bits, una distancia de Hamming de 0 es una coincidencia perfecta, mientras que una distancia de Hamming de 32 muestra una falta total de correlación. En la realización que se describe, debido a que una huella digital siempre contiene una cuenta igual de elementos 1/0, la distancia de Hamming es siempre un número par. Otras realizaciones pueden generar distancias de Hamming no uniformes.

25 Las pruebas con datos de ejemplo sugieren que un umbral de distancia de Hamming de aproximadamente 12-16 es adecuado para identificar coincidencias positivas entre huellas digitales individuales para realizaciones del sistema descrito en la presente memoria.

30 Se apreciará que en la realización que se describe, la distancia de Hamming se usa como una medida para determinar si una huella digital de referencia es similar, o la misma, a una huella digital candidata. En otras realizaciones, se pueden utilizar diferentes técnicas para realizar esta comparación.

35 Por tanto, cada uno de los mecanismos de procesamiento de identificación 216a-d se dispone para hacer una comparación de una huella digital candidata obtenida a partir de una secuencia de huellas digitales de la cola 220 con todas las huellas digitales de referencia RFPn de las que tiene conocimiento el sistema. Esta comparación la realiza típicamente el comparador 128 y se utiliza para localizar huellas digitales de referencia RFPn que coinciden con al menos una de las huellas digitales candidatas CFPn.

40 El mecanismo de procesamiento de identificación 216a-d se dispone así para comparar cada huella digital individual (por ejemplo, cada número de 64 bits en la realización que se describe) en la huella digital candidata CFPn que está procesando con cada una de las huellas digitales de referencia RFPn dentro del almacén 210. El mecanismo de procesamiento de identificación 216a-d se dispone además para acumular pares de huellas digitales candidatas y de referencia que coinciden donde se determina una coincidencia si:

45 1. La distancia de Hamming entre las huellas digitales candidatas y de referencia es inferior o igual a un umbral predeterminado, que es la realización que se describe en 14; y

50 2. La medición de calidad de las huellas digitales tanto candidatas como de referencia es superior o igual a un umbral predeterminado, que en la realización que se describe es 3,0.

55 Por lo tanto, en la realización que se describe, la distancia de Hamming y/o la medición de calidad pueden tomarse como una característica que puede evaluarse para determinar si una huella digital candidata dentro de una secuencia coincide con una huella digital de referencia dentro de una secuencia

60 Esta coincidencia entre las huellas digitales candidatas y de referencia (RFPn) se ejemplifica en la Figura 11a en la que una secuencia de huellas digitales de referencia generadas a partir de un flujo multimedia de referencia 204a se muestra encima de una secuencia de huellas digitales candidatas (CFPn) generadas a partir de un flujo multimedia candidato 219.

65 Por lo tanto, tomando la huella digital CFP1 de la secuencia como ejemplo, entonces se busca la CFP1 de entre todas las huellas digitales de referencia almacenadas en el almacén 210. Las flechas entre las huellas digitales de referencia

RFPn y las huellas digitales candidatas CFPn representan las huellas digitales que se considera que coinciden según los criterios anteriores. Por tanto, en el ejemplo que se da, se puede ver que CFP1 coincide con RFP2; RFP4 y RFP5.

Se recuerda que en la realización que se describe, los flujos multimedia 204a y 219 son flujos de video y, como tales, es probable que las tramas vecinas dentro de los flujos de video no cambien significativamente. Esta baja cantidad de cambio entre tramas es particularmente el caso de periodos en los que hay poco movimiento en el flujo de video. Es probable que se observen efectos similares en flujos de otros tipos de datos.

La similitud entre las tramas puede resultar en múltiples coincidencias de huellas digitales dentro de un flujo. Además, los flujos de multimedia que tienen un contenido similar también pueden resultar en coincidencias en las huellas digitales representativas de esas transmisiones. En un ejemplo, una secuencia similar de movimientos de cámara (por ejemplo, panorámica y zoom y cambio de toma) dentro de dos partidos de fútbol diferentes (u otro contenido) puede generar secuencias similares de huellas digitales. Puede haber coincidencias coincidentes con huellas digitales en otras transmisiones.

También se apreciará que las realizaciones pueden no tener una alineación temporal entre las huellas digitales generadas a partir de los flujos candidato y de referencia. De hecho, es probable que no se conozca la relación temporal entre un flujo multimedia candidato 219 y un flujo de referencia 204a. Por lo tanto, al observar CFP3, se puede ver, en este ejemplo, que se ha comparado con RFP1, RFP6 y RFPn. Se observará que RFP1 está antes de cualquiera de las huellas digitales de referencia (RFPn) con las que se ha comparado CFP1, lo que hace menos claro cómo las huellas digitales candidatas (CFPn) se alinean temporalmente con las huellas digitales de referencia (RFPn), asumiendo que las huellas digitales representadas en las Figuras 11a y 11b representan de hecho el mismo flujo multimedia.

Es probable que la Figura 11a sea una simplificación en cuanto al número de coincidencias entre las huellas digitales candidatas CFP1 y las huellas digitales de referencia RFP1 y también a la naturaleza dispersa ya que, como se analizó anteriormente, pueden coincidir con otros flujos multimedia.

Algunas realizaciones pueden disponerse para añadir una marca de tiempo a cada huella digital que se genere. Además, tales realizaciones pueden disponerse para limitar las huellas digitales de referencia con las cuales se comparan las huellas digitales candidatas con las que se produjeron dentro de un tiempo predeterminado de la huella digital candidata. Tales realizaciones pueden permitir una búsqueda más rápida de las huellas digitales de referencia y también pueden ser útiles cuando el componente multimedia representa un flujo multimedia de Internet de un evento donde es probable que se produzcan flujos copiados dentro de un período de tiempo corto del flujo multimedia original.

Por tanto, los mecanismos de procesamiento de identificación 216a-d generarán un conjunto de coincidencias entre las huellas digitales candidatas CFPn y las huellas digitales de referencia RFPn. En la realización que se describe, que monitoriza un flujo multimedia, cada coincidencia corresponde a un segmento individual de cuatro segundos de un flujo multimedia de referencia 204a-d que coincide con un segmento de 4 segundos del flujo multimedia candidato 219; se recuerda que se genera una única huella digital de referencia o candidata a partir de una duración de 4 segundos de un flujo multimedia y esto se muestra en la Figura 12.

La Figura 12 representa un flujo multimedia (ya sea un flujo multimedia de referencia o candidato) mediante una línea y se puede ver que los flujos están marcados en intervalos largos de segundos desde 0 segundos hasta 10 segundos. A continuación, se muestran nueve huellas digitales debajo del flujo multimedia 204, 219 (FP1 a FP9). La primera de estas huellas digitales FP1 se extiende desde  $t=0$  hasta  $t=4$ ; el segundo FP2 desde  $t=1$  hasta  $t=5$ ; el tercer FP3 desde  $t=2$  hasta  $t=6$ ; etc. Por lo tanto, la primera huella digital FP1 se genera mediante una posición de ventana que se extiende a través del componente multimedia (es decir, el flujo multimedia) desde  $t=0$  a  $t=4$ ; se produce una segunda posición de ventana entre  $t = 1$  y  $t = 5$  y, por lo tanto, existe la diferencia de tiempo de 1 segundo entre las posiciones de ventana analizadas en la presente memoria.

Se puede ver que se requieren 5 huellas digitales para cubrir y un intervalo de 8 segundos. Se cree que la naturaleza superpuesta de las huellas digitales es una ventaja para disminuir la importancia de un cambio en el flujo multimedia dentro de cualquier huella digital FPn. Algunas realizaciones del sistema deben tener una región de 6 huellas digitales coincidentes (es decir,  $M = 6$  como se describe a continuación) que corresponde a las 5 huellas digitales necesarias para dar 8 segundos de cobertura pero con una huella digital adicional que se cree que aumenta la robustez.

En otras realizaciones, se ha encontrado que la diferencia de tiempo entre las posiciones de las ventanas utilizadas para generar las huellas digitales es ventajosamente de 0,5 s. Se apreciará que la reducción de la diferencia de tiempo entre las posiciones de las ventanas aumenta la fidelidad con la que se puede representar el comportamiento de la huella digital a lo largo del tiempo (ya sea candidata o de referencia). Como tal, y con referencia a la realización que se describe, una diferencia de tiempo menor permite que el sistema responda a cambios más rápidos dentro de los flujos multimedia 204, 219.

Por tanto, en la realización que se describe, el generador de ventana se dispone de manera que porciones del componente multimedia generadas por las posiciones vecinas de la ventana se superponen. Por supuesto, pueden

disponerse otras realizaciones de modo que el generador de ventanas se disponga para generar posiciones de ventanas vecinas que no se superpongan.

En particular, se apreciará que en la realización que se describe, se usa una DCT para generar las huellas digitales (ya sean candidatas o de referencia) que extraen información de frecuencia del flujo multimedia. Se ha descubierto que las realizaciones que utilizan las frecuencias más bajas son ventajosas para generar coincidencias sólidas en la comparación entre las huellas digitales candidatas y de referencia. Sin embargo, para capturar esas frecuencias más bajas se requiere un período de tiempo más largo (en comparación con las frecuencias más altas). Se ha descubierto que un período de tiempo sustancialmente de 4 segundos proporciona suficiente información de baja frecuencia para proporcionar coincidencias sólidas mientras que el uso de ventanas de tiempo superpuestas permite una frecuencia de muestreo razonablemente alta. Se apreciará que una frecuencia de muestreo más alta aumenta la capacidad de generar huellas digitales precisas para un flujo de video que cambia rápidamente.

Además, como puede verse en la Figura 12, una vez que se establece el flujo, cualquier intervalo de un segundo está cubierto por 4 huellas digitales que corresponden a la longitud de cuatro segundos de la huella digital. Por tanto, en otras realizaciones en las que la longitud del flujo multimedia cubierto por una huella digital FP es distinta de cuatro segundos, entonces un número diferente de huellas digitales cubriría cada intervalo de tiempo. Se cree que la naturaleza superpuesta de las huellas digitales es ventajosa y permite que el proceso de agregación que se describe a continuación en relación con la Figura 11b omita las huellas digitales y aún tenga la confianza de que se ha encontrado una coincidencia entre un flujo multimedia candidato con un flujo de referencia mantenido dentro del almacén 210 ya que los datos para ese período todavía se están igualando.

Es concebible que las coincidencias entre huellas digitales candidatas (CFPn) y de referencia (RFPn) puedan ser falsos positivos. También es concebible que un segmento de 4 segundos del flujo multimedia candidato 219 que debería haber producido una coincidencia (es decir, coincide con un flujo multimedia de referencia) no haya dado como resultado una coincidencia (es decir, un falso negativo). Como se analizó anteriormente, si los datos se relacionan con un flujo multimedia, como un video, entonces es probable que las tramas vecinas del flujo multimedia (por ejemplo, 300a-d) sean similares entre sí y, por lo tanto, es probable que las huellas digitales, ya sean candidatas o de referencia, que se generan a partir de regiones de tiempo similares del flujo multimedia, sean similares entre sí.

Por consiguiente, cada uno de los mecanismos de procesamiento de identificación se dispone además para agregar un número de coincidencias como se describe ahora con referencia a las Figuras 10 y 11b. Como punto de partida y como se describe en relación con la Figura 11a, se busca en el almacén 210 para localizar cada Huella digital de referencia RFPn que coincida con una Huella digital candidata CFPn dada. A continuación, se lleva a cabo un proceso de agregación para proporcionar una identificación más sólida de un flujo multimedia candidato 219 frente a un flujo multimedia de referencia 204a (u otro componente multimedia) para el cual se mantienen las huellas digitales candidatas RFPn dentro del almacén 210. Como se describió anteriormente, una coincidencia entre una única huella digital de referencia RFPn y las huellas digitales candidatas CFPn no identifica necesariamente una coincidencia entre un flujo multimedia candidato 219 y un flujo multimedia de referencia 204a, o entre cualquier otro componente multimedia candidato y componente multimedia de referencia

Como se describe en relación con la Figura 11a, es probable que haya al menos algunas coincidencias falsas con cualquier huella digital candidata CFPn que se esté procesando. En consecuencia, se cree que las realizaciones que implementan el proceso de agregación son ventajosas ya que aumentan la robustez de la identificación del flujo multimedia candidato 219 (u otro componente multimedia) dentro de los flujos multimedia de referencia 204a (u otro componente multimedia) que han generado las huellas digitales de referencia dentro del almacén 210. Esta mayor robustez proporcionada por el proceso de agregación puede permitir, en algunas realizaciones, tomar una huella digital menos robusta del componente multimedia, lo que puede permitir tomar una huella digital más rápidamente.

Por lo tanto, el proceso de agregación identifica las regiones 1002 de las huellas digitales candidatas CFPn generadas a partir del flujo multimedia candidato 219 que se asignan a regiones de huellas digitales RFPn generadas a partir del flujo multimedia de referencia 204a y utilizan la relación temporal para aumentar la creencia de que existe una coincidencia entre una huella digital candidata CFPn y una Huella digital de referencia RFPn.

Por lo tanto, al observar la Figura 11a se puede ver que existen coincidencias entre las siguientes huellas digitales dentro de la secuencia de huellas digitales que se muestra en la Figura:

CFP2 y RFP5;

CFP3 y RFP6;

CFP4 y RFP7; y

CFP6 y RFPn.

También debe tenerse en cuenta que en el ejemplo dado, no se encontró que CFP5 coincidiera con la huella digital de referencia RFP8.

5 La Figura 11b muestra cómo el proceso de agregación determina que se ha producido una coincidencia entre el flujo multimedia candidato 219 y el flujo multimedia de referencia 204a.

10 El proceso de agregación busca cualquier región de huellas digitales candidatas CFPn dentro de la secuencia que tenga múltiples coincidencias y cualquier región que pase el proceso de agregación se informa como una coincidencia positiva. En una realización, esto puede significar que el flujo multimedia candidato 219 coincide con un flujo multimedia de referencia 204a-d que está siendo monitoreado por el sistema 200. En otras realizaciones, esto puede significar que un componente multimedia candidato se ha identificado frente a huellas digitales de referencia que representan un componente multimedia de referencia que se mantiene dentro del almacén 210. Como se describió anteriormente, se informa una coincidencia positiva si se mantiene un número mínimo de huellas digitales candidatas en una secuencia para hacer coincidir las huellas digitales con la secuencia de huellas digitales de referencia.

15 En el contexto del párrafo anterior, una región de huellas digitales puede comprender cualquier número de huellas digitales consecutivas y puede comprender, por ejemplo, sustancialmente 2, 3, 4, 5, 6, 7, 8, 9, 10, 15, 20, 25, 30 o más huellas digitales consecutivas.

20 Se dice que una secuencia de huellas digitales candidatas es una coincidencia global, agregada de una secuencia de huellas digitales de referencia si, para el mismo número de huellas digitales de referencia y candidatas (es decir, el mismo período de tiempo) - paso 1004:

25 1. La secuencia tiene al menos M huellas digitales de longitud, donde M en la realización que se describe se establece en 6;

2. La secuencia comienza con una coincidencia positiva de huella digital candidata/de referencia;

30 3. No hay ninguna brecha entre coincidencias positivas de candidatas/de referencia superiores a D, en donde la realización actual D se establece en 3. Esta prueba asegura que haya menos de un número predeterminado de brechas dentro de la secuencia de huellas digitales. En la realización que se describe, esto equivale a que es permisible no tener coincidencias entre las huellas digitales candidatas y de referencia de hasta 3 segundos, mientras que un período de 4 segundos daría como resultado una falla para cumplir la prueba de agregación.

35 Poniendo esto en contexto con la Figura 11b, se observa que CFP2 es el comienzo del bloque 704 de coincidencias (identificado por la marca en CFP2) que se asigna a RFP5. El bloque 704 contiene 9 huellas digitales candidatas (CFP2 a CFP10) que incluyen 7 coincidencias (CFP2); CFP3; CFP4; CFP6; CFP8; CFP9; y CFP10) y dos errores (CFP5 y CFP7), es decir, separaciones entre coincidencias vecinas. Por consiguiente, este bloque 704 cumple con los criterios que se han establecido y se considera que se ha producido la coincidencia entre el flujo multimedia candidato 219 y el flujo multimedia de referencia 204a ya que se han cumplido los puntos 1 a 3 anteriores. Se observará que, por razones de claridad, solo se muestran algunos de los vínculos entre las huellas digitales candidatas (CFP10) y las huellas digitales de referencia FRPn.

45 Por tanto, en la realización que se describe, el bloque 704 cumple los criterios y se considera que se ha producido una coincidencia entre ese bloque del flujo multimedia candidato 219 y el flujo multimedia de referencia 204a; etapa 1006.

50 Por tanto, en resumen de lo anterior, el comparador puede utilizar varias características para evaluar si la secuencia de huellas digitales de referencia coincide con la secuencia de huellas digitales candidatas. En la realización que se ha descrito, estas características pueden incluir características sobre las coincidencias entre huellas digitales individuales así como características sobre la comparación de la secuencia de huellas digitales (por ejemplo, las secuencias de cuatro segundos).

Las características de las coincidencias entre huellas digitales individuales pueden incluir al menos una de las siguientes:

- 55 - una comparación de la distancia de Hamming entre las 2 huellas digitales.
- una comparación de la calidad de la huella digital candidata con un umbral de calidad.
- 60 - una comparación de la calidad de la huella digital de referencia con un umbral de calidad.

Por lo tanto, si alguno de estos no se cumple, se considera que la huella digital candidata no coincide con la huella digital de referencia.

65 Las características de las coincidencias entre secuencias de huellas digitales pueden incluir al menos una de las siguientes:

- si un número umbral de huellas digitales agregadas coincide en la secuencia.
- si en la secuencia se supera el umbral más grande de “brecha” de huellas digitales que no coinciden.
- si se alcanza un umbral de distancia de Hamming medio a lo largo de la secuencia.
- si se cumple una diferencia de calidad entre las huellas digitales candidatas y las huellas digitales de referencia en toda la secuencia.

Por tanto, si no se alcanza alguno de estos umbrales de nivel de secuencia, la realización determinará que la secuencia candidata no coincide con la secuencia de referencia.

Para intentar evitar que el sistema sea eludido simplemente por manipulaciones del flujo multimedia candidato, el flujo multimedia se refleja tanto horizontal como verticalmente y cada orientación pasa a través del proceso de toma de huellas digitales. Por lo tanto, en la realización que se describe, cada flujo multimedia candidato pasa por el proceso de toma de huellas digitales 4 veces: sin reflejar; reflejado verticalmente; reflejado horizontalmente; y reflejado tanto horizontal como verticalmente.

En otras realizaciones, los componentes multimedia candidatos pueden sufrir otro procesamiento para intentar evitar la anulación del sistema.

Los mecanismos de procesamiento de identificación 216a-d se disponen entonces para informar de cada región que pasa el proceso de agregación como una coincidencia positiva entre una región del flujo multimedia candidato y del flujo multimedia de referencia.

En una realización, el sistema descrito en relación con las Figuras anteriores puede usarse en el proceso descrito en relación con la Figura 13 que se describe ahora.

Como primer paso 1, las huellas digitales de referencia RFPn se generan y almacenan en el almacén 210 como se describió anteriormente. Además, como se describió anteriormente, el sistema puede disponerse para monitorear muchos tipos de datos pero, en la realización que se describe, está dirigido a monitorear varios tipos de componentes multimedia: particularmente uno o más de los siguientes: archivos de video; video transmitido; archivo de audio; archivos de audio transmitidos.

Las realizaciones que monitorean los componentes multimedia distintos de los archivos de video o audio típicamente utilizarán un algoritmo de huellas digitales distinto al basado en una DCT descrita en la presente memoria.

En el paso 2, un usuario reproduce una versión inicial de un componente multimedia, por ejemplo, reproduciendo el componente multimedia en una página web o similar. Si esta versión inicial del componente multimedia ha sido procesada por el sistema y tiene una huella digital de referencia RFPn almacenada en el almacén 210, puede ser identificada por el sistema. En la Figura 13, se muestra que el paso 2 se origina en el mismo origen que el paso 1, lo que implica que el componente multimedia inicial, reproducido por el usuario, se origina en la misma fuente que el componente multimedia a partir del cual se genera la huella digital de referencia RFPn. Sin embargo, no es necesario que este sea el caso y, en otras realizaciones, un usuario puede obtener y reproducir el componente multimedia inicial desde cualquier fuente.

En el paso 3, el componente multimedia inicial que está siendo reproducido por un usuario se trata como un componente multimedia candidato y, por lo tanto, tiene las huellas digitales candidatas CF-Pn generadas para él como se describió anteriormente. Se apreciará que pueden generarse huellas digitales candidatas CFPn para el componente multimedia independientemente del origen del componente multimedia. El método se dispone entonces para determinar, como se describió anteriormente y en particular con referencia a las Figuras 11a y 11b, si las huellas digitales candidatas CFPn generadas de este modo corresponden a alguna de las huellas digitales de referencia RFPn almacenadas en el almacén 210; es decir, se realiza una búsqueda de la huella digital candidata comparándola con el conjunto de huellas digitales de referencia 210. El experto apreciará que si la secuencia de huellas digitales candidatas CFPn se identifica como correspondiente a una secuencia de huellas digitales de referencia RFPn, se supone que el componente multimedia candidato es el mismo que el componente multimedia de referencia (es decir, el archivo que generó la huella digital de referencia) y así, se ha identificado el componente multimedia candidato.

En un cuarto paso del método mostrado en la Figura 13, el circuito de procesamiento que un usuario está usando para reproducir el componente multimedia inicial se dispone para conectarse a una fuente desde la cual se puede obtener una versión mejorada del archivo multimedia inicial. Se apreciará que tal método se puede hacer que funcione con una variedad de dispositivos de procesamiento y, por ejemplo, un usuario puede ver el archivo multimedia inicial en cualquiera de los siguientes: un ordenador personal (tal como una PC con Windows o LINUX u otro sistema operativo; un MAC; un teléfono móvil (tal como un iPhone, un dispositivo Android, un móvil Windows, una Blackberry o similares); un televisor inteligente; una tableta (tal como un iPad, un Kindle, un Android o similar), etc.

La fuente del componente multimedia inicial puede entonces disponerse para suministrar una versión del componente multimedia inicial al dispositivo de procesamiento que está utilizando el usuario. La versión del componente multimedia inicial puede tener un valor agregado en comparación con la versión inicialmente reproducida por el usuario.

El valor añadido puede incluir cualquiera de los siguientes elementos: mayor calidad; sonido mejorado; un comentario; subtítulos; diferentes ángulos de cámara; más pistas de audio; escenas ocultas o similares.

El sistema puede disponerse para cargar una cuenta, o cobrar de otro modo, al usuario por el suministro del componente multimedia mejorado; por ejemplo, el componente multimedia con el valor añadido.

Además, es conveniente describir el sistema en relación con el monitoreo de flujos de video como el mecanismo para entregar contenido multimedia a un usuario. Sin embargo, en otras realizaciones, las técnicas descritas pueden aplicarse igualmente a otras formas de componente multimedia. Por ejemplo, los archivos de video, tales como los archivos MP4 y similares, pueden considerarse componentes multimedia. En otras realizaciones alternativas o adicionales, los componentes multimedia pueden ser proporcionados por contenido de audio tal como contenido de audio transmitido o archivos de audio discretos (tales como archivos MP3, la pista de audio de un video o similares). El experto en la técnica apreciará cómo modificar los algoritmos de marcas de agua y huellas digitales para procesar datos de audio en lugar de video. Por lo tanto, la referencia a la transmisión en el texto siguiente no debe considerarse como limitada a la transmisión de multimedia y también se pretende que se refiera a los mecanismos de entrega no transmitidos.

Se hace referencia a lo largo de la siguiente descripción a la Figura 16 que ilustra el proceso general que se emplea, por las realizaciones, para monitorear y/o proteger los componentes multimedia.

El sistema 1400 que se muestra en la Figura 14, que puede considerarse como una plataforma de distribución de contenido multimedia, permite tomar las huellas digitales de un flujo multimedia de referencia 204 (es decir, un componente multimedia) para que pueda identificarse como se describe a continuación. En las figuras anteriores se proporcionan más detalles sobre cómo se toma una huella digital.

Además, el sistema 1400 permite que el flujo multimedia de referencia 204 se entregue a una pluralidad de usuarios de manera que la versión entregada a cada uno de los usuarios tenga una marca de agua única incorporada en el mismo. Esta marca de agua única permite que el usuario al que se entregó una versión determinada del flujo multimedia de referencia se identifique como se describe a continuación.

Además, el sistema 1400 permite que una red, que en la realización que se describe es Internet y/o la World Wide Web 1402, sea monitoreada en busca de copias u otras versiones ilícitas (en lo sucesivo citadas como 'copias'), del flujo de contenidos multimedia de referencia 204. La red puede ser o no la misma que la red 251 descrita anteriormente.

Si se detectan copias, el sistema se dispone a determinar, utilizando la marca de agua única, el usuario al que se entregó esa versión del flujo multimedia de referencia. El sistema de monitoreo también se describe con más detalle en relación con las Figuras anteriores.

Por lo tanto, con referencia a la Figura 14, se puede ver que el flujo multimedia de referencia 204 pasa a través de un proceso de toma de huellas digitales 1600 proporcionado en un generador de huellas digitales de referencia 202 que genera una huella digital de referencia RFP que se puede utilizar para determinar si un flujo multimedia candidato coincide con un flujo multimedia de referencia 204 monitoreado por el sistema. Las huellas digitales generadas por el generador 202 se mantienen dentro de un almacén 210. Los componentes del sistema de toma de huellas digitales se describen con más detalle en la Figura 2 y los elementos de la Figura 14 descritos con más detalle en la Figura 2 se muestran dentro de la línea de cadena 1403 y las partes similares se denominan con números de referencia similares.

El flujo multimedia con huellas digitales 1406 pasa por un proceso de marca de agua 1407 que genera al menos dos versiones del flujo de huellas digitales 1406: una primera versión con marca de agua 1408; y una segunda versión 1410 con marca de agua. Mientras que, en la realización que se describe, solo hay dos versiones con marca de agua del flujo multimedia de referencia, otras realizaciones pueden proporcionar más de dos versiones.

Como se describe a continuación, se aplica una marca de agua a cada trama de la versión con marca de agua 1408, 1410 y se aplica la misma marca de agua a cada trama dentro de la misma versión del flujo multimedia con marca de agua. Las versiones con marca de agua se almacenan luego dentro de un sistema de entrega 1412.

En la realización que se describe, se aplica una marca de agua diferente a las tramas de cada una de la primera versión con marca de agua 1408 y la segunda versión con marca de agua 1410. Al menos algunas realizaciones, incluida la que se describe, usan una marca de agua ortogonal de modo que la marca de agua aplicada al primer flujo con marca de agua sea lo más distinguible posible del segundo flujo con marca de agua. En un ejemplo, se puede

usar una secuencia como 1,2,-1,3,-2... para generar la marca de agua para el primer flujo. En la realización que se describe, la secuencia ortogonal -1,-2, 1, -3, 2... se usa entonces para generar la segunda marca de agua.

5 En otras realizaciones, es posible que sólo el segundo flujo con marca de agua 1410 tenga una marca de agua aplicada a las tramas de la misma y el primer flujo con marca de agua 1406 siga siendo una copia del flujo multimedia con huellas digitales 1406 sin que se le aplique ninguna marca de agua. Las realizaciones que aplican marcas de agua solo a algunos de los flujos multimedia pueden ser ventajosas en vista del procesamiento reducido requerido.

10 Para incrustar cada marca de agua, se procesa cada trama de la primera 1408 y la segunda 1410 versiones con marca de agua. Cada trama se transforma en un espacio de bloque de DCT (Transformada de coseno discreta), de manera que se aplica una DCT a cada bloque de píxeles NxN en la trama de entrada, donde N es 8 en la realización descrita. En otras realizaciones, N puede ser cualquier otro valor, pero se cree que N = 8 proporciona una marca de agua relativamente robusta sin presentar una carga demasiado grande para el procesador.

15 Se genera una secuencia pseudoaleatoria de coeficientes de marca de agua, y los coeficientes DCT en cada bloque DCT son modificados por esta secuencia pseudoaleatoria. La cantidad de cambio de cada componente está ponderada por la fuerza de DCT, para reducir los artefactos visuales que de otro modo podrían ser evidentes para un espectador de la trama. Luego, se aplica una DCT inversa a cada bloque para recuperar la versión con marca de agua de la imagen original. Cada trama de entrada (sin marca de agua) se procesará para generar una o más tramas de salida (con marca de agua), una por símbolo de marca de agua, donde cada símbolo de marca de agua tiene una secuencia pseudoaleatoria de coeficientes diferente. Aquí se apreciará que, en la realización descrita, se aplica un símbolo de marca de agua diferente a cada una de la primera 1408 y segunda 1410 versiones con marca de agua, pero que se aplica el mismo símbolo a cada trama de esas versiones.

25 Es concebible que, en algunas realizaciones, algunas tramas dentro del contenido multimedia no tengan marca de agua, de modo que todas las demás, cada enésima trama, o similares, tengan marcas de agua. Sin embargo, se cree que tales realizaciones son menos robustas y llevaría más tiempo recuperar la marca de agua de cualquier flujo multimedia candidato, ya que solo se procesa cada enésima trama, lo que proporciona menos información de la que recuperar la marca de agua.

30 La marca de agua basada en DCT se recupera sin recurrir a la imagen original incorrupta; de A. Piva, M. Barni, F. Bartolini y V. Cappellini; Dipartimento di Ingegneria Elettronica; Università di Firenze via S. Marta 3, 50139, Firenze, Italia describe técnicas similares para incrustar marcas de agua y el experto debe leer este documento para comprender cómo incrustar una marca de agua de modo que pueda recuperarse sin hacer referencia a la imagen original.

35 A Survey of Digital Image Watermarking Techniques”, Vidyasagar M. Potdar, Song Han, Elizabeth Chang también puede ser útil para los expertos.

40 La trama con marca de agua se procesa aún más para reducir los artefactos visuales asociados con la marca de agua. Se calcula una máscara espacial, donde el valor de la máscara representa la tolerancia del HVS (Sistema Visual Humano) a pequeños cambios en el valor de cada píxel en la trama. Por ejemplo, el valor de la máscara en cada píxel se puede establecer en la variación de los 8 píxeles que rodean inmediatamente. La trama con marca de agua enmascarada viene dada por la suma de la trama original y la trama con marca de agua, ponderada por la máscara HVS, de modo que la marca de agua está fuertemente incrustada en las regiones de la versión donde el HVS es menos sensible a las perturbaciones. Se aplica la misma máscara a cada trama con marca de agua, para producir una pluralidad de tramas con marca de agua enmascarados (dos en la realización actual).

50 En la realización que se describe, el proceso de marca de agua 1407 genera adicionalmente metadatos que contienen descriptores de características clave dentro de las tramas del flujo multimedia con marca de agua, características que se utilizan para alinear temporal y espacialmente los flujos multimedia candidatos para recuperar cualquier marca de agua en estos. Con el fin de reducir la complejidad computacional y los requisitos de almacenamiento, las características no se extraen para cada trama y, en esta realización, las características se extraen para cada 10ma trama. Otras realizaciones pueden almacenar características para más tramas (es decir, para tramas de menos de cada 10 tramas) o para menos tramas (es decir, para tramas de más de cada 10 tramas). Las características extraídas en la realización actual están en forma de descriptores de características ORB (Oriented FAST y Rotated BRIEF), que permiten reconstituir un flujo multimedia candidato reescalado y/o rotado a las dimensiones originales durante la extracción de la marca de agua.

60 Es conveniente describir el sistema de entrega 1412 como un protocolo de transmisión adaptativo (tal como transmisión en vivo HTTP (HLS)); transmisión HTTP adaptativa 3GPP; transmisión fluida de Microsoft™; transmisión dinámica adaptativa a través de HTTP (DASH)), pero en otras realizaciones se pueden usar sistemas distintos de los protocolos de transmisión adaptativa. Algunas realizaciones del sistema pueden disponerse para procesar flujos multimedia que no corresponden a un formato conocido. Algunas realizaciones del sistema pueden volver a combinar archivos de audio y video de flujos/archivos separados. Tal funcionalidad surge típicamente del procesamiento de la información de encabezado de las diversas capas del tráfico de red como se describe a continuación.

- 5 En un protocolo de transmisión adaptable, cada versión 1408, 1410 de los contenidos multimedia de referencia con marca de agua se codifica de tal manera que se mantienen múltiples versiones y cada versión es adecuada para diferentes velocidades de bits, diferentes tamaños de tramas, etc. Sin embargo, la marca de agua es tal que sobrevivirá la codificación de los flujos con marca de agua 1408, 1410. La figura muestra un depósito 1414, 1416 respectivamente para cada una de la primera y segunda versiones con marca de agua. El experto en la técnica apreciará que, si bien es conveniente mostrar repositorios separados, no es necesario proporcionar repositorios separados. Por tanto, en otras realizaciones puede haber más o menos repositorios.
- 10 El experto en la técnica apreciará que para un protocolo de transmisión adaptativo, el flujo multimedia se entrega, no como un flujo continuo, sino como una serie de segmentos de duración limitada (por ejemplo, bloques) generados a partir del flujo multimedia original. Por lo tanto, cada una de las múltiples versiones de los flujos multimedia contenidos dentro de los repositorios 1414, 1416 comprende una serie de segmentos de duración limitada. En la realización que se describe, cada uno de los segmentos de duración determinada es típicamente de 5 a 10 segundos de duración.
- 15 Todas las tramas dentro de un segmento de duración limitada determinado tendrán la misma marca de agua aplicada. Sin embargo, es concebible que al menos algunas realizaciones utilicen una longitud para cada segmento de duración limitada de menos de 5 segundos, que es el tiempo de segmento de duración limitada mínimo sugerido para algunos protocolos de transmisión adaptativos para garantizar que no haya demasiado tráfico de red.
- 20 Las realizaciones que no usan protocolos de transmisión adaptativos, en particular, pero no exclusivamente, pueden usar segmentos de duración limitada de menos de 5 segundos y, por ejemplo, pueden usar segmentos de duración limitada en el rango de 0,1 segundos a 15 segundos. El experto en la técnica apreciará que cuanto más corto sea el segmento de duración limitada, más rápidamente se puede extraer una marca de agua del contenido multimedia/flujo multimedia.
- 25 Cada segmento de duración limitada tendrá una hora de inicio y una hora de finalización en relación con el flujo multimedia original, es decir, de referencia. Por ejemplo, en una realización en la que cada segmento de duración limitada tiene una duración de 5 segundos, el tercer segmento de duración limitada comenzará a los 10 segundos en la secuencia y terminará a los 15 segundos en la secuencia. Tanto la hora de inicio como la hora de finalización pueden considerarse información de tiempo.
- 30 El sistema 1400 se dispone para almacenar la información de tiempo en relación con cada uno de los flujos multimedia contenidos en los repositorios 1414, 1416. En algunas realizaciones, la información de tiempo se mantiene dentro del almacén 210, pero no es necesario que este sea el caso.
- 35 Se dispone un proceso de entrega 1418 para atender la solicitud de un flujo multimedia desde un dispositivo cliente. Tres de estos dispositivos cliente se muestran en la Figura: un teléfono móvil 1420 (tal como un iPhone™); un dispositivo Android™; un móvil Windows™; o una Blackberry™); un ordenador 1422 (tal como un PC que ejecute Windows™, LINUX o similares; un ordenador Apple; o similares) y una tableta 1424 (tal como un iPad™; un Kindle™; una tableta Android™ o similares). El experto en la técnica apreciará que otros dispositivos, tales como televisores; relojes; o similares también pueden formar dispositivos cliente. Como se describe a continuación, el proceso de entrega 1418 se dispone para generar una versión única del flujo multimedia para cada uno de los dispositivos cliente 1420, 1422, 1424, cuya versión única se dispone para identificar el dispositivo cliente al que se entregó ese flujo multimedia. Por lo tanto, cada dispositivo cliente 1420, 1422, 1424 puede considerarse como un dispositivo objetivo al que se puede entregar contenido multimedia.
- 40 El proceso de entrega 1418 se dispone para generar una versión única del flujo multimedia para cada uno de los dispositivos cliente 1420, 1422, 1424, cuya versión única se dispone para identificar el dispositivo cliente al que se entregó ese flujo multimedia. Por lo tanto, cada dispositivo cliente 1420, 1422, 1424 puede considerarse como un dispositivo objetivo al que se puede entregar contenido multimedia.
- 45 Específicamente, un usuario 1420-1424 solicita un flujo multimedia (es decir, un componente multimedia) del sistema 1400. Según los protocolos de transmisión adaptativos (como HLS), el cliente, siguiendo la solicitud del flujo multimedia, recibe una lista de texto de URL (localizadores uniformes de recursos) que genera el sistema 1602. Estas URL son para los segmentos de video de duración limitada que componen una versión del flujo multimedia solicitado. Como se analizó anteriormente, cada uno de los repositorios 1414, 1416 contiene múltiples copias del flujo multimedia con el fin de que diferentes condiciones de canal, dispositivos de destino, etc. puedan utilizar el flujo multimedia.
- 50 Se apreciará que en otras realizaciones, un usuario puede solicitar un componente multimedia que no sea un flujo multimedia. Por ejemplo, un usuario podría descargar un archivo que proporcione contenido multimedia (por ejemplo, un archivo de video o un archivo de audio).
- 55 Además, el sistema 1400 determina la identidad del usuario 1420-1424 que realizó la solicitud del flujo multimedia. Por ejemplo, es probable que el usuario tenga una cuenta con el proveedor del flujo multimedia al que está realizando la solicitud y la identidad del usuario puede obtenerse de los detalles de la cuenta de ese usuario 1420-1424. En otras realizaciones, el usuario 1420-1424 puede necesitar ingresar un identificador, etc. cuando solicitan el flujo multimedia.
- 60 El sistema de entrega 1412 se dispone entonces para generar una lista de reproducción para ese usuario 1420-1424 para esa solicitud. Por lo general, esta lista de reproducción será única para esa solicitud particular del flujo multimedia. Las URL que componen la lista de reproducción apuntan a segmentos de duración limitada en cada uno de los repositorios 1414, 1416 dentro del sistema de entrega 1412 de manera que un conjunto de segmentos de duración
- 65

limitada que cuando se unen por el protocolo generan el flujo multimedia. Como tal, cuando el navegador u otro software utilizado por el usuario 1420-1424 accede a la lista de URL, se entregan segmentos de duración limitada que se generan a partir de cada una de la primera 1408 y segunda 1410 versiones con marca de agua del flujo multimedia. Debido a que la lista de URL ha sido generada por el sistema específicamente para cada uno de los usuarios 1420-1424, cada usuario recibe, en la realización que se describe, una combinación única de segmentos de duración limitada para formar el flujo multimedia entregado que se le entrega. Tales flujos multimedia entregados están ejemplificados por las líneas 1419a, 1419b, 1419c en la Figura 14.

En la realización que se describe, el sistema de entrega 1412 usa un número binario de 32 bits para generar la secuencia de URL. En la realización que se describe, se puede pensar que la elección del símbolo de marca de agua dentro de los segmentos de duración limitada del flujo multimedia enviado a un usuario 1420-1424 proporciona un cero o uno de un número binario. Aquí se apreciará que cada uno del primer 1408 y segundo 1410 flujos con marca de agua tiene un símbolo diferente aplicado a estos. Por ejemplo, la presencia de un segmento de duración limitada de la primera versión con marca de agua 1408 da un cero y la presencia de un segmento de duración limitada de la segunda versión con marca de agua 1410 da un uno. Por lo tanto, en la realización que se describe, se requieren 32 segmentos de duración limitada para identificar al usuario 1420-1424 a quien se transmitió el flujo multimedia. Por tanto, el orden en el que se entregan los segmentos de duración limitada proporciona un código de identidad que se ha asignado a un usuario determinado.

Por tanto, en la realización que se está describiendo, se necesitan 32 segmentos de duración limitada para proporcionar el código de identidad que se ha insertado en el flujo multimedia entregado 1419. Por lo tanto, suponiendo que un segmento de duración limitada tiene una duración de 5 segundos, se necesitan 2 minutos y 40 segundos de flujo multimedia para recuperar el código de identidad. Por lo tanto, la longitud de los segmentos de duración limitada y el número de bits que codifican el código de identidad son un equilibrio entre proporcionar una identificación robusta, permitir indexar una gran población de usuarios y permitir que el código de identidad se recupere rápidamente de un flujo multimedia.

El experto en la técnica apreciará que otras realizaciones pueden utilizar un número diferente de bits. Además, otras realizaciones pueden usar bases numéricas distintas de la base 2 para la generación del código y tales realizaciones comprenderían más de dos versiones de los flujos multimedia con marca de agua: para generar un número de base 3 se requeriría la versión 3 del flujo multimedia con marca de agua, etc.

A medida que el contenido multimedia se entrega a los dispositivos cliente 1420-1424, un sistema de monitoreo 218 monitorea 1604 la red 1402/251 en busca de contenidos multimedia que tengan una huella digital almacenada dentro del almacén 210. Si el sistema de monitoreo 218 detecta, en la red, copias de un flujo multimedia de referencia 204 que tienen una huella digital reconocida (es decir, una huella digital almacenada en el almacén 210), entonces un controlador 1426 se dispone para identificar, usando cualquier marca de agua dentro del flujo multimedia, el usuario 1420-1424 al que se envió el contenido multimedia. El controlador 1426 se dispone entonces, en algunas realizaciones, para detener la entrega del flujo multimedia entregado a ese usuario 1420-1424 controlando 1428 el proceso de entrega 1418.

Aunque solo se muestra un generador de huellas digitales 202 en la Figura 1, el sistema de huellas digitales 200 típicamente tendrá varios generadores de huellas digitales de referencia 202a, 202b, 202c, 202d. La Figura 2 muestra cuatro generadores de huellas digitales 202a-d, pero esto es meramente indicativo y otras realizaciones pueden contener cualquier número.

Con referencia a los sistemas descritos anteriormente, luego los flujos multimedia candidatos proporcionan contenido multimedia candidato a decodificar del que se realiza un intento de extraer una marca de agua como se describe a continuación.

El proceso de detección de huellas digitales 1606 como se describe en relación con la Figura anterior identifica los flujos multimedia candidatos que se cree que coinciden con los flujos multimedia de referencia. Estos flujos multimedia candidatos identificados pueden considerarse como flujos multimedia coincidentes. El siguiente proceso se realiza en flujos multimedia coincidentes para determinar si ese flujo lleva una marca de agua. Aquí se apreciará que el sistema de monitoreo descrito anteriormente puede usarse para monitorear flujos multimedia tanto con marca de agua como sin marca de agua.

La salida del proceso anterior proporciona lo que puede considerarse una sincronización temporal aproximada entre el flujo multimedia de referencia y el flujo multimedia candidato. En este caso, puede pensarse que la sincronización temporal identifica qué tan remota está una porción del flujo multimedia candidato desde el comienzo del flujo multimedia de referencia original. La referencia de porción aquí podría ser una trama de un flujo de video u otra porción. Tal alineación temporal es útil en realizaciones que detectan marcas de agua, ya que se apreciará que se aplican diferentes marcas de agua a segmentos de duración limitada de 5 segundos del flujo multimedia. Por tanto, para identificar el código de identidad proporcionado por la secuencia de marca de agua, el sistema identifica al menos los puntos de inicio y finalización de los segmentos de duración limitada dentro de los flujos multimedia coincidentes.

La alineación del flujo multimedia coincidente 1500 contra un flujo multimedia de referencia 204 se muestra en la Figura 15a y 15b. La Figura 15a muestra un flujo multimedia coincidente 1500 que está temporalmente alineado con el flujo multimedia de referencia 204 y se puede ver que los segmentos de duración limitada, cada uno representado por un bloque, tal como 1502a, 1502b, 1504a, 1504b. Como en la Figura 141, el sombreado de cada bloque muestra la marca de agua que se ha aplicado a ese bloque. Se apreciará que el flujo multimedia candidato 1500 no puede dividirse en segmentos de duración limitada del flujo multimedia 204 original, de referencia, cuyos segmentos de duración limitada probablemente se hayan eliminado durante cualquier copia, modificación, etc. de uno de los flujos multimedia entregados 1419. Sin embargo, las regiones con marca de agua probablemente retendrán la sincronización del flujo multimedia con marca de agua y, por lo tanto, en la realización que se describe, aparecerán en bloques de 5 segundos dentro de un flujo multimedia candidato 1500.

La Figura 15b muestra un flujo multimedia coincidente 1506 que no está alineado temporalmente con su flujo multimedia de referencia 204 asociado. Como se analizó anteriormente, la alineación temporal significa que el sistema no ha identificado porciones del flujo multimedia coincidente con respecto a un punto, típicamente el comienzo, del flujo multimedia de referencia 204. Como tal, puede verse que el segmento 1502b no está alineado con el bloque 1502a y el bloque 1504a no está alineado con el bloque 1504b. Por tanto, en el ejemplo de la Figura 15b, si se comparan las marcas de agua, la comparación probablemente fallaría ya que el sistema ha identificado incorrectamente los segmentos de duración limitada con respecto al inicio del flujo multimedia de referencia 204.

En el caso de un flujo multimedia de video, la salida del proceso de toma de huellas digitales anterior también identificará si el flujo multimedia coincidente se ha invertido horizontal y/o verticalmente. Las realizaciones pueden procesar el flujo multimedia coincidente para revertir cualquier cambio detectado antes de detectar la marca de agua dentro del flujo multimedia coincidente.

Es concebible que la alineación temporal del flujo multimedia coincidente del proceso de toma de huellas digitales no sea suficiente. Como tal, al menos algunas realizaciones, incluida la que se describe, realizan una alineación adicional. En particular, se apreciará que la tasa de cambio en algunos flujos multimedia puede ser bastante baja y tales flujos multimedia, especialmente, pueden necesitar una alineación más precisa. Se apreciará que cuanto más precisos puedan alinearse los puntos de inicio y final de los segmentos de duración limitada dentro de un flujo multimedia coincidente con el flujo multimedia de referencia, entonces es más probable que el proceso de extracción de marcas de agua descrito a continuación dé el resultado correcto.

La realización que se describe extrae descriptores de características del flujo multimedia coincidente del mismo formato en el flujo multimedia de referencia 204 original (descriptores de características ORB). A continuación, se realiza un registro de video entre las características extraídas del flujo multimedia coincidente y el flujo multimedia de referencia 204 con el que se ha encontrado una coincidencia. El registro de vídeo identifica una sincronización temporal más precisa entre el flujo multimedia coincidente 1500 y el flujo multimedia de referencia, así como cualquier transformación geométrica (por ejemplo, rotación); escalado; y/o traslación) entre el flujo multimedia coincidente 1500 y el flujo multimedia de referencia 204. El flujo multimedia coincidente 1500 se transforma para revertir cualquier transformación de este tipo, de modo que tenga la misma geometría que el flujo multimedia de referencia 204 original.

A continuación, el proceso identifica las tramas dentro de cada uno de los segmentos de duración limitada (por ejemplo, 1502b, 1504b) y procesa cada una de esas tramas para extraer las marcas de agua que se encuentran dentro de cada trama.

En la realización que se describe, hay dos posibles marcas de agua que podrían haberse aplicado a cualquier trama dada. En otras realizaciones, es posible que se pueda aplicar un número diferente de símbolos de marca de agua, o simplemente la presencia o ausencia de una marca de agua. Como tal, el proceso asigna una creencia a cada una de las posibles marcas de agua y determina que la trama que se está procesando contiene la marca de agua en la que se asigna la creencia más alta.

Para extraer la marca de agua, la trama se transforma en un espacio DCT de bloque como se describe en el proceso de incrustación de la marca de agua. La puntuación de cada bloque DCT viene dada por la correlación sumada entre los componentes DCT y las secuencias aleatorias correspondientes a cada código de marca de agua candidato (como se describió anteriormente). Los puntajes generales para la trama vienen dados por la suma de los puntajes de los bloques DCT individuales.

A continuación, se asigna un valor de marca de agua a cada uno de los segmentos de duración limitada. En la realización que se describe, esta determinación comprende determinar cuál de los dos símbolos de marca de agua tiene la puntuación más alta en cada segmento de duración limitada. Esta determinación se realiza sobre una base de puntuación simple: el símbolo de marca de agua con la puntuación más alta sumada a las tramas que componen el segmento de duración limitada se toma para representar el valor de la marca de agua en ese segmento.

A continuación, se determina 1608 el código de identidad que se ha codificado en el flujo multimedia entregado, del cual se ha derivado el flujo coincidente. Se apreciará desde arriba que el código de identidad está codificado, en esta realización, en 32 segmentos consecutivos de duración limitada. Como tal, la realización coteja las marcas de agua

de 32 segmentos consecutivos de duración limitada. La información de tiempo que el sistema ha almacenado se utiliza para ayudar a garantizar que los puntos de inicio y finalización de la secuencia de segmentos de duración limitada se identifiquen correctamente. El resultado de este proceso es la generación de un número de 32 bits proporcionado por la marca de agua que se ha aplicado a los segmentos de duración limitada del flujo multimedia coincidente. En otras realizaciones, es igualmente posible cualquier otra longitud de número.

Finalmente, el número de 32 bits se compara con códigos de identidad válidos que se han asignado a los usuarios 1420-1424. Si el número de 32 bits corresponde a un código válido, entonces la extracción del código de identidad ha sido exitosa, mientras que si este no es el caso, la extracción del código de identidad ha fallado.

Algunas realizaciones pueden disponerse para extraer continuamente códigos de identidad de un flujo multimedia candidato, o al menos realizar extracciones múltiples del código de identidad de un flujo multimedia candidato, y posteriormente promediar o combinar de otro modo los códigos de identidad. Tales realizaciones pueden ayudar a aumentar la creencia en el código de identidad que se extrae del flujo multimedia coincidente.

Una vez que se ha extraído un código de identidad válido de un flujo multimedia coincidente, se determina la identidad del usuario al que se envió el flujo multimedia entregado 1419a-c, típicamente accediendo a una base de datos de usuarios que contiene el código de identidad asignado a cada uno de ellos. Las realizaciones del sistema pueden entonces tomar la acción apropiada, incluyendo pero sin limitarse a terminar automáticamente el flujo multimedia entregado al usuario identificado (por ejemplo, el usuario 1420).

El proceso de determinar la identidad de un usuario (por ejemplo, el usuario 1420) a partir de un flujo multimedia coincidente se realiza en tiempo real de manera que es posible terminar el flujo multimedia mientras se transmite el flujo multimedia entregado. En la realización que se describe, se apreciará que el proceso de identificación de huellas digitales realizado en el flujo multimedia candidato 219 toma el orden de 1 minuto. Además, el proceso de extracción de la marca de agua tomará el orden de 5 minutos (32 bits, con 10 segundos por bit). Por lo tanto, aquí, en tiempo real, puede significar en el orden de 5 a 10 minutos. Por ejemplo, en tiempo real puede significar sustancialmente 8 minutos.

Como se describió anteriormente, un número de 32 bits se codifica en los flujos entregados a través de la primera y segunda marcas de agua para proporcionar un código de identidad, típicamente, único en cada uno de los flujos entregados. En algunas realizaciones, al menos algunos de los bits del código incrustado se utilizan para proporcionar bits de encabezado y/o de verificación. Como tal, es posible que el código de identidad no sea un número de 32 bits. El experto en la técnica apreciará que el número de bits usados para codificar información en los flujos multimedia entregados 1419a-c puede variar entre realizaciones.

En algunas realizaciones, los bits de verificación pueden proporcionar un código de corrección de errores. Por ejemplo, los bits de verificación pueden permitir cualquiera de los siguientes códigos: un código de Hamming; un código de Reed-Solomon; una verificación de paridad de baja densidad; o similares. El experto en la técnica apreciará que el uso de dichos códigos aumentará la robustez, para las realizaciones que usan el código, con el que el código de identidad puede extraerse del flujo multimedia coincidente.

En algunas realizaciones, el generador de huellas digitales de referencia puede denominarse aparato de monitoreo.

En algunas realizaciones, los mecanismos de procesamiento de identificación 216a-d pueden considerarse como un nodo.

En lo anterior, el o cada componente multimedia puede considerarse como un conjunto de datos.

Como lo apreciará el experto en la técnica, las conexiones 208a-d, 212, 214a-d, o cualquier otra de las conexiones de red a las que se hace referencia en la presente memoria puede ser cualquier forma de conexión de red inalámbrica o por cable. Típicamente, estas conexiones se utilizarán en protocolos TCP/IP, pero no es necesario que sea así.

El proceso también se ejemplifica en relación con las Figuras 7 y 8. La Figura 7 representa el flujo multimedia entrante como una secuencia de tramas; por ejemplo, 300a, 300b, 300c, etc. (etapa 800). Cada trama del flujo multimedia puede considerarse como un elemento de datos del componente multimedia. Otras realizaciones pueden procesar otras formas de datos.

También se realiza un proceso similar dentro del sistema de monitoreo 218 con el fin de generar la secuencia de huellas digitales candidatas CFPn y de nuevo la Figura 7 representa el flujo multimedia candidato 219 como una secuencia de tramas 306a-n.

En la realización que se describe, los generadores de huellas digitales de referencia 202a-d se disponen como un primer paso para eliminar información del flujo multimedia. En la realización que se describe, el generador de huellas digitales de referencia se dispone para convertir las tramas a la escala de grises 802 con el fin de reducir el tamaño de los datos a procesar. Se apreciará que para las realizaciones que se disponen para procesar video en tiempo real,

puede resultar útil reducir la cantidad de datos para aumentar la velocidad a la que se pueden generar huellas digitales de referencia. Ha sido un problema generar huellas digitales (es decir, huellas digitales de referencia) para datos, como un video o transmisiones de video, que se pueden generar en tiempo real y, sin embargo, permiten una detección sólida de transmisiones de multimedia que dan lugar a las mismas, o al menos a huellas digitales similares, (es decir, huellas digitales candidatas). Se cree que al menos algunas de las realizaciones descritas en la presente memoria resuelven este problema.

En realizaciones en las que el componente multimedia (el flujo multimedia en la realización que se describe) es distinto de los datos de video, la calidad puede eliminarse eliminando una o más frecuencias, eliminando el estéreo u otra información multicanal, o similar. Alternativamente, o adicionalmente, otras realizaciones podrían procesar canales de un contenido multimedia dentro de un componente multimedia por separado.

A continuación, los generadores de huellas digitales de referencia 202a-d se disponen para cambiar el tamaño de las tramas 300a-n con el fin de que las tramas 300a-n tengan dimensiones uniformes 804. En la realización que se describe, las tramas se redimensionan para que sean de 32x32 píxeles, reduciendo así la calidad del flujo multimedia. Los generadores de huellas digitales de referencia se disponen además para mantener la relación de aspecto de las tramas de manera que se añaden barras negras en la parte superior/inferior o en los bordes de las tramas para mantener la relación de aspecto de las tramas en el estado redimensionado.

En otras realizaciones, los flujos multimedia se pueden redimensionar para que tengan un tamaño distinto de 32x32 píxeles. Se apreciará a partir del análisis que sigue que las realizaciones que usan un tamaño mayor generarán un mayor número de posibles huellas digitales reduciendo así la posibilidad de falsos positivos pero también que aumenta el procesamiento involucrado para generar las huellas digitales de referencia. En consecuencia, se ha descubierto que las realizaciones que utilizan 32x32 píxeles ofrecen un buen equilibrio entre varios factores, incluido el número de huellas digitales potenciales, el procesamiento requerido para procesar los flujos multimedia en tiempo real (o al menos en tiempo pseudo-real) y el almacenamiento necesario para almacenar el flujo multimedia redimensionado.

En la realización que se describe, una segunda copia de las tramas 300a-n se redimensiona a 16x16 píxeles y se usa para determinar una cantidad de movimiento (es decir, cambio) entre tramas consecutivas 300a-n.

Las tramas de escala de grises redimensionadas se alimentan luego a dos memorias intermedias separadas, que en la realización que se describe son circulares, cada una dispuesta para contener porciones del componente multimedia (etapa 806). Las tramas de escala de grises redimensionadas a 32x32 píxeles se suministran a una de estas memorias intermedias y se utilizan para generar una huella digital y las tramas de escala de grises redimensionadas a 16x16 píxeles se introducen en una segunda de estas y se utilizan para detectar cambios. En la realización que se describe, cada una de las partes alimentadas a las dos memorias intermedias dura 4 segundos y comprende partes del flujo de medios de referencia 204a-d. Ahora se describe la generación de las huellas digitales utilizando las tramas redimensionadas que se encuentran en la primera de las memorias intermedias.

Una vez que se recibe el flujo multimedia 20a4a-d, no ocurre ninguna acción hasta que se llenan las memorias intermedias, pero luego se activa la generación de las huellas digitales de referencia RFPn y se genera una huella digital de referencia a intervalos de medio segundo. Otras realizaciones pueden estar provistas de memorias intermedias de diferente longitud y/o pueden generar huellas digitales de referencia en intervalos distintos de medio segundo. Otras realizaciones pueden tener, por ejemplo, memorias intermedias que son sustancialmente cualquiera de las siguientes longitudes (en segundos); 2, 3, 5, 6, 7, 8, 10, 20, 30, 60. Pueden disponerse otras realizaciones adicionales para generar huellas digitales de referencia sustancialmente en cualquiera de los siguientes periodos (en segundos): 0,1; 0,25; 1; 2; 3; 5; 10;60.

Por tanto, el flujo de tramas almacenado en memoria intermedia comprende una secuencia de tramas que comprende cada una X por Y píxeles. Por tanto, un píxel en (x, y) dentro de una trama se puede ubicar en cada una de las tramas y esto puede pensarse como un tubo de píxeles que atraviesan las tramas dentro de la memoria intermedia. Como tal, la memoria intermedia contiene X por Y tubos de píxeles y en la realización que se describe hay 32 x 32 tubos; 1024 tubos.

Como primer paso en la generación de huellas digitales de referencia, se suaviza cada uno de los tubos de píxeles. En la realización que se está llevando a cabo, esto se logra con un paso de perfil gaussiano 808.

A continuación, el contenido de la memoria intermedia se muestrea a una frecuencia de trama de 8 tramas por segundo, paso 810. Como tal, una memoria intermedia de cuatro segundos contiene 32 tramas de flujo multimedia muestreados. Por lo tanto, ahora hay 32 tramas cada una de 32 por 32 píxeles; un total de 32.768 píxeles.

Por tanto, se puede pensar que cada uno de los generadores de huellas digitales 202 de referencia contiene un reductor de resolución 134 que se dispone para procesar el flujo multimedia entrante u otros datos antes de que ese flujo multimedia se procese como se describe a continuación. Dicho submuestreador 134 puede estar dispuesto para realizar una o más de las etapas de procesamiento descritas anteriormente: conversión del flujo multimedia a escala

de grises; redimensionar el flujo multimedia, que en la realización que se describe era de 32 x 32 píxeles; adición de las barras negras; y una reducción de la velocidad de tramas, que en la realización que se describe era de 8 fps.

A continuación, los generadores de huellas digitales de referencia se disponen para realizar una Transformación de Coseno Discreta (DCT) en los cubos de 32x32x32 (etapa 812); es decir, en esta realización se realiza una DCT tridimensional (3D). Se apreciará que las realizaciones que realizan la DCT en cubos de píxeles de lados iguales son ventajosas en vista de la optimización que puede producirse en la DCT, reduciendo así la cantidad de procesamiento. Como tal, en otras realizaciones, si se alteraran las dimensiones X o Y a las que se redimensionan las tramas 300a-n, entonces también puede ser conveniente alterar el número de tramas a las que se le redujo la resolución del flujo multimedia.

Por tanto, se apreciará que se extraen los componentes de frecuencia del componente multimedia (el flujo multimedia en esta realización). En la realización que se describe, esta extracción del componente de frecuencia la realiza la DCT, pero otras realizaciones pueden usar otros mecanismos para lograr esto. Por ejemplo, otras realizaciones también pueden usar una Transformada de Fourier, una Transformada de Fourier discreta, una Transformada de coseno o similares.

La DCT proporciona una serie de coeficientes, cada uno de los cuales representa un componente de frecuencia del bloque de 4 segundos del flujo multimedia de referencia 204a-d que ha sido procesado. Debido a que la DCT es una DCT 3D y se ha realizado no solo en dos dimensiones (XY de la trama) sino también en la dimensión del tiempo, a través de las tramas, la huella digital de referencia resultante se denomina huella digital espacio-temporal. En la realización que se describe, el 2do y 5to coeficientes, en cada una de las tres dimensiones, se seleccionan para su procesamiento adicional (etapa 814). En otras realizaciones, se puede seleccionar un conjunto diferente de coeficientes, sin embargo, las realizaciones que utilizan el 2do y 5to coeficientes se consideran ventajosas ya que estos coeficientes tienden a capturar el mayor cambio dentro de las tramas del flujo multimedia de referencia 204a-d proporcionando así una identificación sólida del componente multimedia candidato frente al componente multimedia de referencia. Estos componentes de frecuencia extraídos, en la realización que se describe, proporcionan la base de la huella digital que se genera.

A continuación, los generadores de huellas digitales de referencia se disponen para cuantificar los coeficientes seleccionados (en esta realización, estos son el 2do y 5to coeficientes) alrededor del valor medio de los coeficientes seleccionados (paso 816). Por lo tanto, hay coeficientes de 4x4x4, cada uno de los cuales se almacena como un número de un bit. Tal cuantificación da como resultado un número igual de unos y ceros y, en la realización que se describe, se dispone para generar 32 unos y 32 ceros. Las huellas digitales se muestran esquemáticamente en 304 en la Figura 7.

Por tanto, se genera una huella digital de referencia 304 según el contenido de una de las dos memorias intermedias. La otra memoria intermedia se usa para detectar movimiento dentro de la porción de 4 segundos del flujo multimedia como se describe ahora con referencia a la Figura 9. En la realización que se describe, esta reducción se realiza en la frecuencia de trama completa del componente multimedia de entrada, pero en otras realizaciones esto podría realizarse en las tramas de resolución reducida. Para realizaciones que procesan componentes multimedia que no son de video, la segunda memoria intermedia puede disponerse para monitorear la presencia de audio.

Por tanto, se puede pensar que el generador de huellas digitales 202a-d de referencia comprende un generador de huellas digitales 130 que se dispone para realizar los pasos descritos. De forma similar, también se puede pensar que el sistema de monitoreo 218 comprende un generador de huellas digitales 130. Estos generadores de huellas digitales 130 se disponen para generar una secuencia de huellas digitales a partir de una porción de los componentes multimedia candidatos o de referencia; es decir, los flujos multimedia respectivos en la realización que se describe.

Entonces se genera una medición de calidad, que en la realización que se describe puede ser realizada por el evaluador de calidad 132. Como primer paso, se determina un cambio en una propiedad entre al menos algunos de las tramas. En la realización que se describe, el cambio de propiedad es la diferencia de intensidad entre píxeles en tramas vecinas (es decir, se calcula la diferencia de intensidad entre píxeles (x1, y1) en tramas vecinas), paso 900, para generar un cambio de intensidad entre tramas. En la realización que se describe, este cambio de intensidad entre los píxeles vecinos se suma en todos los píxeles de las tramas para cada período entre tramas. En otras realizaciones, se puede sumar un subconjunto de píxeles.

Por lo tanto, se genera una matriz de 31 números, con un número para cada período entre tramas, y se utilizan 32 tramas para generar cada medición de calidad.

A continuación, la matriz se clasifica en orden de magnitud (paso 902) creando así una matriz ordenada que va en un extremo con el valor más alto en un extremo de la matriz al valor más bajo en el otro extremo de la matriz.

A continuación, el generador de huellas digitales de referencia 202 o el sistema de monitoreo 218 se dispone para calcular un percentil predeterminado de la matriz. En la realización que se describe, este percentil predeterminado es el percentil 95 de la matriz (paso 904). Posteriormente, el valor dado por el percentil predeterminado se normaliza por

el número de píxeles (paso 906). Esto proporciona una medición de calidad que da una indicación de la cantidad de movimiento entre tramas que se utiliza durante las comparaciones de huellas digitales realizadas por el sistema 200.

5 Por último, los generadores de huellas digitales de referencia 202a-d se disponen para empaquetar las huellas digitales generadas en el paso 816 junto con las mediciones de calidad y cargadas en el almacén 210 (paso 908). Por tanto, las mediciones de calidad se almacenan en asociación con las huellas digitales. Algunas realizaciones pueden almacenar las huellas digitales y las mediciones de calidad por separado de una manera en la que están vinculadas o asociadas entre sí.

10 En la realización que se describe, las secuencias de huellas digitales de referencia RFPn se cargan en intervalos de minutos y, como tales, se cargan 57 huellas digitales a la vez, no habiendo huellas digitales generadas durante los primeros 3 segundos del flujo multimedia 204a-d hasta que la memoria intermedia se llena. Se asocia con cada una de las huellas digitales de referencia RFPn una marca de tiempo y la medición de calidad derivada para esa huella digital.

15 Se considera que la carga en el almacén 210 en lotes de un minuto proporciona un equilibrio entre la obtención rápida de huellas digitales en el almacén 210 y la reducción de la sobrecarga de la red asociada con la transmisión de pequeñas cantidades de información. Como tal, se pueden disponer otras realizaciones para cargar huellas digitales en el almacén 210 a diferentes intervalos. Por ejemplo, se pueden organizar otras realizaciones para cargar las huellas digitales de forma individual; o prácticamente en cualquiera de los siguientes períodos de tiempo: 10 segundos; 30 segundos; 45 segundos; 90 segundos; 2 minutos; 5 minutos; o más.

25 Cabe señalar que, en la realización que se describe, aunque cada huella digital representa 4 segundos del flujo multimedia 204a, se genera una huella digital mediante el método cada medio segundo. Por lo tanto, la generación de las huellas digitales de referencia 300a-n se puede considerar como si hubiera una ventana 305 movida a través de las tramas del flujo multimedia 204a (es decir, para moverse a través de la referencia o los componentes multimedia candidatos) para tener una pluralidad de posiciones relativas al componente multimedia (es decir, el flujo multimedia en la realización que se describe). Entonces se genera una huella digital para posiciones predeterminadas de la ventana 305. Por tanto, en la realización descrita, la memoria intermedia de cada uno del generador de huellas digitales de referencia 202 y el sistema de monitoreo 218 proporciona generadores de ventana dispuestos para seleccionar respectivamente una porción del componente multimedia de referencia y los componentes multimedia candidatos (es decir, los flujos multimedia) y generar respectivamente una huella digital de referencia o candidata de la porción del componente multimedia cubierto por la ventana.

35 Por tanto, las posiciones de ventana generadas en períodos de tiempo consecutivos pueden considerarse como posiciones de ventana vecinas. Por tanto, se puede ver que las posiciones de ventana 305a en un primer período de tiempo que generó RFP1 es una posición de ventana vecina a la posición de ventana 305b usada para generar huellas digitales RFP2.

40 Por tanto, en la realización que se describe, la información de cualquiera de las tramas de resolución reducida está contenida en varias huellas digitales (por ejemplo, 4 en la realización que se describe). Esto se puede ver en las Figuras 7 y 8, donde se puede ver que la Huella Digital de Referencia 1 (RFPi) contiene información relacionada con las tramas 300a, 300b y 300c. La segunda huella digital de referencia (RFP2) se inicia desde la trama 300c. El experto apreciará que la Figura 7 es una simplificación y que el número de tramas se ha reducido para aumentar la legibilidad de la Figura y que en la realización que se describe, se genera una huella digital (por ejemplo, RFP1) a partir de 32 tramas (después del muestreo descendente).

50 En la realización que se describe, el almacén 210 contiene dos bases de datos: una primera base de datos que contiene huellas digitales de referencia RFPn para flujos multimedia de referencia 204a-d que están actualmente en vivo y una segunda base de datos que contiene huellas digitales de referencia RFPn para flujos multimedia que no están en vivo. Otras realizaciones pueden proporcionar diferentes disposiciones de datos con el almacén 210, pero se cree que es ventajoso proporcionar una base de datos para las huellas digitales asociadas con los flujos multimedia de referencia 204a-d en vivo para garantizar que la base de datos se reduzca en tamaño, aumentando así la velocidad de acceso a los flujos multimedia que se están transmitiendo actualmente (es decir, flujos multimedia en vivo).

55 En la realización que se describe, la o cada base de datos contenida en el almacén 201 (en esta realización dos bases de datos) se replican luego en los mecanismos de procesamiento de identificación 216a-d de manera que cada uno de los mecanismos de procesamiento de identificación mantiene una copia de la, o de cada, base de datos de huellas digitales de referencia RFPn. Se cree que tales realizaciones son ventajosas ya que es altamente escalable y el tamaño del sistema puede aumentarse añadiendo mecanismos de procesamiento de identificación adicionales 216a-d.

60 En la realización que se describe, el almacén 210 y los mecanismos de procesamiento de identificación 216a-d se alojan usando un sistema LINUX y el comando de shell LINUX rsync se usa para replicar la, o cada, base de datos en los mecanismos de procesamiento de identificación.

65

Por lo tanto, típicamente, una vez que uno de los generadores de huellas digitales de referencia 202a-d ha generado una huella digital de referencia 304 (RFPn), se carga en el almacén 201 dentro del orden de segundos después de que haya transcurrido el ciclo de carga del intervalo de minutos. Además, se cree que las realizaciones que utilizan el comando rsync son ventajosas ya que solo replican los cambios en la, o en cada, base de datos del almacén 210 en los mecanismos de procesamiento de identificación 216a-d aumentando así la velocidad y la eficiencia a la que se realiza la carga.

Además de la generación de huellas digitales de referencia a partir de los flujos multimedia de referencia 204a-d, el almacenamiento puede disponerse, en algunas realizaciones, para recibir huellas digitales de referencia de otras fuentes 222. Dichas huellas digitales de referencia pueden incluirse posteriormente en el proceso de análisis descrito a continuación para determinar si los componentes multimedia candidatos corresponden a los componentes multimedia de referencia (por ejemplo, flujos multimedia) que están siendo monitorizados por el sistema 200; los medios u otros datos que tienen una huella digital de referencia almacenada en el almacén 210 pueden considerarse medios (u otros datos) monitoreados por el sistema.

Algunas realizaciones pueden disponerse para monitorear múltiples tipos de contenido multimedia a la vez. Por ejemplo, algunas de las huellas digitales de referencia pueden generarse a partir de uno o más de los siguientes: flujos multimedia; videoclips; transmisiones de audio; clips de audio.

Las huellas digitales de referencia RFPn cargadas en el almacén 210 desde la fuente externa 222 pueden ser representativas de cualquier tipo de componente multimedia que esa realización sea capaz de monitorear identificando de otro modo los contenidos multimedia candidatos.

El sistema de monitoreo 218 se dispone para monitorear una o más fuentes de flujos multimedia candidatos 219. Estas fuentes pueden ser típicamente sitios web y el sistema de monitoreo típicamente puede disponerse para monitorear una pluralidad de dichas fuentes en cualquier momento. Cada flujo multimedia candidato 219 puede considerarse como un componente multimedia candidato; se apreciará que algunas realizaciones pueden monitorizar datos distintos de los flujos multimedia.

Convenientemente, algunas realizaciones se disponen para procesar porciones del flujo multimedia candidato de una longitud predeterminada. En particular, pueden disponerse realizaciones para generar huellas digitales para porciones de sustancialmente un minuto de un flujo multimedia candidato. Se apreciará que si el flujo multimedia se relaciona con algo como un partido de fútbol, entonces será significativamente más largo que esto. Sin embargo, se apreciará que si la realización se dispone para monitorear e intentar prevenir la transmisión de flujos multimedia no autorizados, entonces es deseable detectar la transmisión de ese flujo multimedia tan pronto como sea posible y, como tal, no es deseable esperar hasta que la transmisión del flujo multimedia haya terminado antes de generar huellas digitales y, como tal, es ventajoso procesar los flujos multimedia candidatos en porciones de la misma manera que es ventajoso generar huellas digitales para porciones de los flujos multimedia de referencia. Sin embargo, tampoco es deseable generar demasiado tráfico de red enviando cantidades excesivas de huellas digitales y se considera que proporcionar realizaciones que procesen porciones del flujo multimedia candidato en porciones de sustancialmente de 60 segundos proporciona un equilibrio. De nuevo, otras realizaciones pueden utilizar un período de tiempo distinto de 60 segundos.

Durante el uso, el sistema de monitoreo 218 se dispone para generar, usando el mismo método que se describe en las Figuras 8 y 9, una secuencia de huellas digitales candidatas CFPn para el o cada flujo multimedia candidato 219. Es decir, se utiliza el mismo método para generar la huella digital candidata CFPn que se utiliza para generar la huella digital de referencia. Por lo tanto, se generan una o más huellas digitales candidatas CFPn que son iguales, o al menos similares, a una o más huellas digitales de referencia RFPn, siempre que el flujo multimedia candidato subyacente a partir del cual se generan las huellas digitales candidatas CFPn (u otros datos) sea el mismo, o al menos similar, a un flujo multimedia de referencia que generó la huella digital de referencia RFPn. Por lo tanto, las realizaciones del sistema 200 podrían encontrar aplicación para determinar si los flujos multimedia candidatos corresponden a los flujos multimedia de referencia en vivo 204a-d, lo que en sí mismo es útil para determinar e intentar detener la piratería en línea y detectar copias de los flujos multimedia de referencia 204 monitoreados por el sistema.

Otras realizaciones pueden encontrar otra aplicación. Por ejemplo, el almacén 210 puede proporcionar una librería en la que las huellas digitales candidatas CFPn generadas a partir de un componente multimedia candidato pueden identificarse frente a las huellas digitales de referencia RFPn contenidas en el almacén 210. Dichas realizaciones permitirían, por lo tanto, que un componente multimedia (tal como un archivo de vídeo; una imagen, un archivo de audio o similar) se verificara con el almacén 210 en un intento de identificar ese componente multimedia candidato como una coincidencia con los componentes multimedia a los que se hace referencia en el almacén 210.

La generación de las huellas digitales candidatas se representa de nuevo esquemáticamente en la Figura 7, que muestra un flujo multimedia candidato 219 entrante representado como una secuencia de tramas 306a, 306b, 306c, 306d. Como es el caso con la generación de las huellas digitales de referencia, una ventana móvil 308 se mueve a lo largo del flujo multimedia 219 y se genera una Huella digital Candidata CFP<sub>1</sub> - CFP<sub>n</sub> para posiciones predeterminadas de la ventana 308. Al igual que con la generación de huellas digitales de referencia, en la realización que se describe, se genera una huella digital sustancialmente cada medio segundo, aunque cada huella digital cubre 4 segundos de

las tramas 306a-n. Como se ha descrito anteriormente, una memoria intermedia circular proporciona un generador de ventanas diseñado para proporcionar la ventana necesaria para recibir y procesar una parte del componente multimedia.

5 Al igual que con la generación de las huellas digitales de referencia, las huellas digitales candidatas se agrupan antes de cargarse en la cola 220. De nuevo, las huellas digitales candidatas, una marca de tiempo y la medición de calidad asociada se cargan en grupos de 57 huellas digitales candidatas que representan la longitud de un minuto del flujo multimedia candidato 219 que se ha procesado.

10 Como se describió anteriormente, el sistema genera una cola 220 de secuencias de huellas digitales candidatas CFPn para ser comparadas con las huellas digitales de referencia RFPn dentro del almacén 210. Además de las huellas digitales candidatas CFPn generadas por el sistema de monitoreo 218, se pueden añadir huellas digitales candidatas, en al menos algunas realizaciones, a la cola 220 desde una fuente externa 224 capaz de generar huellas digitales candidatas CFPn.

15 A continuación, se accede a las secuencias en cola de huellas digitales candidatas y se eliminan de la cola 220 mediante uno de los mecanismos de procesamiento de identificación 216a-dy ese mecanismo de procesamiento de identificación busca luego en las huellas digitales de referencia contenidas en el almacén 210 para la aparición de una huella digital de referencia RFPn correspondiente a una o más huellas digitales candidatas CFPn de la secuencia de huellas digitales candidatas CFPn. Esta búsqueda de huellas digitales candidatas CFPn intenta localizar huellas digitales de referencia RFPn coincidentes dentro del almacén 210 identificando así el componente multimedia (por ejemplo, flujo multimedia) que generó la huella digital candidata CFPn como el mismo componente multimedia (por ejemplo, flujo multimedia) que generó la huella digital de referencia RFPn.

25 Cada uno de los mecanismos de procesamiento de identificación 216a-d mantiene una copia completa de la, o de cada base de datos del almacén 210. En la realización que se describe, esta copia se mantiene dentro de la memoria de acceso aleatorio (RAM) del mecanismo de procesamiento de identificación 216a-d. Las realizaciones con una disposición de este tipo son ventajosas debido a la velocidad de procesamiento que esto permite. Sin embargo, otras realizaciones pueden contener la copia de la, o de cada, base de datos dentro de cualquier forma de almacenamiento al que pueda acceder el circuito de procesamiento del mecanismo de procesamiento de identificación 216a-d, ya sea en red o local al circuito de procesamiento.

30 La copia de la o de cada base de datos dentro del mecanismo de procesamiento de identificación 216a-d se ve obligada de vez en cuando a actualizarse para reflejar cualquier cambio dentro de la(s) base(s) de datos en el almacén 210. La actualización puede, por ejemplo, obligarse a que se produzca sustancialmente en el orden de minutos, tal vez sustancialmente en el orden de cualquiera de los 30 segundos siguientes; cada minuto; cada 5 minutos; cada 10 minutos; en cualquier momento entre estos tiempos.

35 Una vez que el mecanismo de procesamiento de identificación 216a-d ha obtenido una secuencia de huellas digitales candidatas CFPn de la cola 220, este busca al menos algunas, y típicamente cada una de las bases de datos a las que tiene acceso para determinar si las huellas digitales candidatas CFPn de la cola coinciden con alguna de las huellas digitales de referencia RFPn que se encuentran dentro de las bases de datos que se buscan. Aunque las huellas digitales se cargan en la cola 220 en grupos relacionados con porciones del flujo multimedia candidato (como una porción de un minuto de duración), cada una de las huellas digitales de 64 bits relacionadas con 4 segundos del flujo multimedia candidato se busca en la o en cada base de datos.

40 Con el fin de aumentar la seguridad de una coincidencia y reducir las posibilidades de un falso positivo y/o un falso negativo, se pueden agregar huellas digitales como se describe a continuación. La agregación puede tomarse como una característica que se evalúa para determinar si una huella digital candidata dentro de una secuencia coincide con una huella digital de referencia dentro de una secuencia.

45 La distancia de Hamming entre dos secuencias de bits es igual al número de bits que difieren entre las dos secuencias. En el caso de una huella digital de 64 bits, una distancia de Hamming de 0 es una coincidencia perfecta, mientras que una distancia de Hamming de 32 muestra una falta total de correlación. En la realización que se describe, debido a que una huella digital siempre contiene una cuenta igual de elementos 1/0, la distancia de Hamming es siempre un número par. Otras realizaciones pueden generar distancias de Hamming no uniformes.

50 Las pruebas con datos de ejemplo sugieren que un umbral de distancia de Hamming de aproximadamente 12-16 es adecuado para identificar coincidencias positivas entre huellas digitales individuales para realizaciones del sistema descrito en la presente memoria.

55 Se apreciará que en la realización que se describe, la distancia de Hamming se usa como una medida para determinar si una huella digital de referencia es similar, o la misma, a una huella digital candidata. En otras realizaciones, se pueden utilizar diferentes técnicas para realizar esta comparación.

60

65

Por tanto, cada uno de los mecanismos de procesamiento de identificación 216a-d se dispone para hacer una comparación de una huella digital candidata obtenida a partir de una secuencia de huellas digitales de la cola 220 con todas las huellas digitales de referencia RFPn de las que tiene conocimiento el sistema. Esta comparación la realiza típicamente el comparador 128 y se utiliza para localizar huellas digitales de referencia RFPn que coinciden con al menos una de las huellas digitales candidatas CFPn.

El mecanismo de procesamiento de identificación 216a-d se dispone así para comparar cada huella digital individual (por ejemplo, cada número de 64 bits en la realización que se describe) en la huella digital candidata CFPn que está procesando con cada una de las huellas digitales de referencia RFPn dentro del almacén 210. El mecanismo de procesamiento de identificación 216a-d se dispone además para acumular pares de huellas digitales candidatas y de referencia que coinciden donde se determina una coincidencia si:

1. La distancia de Hamming entre las huellas digitales candidatas y de referencia es inferior o igual a un umbral predeterminado, que es la realización que se describe en 14; y

2. La medición de calidad de las huellas digitales tanto candidatas como de referencia es superior o igual a un umbral predeterminado, que en la realización que se describe es 3,0.

Por lo tanto, en la realización que se describe, la distancia de Hamming y/o la medición de calidad pueden tomarse como una característica que puede evaluarse para determinar si una huella digital candidata dentro de una secuencia coincide con una huella digital de referencia dentro de una secuencia

Esta coincidencia entre las huellas digitales candidatas y de referencia (RFPn) se ejemplifica en la Figura 11a en la que una secuencia de huellas digitales de referencia generadas a partir de un flujo multimedia de referencia 204a se muestra encima de una secuencia de huellas digitales candidatas (CFPn) generadas a partir de un flujo multimedia candidato 219.

Por lo tanto, tomando la huella digital CFP1 de la secuencia como ejemplo, entonces se busca la CFP1 de entre todas las huellas digitales de referencia almacenadas en el almacén 210. Las flechas entre las huellas digitales de referencia RFPn y las huellas digitales candidatas CFPn representan las huellas digitales que se considera que coinciden según los criterios anteriores. Por tanto, en el ejemplo que se da, se puede ver que CFP1 coincide con RFP2; RFP4 y RFP5.

Se recuerda que en la realización que se describe, los flujos multimedia 204a y 219 son flujos de video y, como tales, es probable que las tramas vecinas dentro de los flujos de video no cambien significativamente. Esta baja cantidad de cambio entre tramas es particularmente el caso de periodos en los que hay poco movimiento en el flujo de video. Es probable que se observen efectos similares en flujos de otros tipos de datos.

La similitud entre las tramas puede resultar en múltiples coincidencias de huellas digitales dentro de un flujo. Además, los flujos de multimedia que tienen un contenido similar también pueden resultar en coincidencias en las huellas digitales representativas de esas transmisiones. En un ejemplo, una secuencia similar de movimientos de cámara (por ejemplo, panorámica y zoom y cambio de toma) dentro de dos partidos de fútbol diferentes (u otro contenido) puede generar secuencias similares de huellas digitales. Puede haber coincidencias coincidentes con huellas digitales en otras transmisiones.

También se apreciará que las realizaciones pueden no tener una alineación temporal entre las huellas digitales generadas a partir de los flujos candidato y de referencia. De hecho, es probable que no se conozca la relación temporal entre un flujo multimedia candidato 219 y un flujo de referencia 204a. Por lo tanto, al observar CFP3, se puede ver, en este ejemplo, que se ha comparado con RFP1, RFP6 y RFPn. Se observará que RFP1 está antes de cualquiera de las huellas digitales de referencia (RFPn) con las que se ha comparado CFP1, lo que hace menos claro cómo las huellas digitales candidatas (CFPn) se alinean temporalmente con las huellas digitales de referencia (RFPn), asumiendo que las huellas digitales representadas en las Figuras 11a y 11b representan de hecho el mismo flujo multimedia.

Es probable que la Figura 11a sea una simplificación en cuanto al número de coincidencias entre las huellas digitales candidatas CFP1 y las huellas digitales de referencia RFP1 y también a la naturaleza dispersa ya que, como se analizó anteriormente, pueden coincidir con otros flujos multimedia.

Algunas realizaciones pueden disponerse para añadir una marca de tiempo a cada huella digital que se genere. Además, tales realizaciones pueden disponerse para limitar las huellas digitales de referencia con las cuales se comparan las huellas digitales candidatas con las que se produjeron dentro de un tiempo predeterminado de la huella digital candidata. Tales realizaciones pueden permitir una búsqueda más rápida de las huellas digitales de referencia y también pueden ser útiles cuando el componente multimedia representa un flujo multimedia de Internet de un evento en el que es probable que se produzcan flujos copiados dentro de un período de tiempo corto del flujo multimedia original.

Por tanto, los mecanismos de procesamiento de identificación 216a-d generarán un conjunto de coincidencias entre las huellas digitales candidatas CFPn y las huellas digitales de referencia RFPn. En la realización que se describe, que monitoriza un flujo multimedia, cada coincidencia corresponde a un segmento individual de cuatro segundos de un flujo multimedia de referencia 204a-d que coincide con un segmento de 4 segundos del flujo multimedia candidato 219; se recuerda que se genera una única huella digital de referencia o candidata a partir de una duración de 4 segundos de un flujo multimedia y esto se muestra en la Figura 12.

La Figura 12 representa un flujo multimedia (ya sea un flujo multimedia de referencia o candidato) mediante una línea y se puede ver que los flujos están marcados en intervalos largos de segundos desde 0 segundos hasta 10 segundos. A continuación, se muestran nueve huellas digitales debajo del flujo multimedia 204, 219 (FP1 a FP9). La primera de estas huellas digitales FP1 se extiende desde  $t=0$  hasta  $t=4$ ; el segundo FP2 desde  $t=1$  hasta  $t=5$ ; el tercer FP3 desde  $t=2$  hasta  $t=6$ ; etc. Por lo tanto, la primera huella digital FP1 se genera mediante una posición de ventana que se extiende a través del componente multimedia (es decir, el flujo multimedia) desde  $t=0$  a  $t=4$ ; se produce una segunda posición de ventana entre  $t = 1$  y  $t = 5$  y, por lo tanto, existe la diferencia de tiempo de 1 segundo entre las posiciones de ventana analizadas en la presente memoria.

Se puede ver que se requieren 5 huellas digitales para cubrir y un intervalo de 8 segundos. Se cree que la naturaleza superpuesta de las huellas digitales es una ventaja para disminuir la importancia de un cambio en el flujo multimedia dentro de cualquier huella digital FPn. Algunas realizaciones del sistema deben tener una región de 6 huellas digitales coincidentes (es decir,  $M = 6$  como se describe a continuación) que corresponde a las 5 huellas digitales necesarias para dar 8 segundos de cobertura pero con una huella digital adicional que se cree que aumenta la robustez.

En otras realizaciones, se ha encontrado que la diferencia de tiempo entre las posiciones de las ventanas utilizadas para generar las huellas digitales es ventajosamente de 0,5 s. Se apreciará que la reducción de la diferencia de tiempo entre las posiciones de las ventanas aumenta la fidelidad con la que se puede representar el comportamiento de la huella digital a lo largo del tiempo (ya sea candidata o de referencia). Como tal, y con referencia a la realización que se describe, una diferencia de tiempo menor permite que el sistema responda a cambios más rápidos dentro de los flujos multimedia 204, 219.

Por tanto, en la realización que se describe, el generador de ventana se dispone de manera que porciones del componente multimedia generadas por las posiciones vecinas de la ventana se superponen. Por supuesto, pueden disponerse otras realizaciones de modo que el generador de ventanas se disponga para generar posiciones de ventanas vecinas que no se superpongan.

En particular, se apreciará que en la realización que se describe, se usa una DCT para generar las huellas digitales (ya sean candidatas o de referencia) que extraen información de frecuencia del flujo multimedia. Se ha descubierto que las realizaciones que utilizan las frecuencias más bajas son ventajosas para generar coincidencias sólidas en la comparación entre las huellas digitales candidatas y de referencia. Sin embargo, para capturar esas frecuencias más bajas se requiere un período de tiempo más largo (en comparación con las frecuencias más altas). Se ha descubierto que un período de tiempo sustancialmente de 4 segundos proporciona suficiente información de baja frecuencia para proporcionar coincidencias sólidas mientras que el uso de ventanas de tiempo superpuestas permite una frecuencia de muestreo razonablemente alta. Se apreciará que una frecuencia de muestreo más alta aumenta la capacidad de generar huellas digitales precisas para un flujo de video que cambia rápidamente.

Además, como puede verse en la Figura 12, una vez que se establece el flujo, cualquier intervalo de un segundo está cubierto por 4 huellas digitales que corresponden a la longitud de cuatro segundos de la huella digital. Por tanto, en otras realizaciones en las que la longitud del flujo multimedia cubierto por una huella digital FP es distinta de cuatro segundos, entonces un número diferente de huellas digitales cubriría cada intervalo de tiempo. Se cree que la naturaleza superpuesta de las huellas digitales es ventajosa y permite que el proceso de agregación que se describe a continuación en relación con la Figura 11b omita las huellas digitales y aún tenga la confianza de que se ha encontrado una coincidencia entre un flujo multimedia candidato con un flujo de referencia mantenido dentro del almacén 210 ya que los datos para ese período todavía se están igualando.

Es concebible que las coincidencias entre huellas digitales candidatas (CFPn) y de referencia (RFPn) puedan ser falsos positivos. También es concebible que un segmento de 4 segundos del flujo multimedia candidato 219 que debería haber producido una coincidencia (es decir, coincide con un flujo multimedia de referencia) no haya dado como resultado una coincidencia (es decir, un falso negativo). Como se analizó anteriormente, si los datos se relacionan con un flujo multimedia, como un video, entonces es probable que las tramas vecinas del flujo multimedia (por ejemplo, 300a-d) sean similares entre sí y, por lo tanto, es probable que las huellas digitales, ya sean candidatas o de referencia, que se generan a partir de regiones de tiempo similares del flujo multimedia, sean similares entre sí.

Por consiguiente, cada uno de los mecanismos de procesamiento de identificación se dispone además para agregar un número de coincidencias como se describe ahora con referencia a las Figuras 10 y 11b. Como punto de partida y como se describe en relación con la Figura 11a, se busca en el almacén 210 para localizar cada Huella digital de referencia RFPn que coincida con una Huella digital candidata CFPn dada. A continuación, se lleva a cabo un proceso de agregación para proporcionar una identificación más sólida de un flujo multimedia candidato 219 frente a un flujo

5 multimedia de referencia 204a (u otro componente multimedia) para el cual se mantienen las huellas digitales candidatas RFPn dentro del almacén 210. Como se describió anteriormente, una coincidencia entre una única huella digital de referencia RFPn y las huellas digitales candidatas CFPn no identifica necesariamente una coincidencia entre un flujo multimedia candidato 219 y un flujo multimedia de referencia 204a, o entre cualquier otro componente multimedia candidato y componente multimedia de referencia

10 Como se describe en relación con la Figura 11a, es probable que haya al menos algunas coincidencias falsas con cualquier huella digital candidata CFPn que se esté procesando. En consecuencia, se cree que las realizaciones que implementan el proceso de agregación son ventajosas ya que aumentan la robustez de la identificación del flujo multimedia candidato 219 (u otro componente multimedia) dentro de los flujos multimedia de referencia 204a (u otro componente multimedia) que han generado las huellas digitales de referencia dentro del almacén 210. Esta mayor robustez proporcionada por el proceso de agregación puede permitir, en algunas realizaciones, tomar una huella digital menos robusta del componente multimedia, lo que puede permitir tomar una huella digital más rápidamente.

15 Por lo tanto, el proceso de agregación identifica las regiones 1002 de las huellas digitales candidatas CFPn generadas a partir del flujo multimedia candidato 219 que se asignan a regiones de huellas digitales RFPn generadas a partir del flujo multimedia de referencia 204a y utilizan la relación temporal para aumentar la creencia de que existe una coincidencia entre una huella digital candidata CFPn y una Huella digital de referencia RFPn.

20 Por lo tanto, al observar la Figura 11a se puede ver que existen coincidencias entre las siguientes huellas digitales dentro de la secuencia de huellas digitales que se muestra en la Figura:

CFP2 y RFP5;

25 CFP3 y RFP6;

CFP4 y RFP7; y

30 CFP6 y RFPn.

También debe tenerse en cuenta que en el ejemplo dado, no se encontró que CFP5 coincidiera con la huella digital de referencia RFP8.

35 La Figura 11b muestra cómo el proceso de agregación determina que se ha producido una coincidencia entre el flujo multimedia candidato 219 y el flujo multimedia de referencia 204a.

40 El proceso de agregación busca cualquier región de huellas digitales candidatas CFPn dentro de la secuencia que tenga múltiples coincidencias y cualquier región que pase el proceso de agregación se informa como una coincidencia positiva. En una realización, esto puede significar que el flujo multimedia candidato 219 coincide con un flujo multimedia de referencia 204a-d que está siendo monitoreado por el sistema 200. En otras realizaciones, esto puede significar que un componente multimedia candidato se ha identificado frente a huellas digitales de referencia que representan un componente multimedia de referencia que se mantiene dentro del almacén 210. Como se describió anteriormente, se informa una coincidencia positiva si se mantiene un número mínimo de huellas digitales candidatas en una secuencia para hacer coincidir las huellas digitales con la secuencia de huellas digitales de referencia.

45 En el contexto del párrafo anterior, una región de huellas digitales puede comprender cualquier número de huellas digitales consecutivas y puede comprender, por ejemplo, sustancialmente 2, 3, 4, 5, 6, 7, 8, 9, 10, 15, 20, 25, 30 o más huellas digitales consecutivas.

50 Se dice que una secuencia de huellas digitales candidatas es una coincidencia global, agregada de una secuencia de huellas digitales de referencia si, para el mismo número de huellas digitales de referencia y candidatas (es decir, el mismo período de tiempo) - paso 1004:

55 1. La secuencia tiene al menos M huellas digitales de longitud, donde M en la realización que se describe se establece en 6;

2. La secuencia comienza con una coincidencia positiva de huella digital candidata/de referencia;

60 3. No hay ninguna brecha entre coincidencias positivas de candidatas/de referencia superiores a D, en donde la realización actual D se establece en 3. Esta prueba asegura que haya menos de un número predeterminado de brechas dentro de la secuencia de huellas digitales. En la realización que se describe, esto equivale a que es permisible no tener coincidencias entre las huellas digitales candidatas y de referencia de hasta 3 segundos, mientras que un período de 4 segundos daría como resultado una falla para cumplir la prueba de agregación.

65 Poniendo esto en contexto con la Figura 11b, se observa que CFP2 es el comienzo del bloque 704 de coincidencias (identificado por la marca en CFP2) que se asigna a RFP5. El bloque 704 contiene 9 huellas digitales candidatas

(CFP2 a CFP10) que incluyen 7 coincidencias (CFP2); CFP3; CFP4; CFP6; CFP8; CFP9; y CFP10) y dos errores (CFP5 y CFP7), es decir, separaciones entre coincidencias vecinas. Por consiguiente, este bloque 704 cumple con los criterios que se han establecido y se considera que se ha producido la coincidencia entre el flujo multimedia candidato 219 y el flujo multimedia de referencia 204a ya que se han cumplido los puntos 1 a 3 anteriores. Se observará que, por razones de claridad, solo se muestran algunos de los vínculos entre las huellas digitales candidatas (CFP10) y las huellas digitales de referencia FRPn.

Por tanto, en la realización que se describe, el bloque 704 cumple los criterios y se considera que se ha producido una coincidencia entre ese bloque del flujo multimedia candidato 219 y el flujo multimedia de referencia 204a; etapa 1006.

Por tanto, en resumen de lo anterior, el comparador puede utilizar varias características para evaluar si la secuencia de huellas digitales de referencia coincide con la secuencia de huellas digitales candidatas. En la realización que se ha descrito, estas características pueden incluir características sobre las coincidencias entre huellas digitales individuales así como características sobre la comparación de la secuencia de huellas digitales (por ejemplo, las secuencias de cuatro segundos).

Las características de las coincidencias entre huellas digitales individuales pueden incluir al menos una de las siguientes:

- una comparación de la distancia de Hamming entre las 2 huellas digitales.
- una comparación de la calidad de la huella digital candidata con un umbral de calidad.
- una comparación de la calidad de la huella digital de referencia con un umbral de calidad.

Por lo tanto, si alguno de estos no se cumple, se considera que la huella digital candidata no coincide con la huella digital de referencia.

Las características de las coincidencias entre secuencias de huellas digitales pueden incluir al menos una de las siguientes:

- si un número umbral de huellas digitales agregadas coincide en la secuencia.
- si en la secuencia se supera el umbral más grande de "brecha" de huellas digitales que no coinciden.
- si se alcanza un umbral de distancia de Hamming medio a lo largo de la secuencia.
- si se cumple una diferencia de calidad entre las huellas digitales candidatas y las huellas digitales de referencia en toda la secuencia.

Por tanto, si no se alcanza alguno de estos umbrales de nivel de secuencia, la realización determinará que la secuencia candidata no coincide con la secuencia de referencia.

Para intentar evitar que el sistema sea eludido simplemente por manipulaciones del flujo multimedia candidato, el flujo multimedia se refleja tanto horizontal como verticalmente y cada orientación pasa a través del proceso de toma de huellas digitales. Por lo tanto, en la realización que se describe, cada flujo multimedia candidato pasa por el proceso de toma de huellas digitales 4 veces: sin reflejar; reflejado verticalmente; reflejado horizontalmente; y reflejado tanto horizontal como verticalmente.

En otras realizaciones, los componentes multimedia candidatos pueden sufrir otro procesamiento para intentar evitar la anulación del sistema.

Los mecanismos de procesamiento de identificación 216a-d se disponen entonces para informar de cada región que pasa el proceso de agregación como una coincidencia positiva entre una región del flujo multimedia candidato y del flujo multimedia de referencia.

En una realización, el sistema descrito en relación con las Figuras anteriores puede usarse en el proceso descrito en relación con la Figura 13 que se describe ahora.

Como primer paso 1, las huellas digitales de referencia RFPn se generan y almacenan en el almacén 210 como se describió anteriormente. Además, como se describió anteriormente, el sistema puede disponerse para monitorear muchos tipos de datos pero, en la realización que se describe, está dirigido a monitorear varios tipos de componentes multimedia: particularmente uno o más de los siguientes: archivos de vídeo; vídeo transmitido; archivo de audio; archivos de audio transmitidos.

Las realizaciones que monitorean los componentes multimedia distintos de los archivos de video o audio típicamente utilizarán un algoritmo de huellas digitales distinto al basado en una DCT descrita en la presente memoria.

5 En el paso 2, un usuario reproduce una versión inicial de un componente multimedia, por ejemplo, reproduciendo el componente multimedia en una página web o similar. Si esta versión inicial del componente multimedia ha sido procesada por el sistema y tiene una huella digital de referencia RFPn almacenada en el almacén 210, puede ser identificada por el sistema. En la Figura 13, se muestra que el paso 2 se origina en el mismo origen que el paso 1, lo que implica que el componente multimedia inicial, reproducido por el usuario, se origina en la misma fuente que el componente multimedia a partir del cual se genera la huella digital de referencia RFPn. Sin embargo, no es necesario que este sea el caso y, en otras realizaciones, un usuario puede obtener y reproducir el componente multimedia inicial desde cualquier fuente.

15 En el paso 3, el componente multimedia inicial que está siendo reproducido por un usuario se trata como un componente multimedia candidato y, por lo tanto, tiene las huellas digitales candidatas CF-Pn generadas para él como se describió anteriormente. Se apreciará que pueden generarse huellas digitales candidatas CFPn para el componente multimedia independientemente del origen del componente multimedia. El método se dispone entonces para determinar, como se describió anteriormente y en particular con referencia a las Figuras 11a y 11b, si las huellas digitales candidatas CFPn generadas de este modo corresponden a alguna de las huellas digitales de referencia RFPn almacenadas en el almacén 210; es decir, se realiza una búsqueda de la huella digital candidata comparándola con el conjunto de huellas digitales de referencia 210. El experto apreciará que si la secuencia de huellas digitales candidatas CFPn se identifica como correspondiente a una secuencia de huellas digitales de referencia RFPn, se supone que el componente multimedia candidato es el mismo que el componente multimedia de referencia (es decir, el archivo que generó la huella digital de referencia) y así, se ha identificado el componente multimedia candidato.

25 En un cuarto paso del método mostrado en la Figura 13, el circuito de procesamiento que un usuario está usando para reproducir el componente multimedia inicial se dispone para conectarse a una fuente desde la cual se puede obtener una versión mejorada del archivo multimedia inicial. Se apreciará que tal método se puede hacer que funcione con una variedad de dispositivos de procesamiento y, por ejemplo, un usuario puede ver el archivo multimedia inicial en cualquiera de los siguientes: un ordenador personal (tal como una PC con Windows o LINUX u otro sistema operativo; un MAC; un teléfono móvil (tal como un iPhone, un dispositivo Android, un móvil Windows, una Blackberry o similares); un televisor inteligente; una tableta (tal como un iPad, un Kindle, un Android o similar), etc.

35 La fuente del componente multimedia inicial puede entonces disponerse para suministrar una versión del componente multimedia inicial al dispositivo de procesamiento que está utilizando el usuario. La versión del componente multimedia inicial puede tener un valor agregado en comparación con la versión inicialmente reproducida por el usuario.

El valor añadido puede incluir cualquiera de los siguientes elementos: mayor calidad; sonido mejorado; un comentario; subtítulos; diferentes ángulos de cámara; más pistas de audio; escenas ocultas o similares.

40 El sistema puede disponerse para cargar una cuenta, o cobrar de otro modo, al usuario por el suministro del componente multimedia mejorado; por ejemplo, el componente multimedia con el valor añadido.

45 Además, es conveniente describir el sistema en relación con el monitoreo de flujos de video como el mecanismo para entregar contenido multimedia a un usuario. Sin embargo, en otras realizaciones, las técnicas descritas pueden aplicarse igualmente a otras formas de componente multimedia. Por ejemplo, los archivos de video, tales como los archivos MP4 y similares, pueden considerarse componentes multimedia. En otras realizaciones alternativas o adicionales, los componentes multimedia pueden ser proporcionados por contenido de audio tal como contenido de audio transmitido o archivos de audio discretos (tales como archivos MP3, la pista de audio de un video o similares). El experto en la técnica apreciará cómo modificar los algoritmos de marcas de agua y huellas digitales para procesar datos de audio en lugar de video. Por lo tanto, la referencia a la transmisión en el texto siguiente no debe considerarse como limitada a la transmisión de multimedia y también se pretende que se refiera a los mecanismos de entrega no transmitidos.

55 Se hace referencia a lo largo de la siguiente descripción a la Figura 16 que ilustra el proceso general que se emplea, por las realizaciones, para monitorear y/o proteger los componentes multimedia.

60 El sistema 1400 que se muestra en la Figura 14, que puede considerarse como una plataforma de distribución de contenido multimedia, permite tomar las huellas digitales de un flujo multimedia de referencia 204 (es decir, un componente multimedia) para que pueda identificarse como se describe a continuación. En las figuras anteriores se proporcionan más detalles sobre cómo se toma una huella digital.

65 Además, el sistema 1400 permite que el flujo multimedia de referencia 204 se entregue a una pluralidad de usuarios de manera que la versión entregada a cada uno de los usuarios tenga una marca de agua única incorporada en el mismo. Esta marca de agua única permite que el usuario al que se entregó una versión determinada del flujo multimedia de referencia se identifique como se describe a continuación.

Además, el sistema 1400 permite que una red, que en la realización que se describe es Internet y/o la World Wide Web 1402, sea monitoreada en busca de copias u otras versiones ilícitas (en lo sucesivo citadas como 'copias'), del flujo de contenidos multimedia de referencia 204. La red puede ser o no la misma que la red 251 descrita anteriormente.

5 Si se detectan copias, el sistema se dispone a determinar, utilizando la marca de agua única, el usuario al que se entregó esa versión del flujo multimedia de referencia. El sistema de monitoreo también se describe con más detalle en relación con las figuras anteriores.

10 Por lo tanto, con referencia a la Figura 14, se puede ver que el flujo multimedia de referencia 204 pasa a través de un proceso de toma de huellas digitales 1600 proporcionado en un generador de huellas digitales de referencia 202 que genera una huella digital de referencia RFP que se puede utilizar para determinar si un flujo multimedia candidato coincide con un flujo multimedia de referencia 204 monitoreado por el sistema. Las huellas digitales generadas por el generador 202 se mantienen dentro de un almacén 210. Los componentes del sistema de toma de huellas digitales se describen con más detalle en la Figura 2 y los elementos de la Figura 14 descritos con más detalle en la Figura 2 se muestran dentro de la línea de cadena 1403 y las partes similares se denominan con números de referencia similares.

20 El flujo multimedia con huellas digitales 1406 pasa por un proceso de marca de agua 1407 que genera al menos dos versiones del flujo de huellas digitales 1406: una primera versión con marca de agua 1408; y una segunda versión 1410 con marca de agua. Mientras que, en la realización que se describe, solo hay dos versiones con marca de agua del flujo multimedia de referencia, otras realizaciones pueden proporcionar más de dos versiones.

25 Como se describe a continuación, se aplica una marca de agua a cada trama de la versión con marca de agua 1408, 1410 y se aplica la misma marca de agua a cada trama dentro de la misma versión del flujo multimedia con marca de agua. Las versiones con marca de agua se almacenan luego dentro de un sistema de entrega 1412.

30 En la realización que se describe, se aplica una marca de agua diferente a las tramas de cada una de la primera versión con marca de agua 1408 y la segunda versión con marca de agua 1410. Al menos algunas realizaciones, incluida la que se describe, usan una marca de agua ortogonal de modo que la marca de agua aplicada al primer flujo con marca de agua sea lo más distinguible posible del segundo flujo con marca de agua. En un ejemplo, se puede usar una secuencia como 1,2,-1,3,-2... para generar la marca de agua para el primer flujo. En la realización que se describe, la secuencia ortogonal -1,-2, 1, -3, 2... se usa entonces para generar la segunda marca de agua.

35 En otras realizaciones, es posible que sólo el segundo flujo con marca de agua 1410 tenga una marca de agua aplicada a las tramas de la misma y el primer flujo con marca de agua 1406 siga siendo una copia del flujo multimedia con huellas digitales 1406 sin que se le aplique ninguna marca de agua. Las realizaciones que aplican marcas de agua solo a algunos de los flujos multimedia pueden ser ventajosas en vista del procesamiento reducido requerido.

40 Para incrustar cada marca de agua, se procesa cada trama de la primera 1408 y la segunda 1410 versiones con marca de agua. Cada trama se transforma en un espacio de bloque de DCT (Transformada de coseno discreta), de manera que se aplica una DCT a cada bloque de píxeles NxN en la trama de entrada, donde N es 8 en la realización descrita. En otras realizaciones, N puede ser cualquier otro valor, pero se cree que N = 8 proporciona una marca de agua relativamente robusta sin presentar una carga demasiado grande para el procesador.

45 Se genera una secuencia pseudoaleatoria de coeficientes de marca de agua, y los coeficientes DCT en cada bloque DCT son modificados por esta secuencia pseudoaleatoria. La cantidad de cambio de cada componente está ponderada por la fuerza de DCT, para reducir los artefactos visuales que de otro modo podrían ser evidentes para un espectador de la trama. Luego, se aplica una DCT inversa a cada bloque para recuperar la versión con marca de agua de la imagen original. Cada trama de entrada (sin marca de agua) se procesará para generar una o más tramas de salida (con marca de agua), una por símbolo de marca de agua, donde cada símbolo de marca de agua tiene una secuencia pseudoaleatoria de coeficientes diferente. Aquí se apreciará que, en la realización descrita, se aplica un símbolo de marca de agua diferente a cada una de la primera 1408 y segunda 1410 versiones con marca de agua, pero que se aplica el mismo símbolo a cada trama de esas versiones.

55 Es concebible que, en algunas realizaciones, algunas tramas dentro del contenido multimedia no tengan marca de agua, de modo que todas las demás, cada enésima trama, o similares, tengan marcas de agua. Sin embargo, se cree que tales realizaciones son menos robustas y llevaría más tiempo recuperar la marca de agua de cualquier flujo multimedia candidato, ya que solo se procesa cada enésima trama, lo que proporciona menos información de la que recuperar la marca de agua.

60 La marca de agua basada en DCT se recupera sin recurrir a la imagen original incorrupta; de A. Piva, M. Barni, F. Bartolini y V. Cappellini; Dipartimento di Ingegneria Elettronica; Universit'a di Firenze via S. Marta 3, 50139, Firenze, Italia describe técnicas similares para incrustar marcas de agua y el experto debe leer este documento para comprender cómo incrustar una marca de agua de modo que pueda recuperarse sin hacer referencia a la imagen original.

65

A Survey of Digital Image Watermarking Techniques”, Vidyasagar M. Potdar, Song Han, Elizabeth Chang también puede ser útil para los expertos.

La trama con marca de agua se procesa aún más para reducir los artefactos visuales asociados con la marca de agua. Se calcula una máscara espacial, donde el valor de la máscara representa la tolerancia del HVS (Sistema Visual Humano) a pequeños cambios en el valor de cada píxel en la trama. Por ejemplo, el valor de la máscara en cada píxel se puede establecer en la variación de los 8 píxeles que rodean inmediatamente. La trama con marca de agua enmascarada viene dada por la suma de la trama original y la trama con marca de agua, ponderada por la máscara HVS, de modo que la marca de agua está fuertemente incrustada en las regiones de la versión donde el HVS es menos sensible a las perturbaciones. Se aplica la misma máscara a cada trama con marca de agua, para producir una pluralidad de tramas con marca de agua enmascarados (dos en la realización actual).

En la realización que se describe, el proceso de marca de agua 1407 genera adicionalmente metadatos que contienen descriptores de características clave dentro de las tramas del flujo multimedia con marca de agua, características que se utilizan para alinear temporal y espacialmente los flujos multimedia candidatos para recuperar cualquier marca de agua en estos. Con el fin de reducir la complejidad computacional y los requisitos de almacenamiento, las características no se extraen para cada trama y, en esta realización, las características se extraen para cada 10ma trama. Otras realizaciones pueden almacenar características para más tramas (es decir, para tramas de menos de cada 10 tramas) o para menos tramas (es decir, para tramas de más de cada 10 tramas). Las características extraídas en la realización actual están en forma de descriptores de características ORB (Oriented FAST y Rotated BRIEF), que permiten reconstituir un flujo multimedia candidato reescalado y/o rotado a las dimensiones originales durante la extracción de la marca de agua.

Es conveniente describir el sistema de entrega 1412 como un protocolo de transmisión adaptativo (tal como transmisión en vivo HTTP (HLS)); transmisión HTTP adaptativa 3GPP; transmisión fluida de Microsoft™; transmisión dinámica adaptativa a través de HTTP (DASH)), pero en otras realizaciones se pueden usar sistemas distintos de los protocolos de transmisión adaptativa. Algunas realizaciones del sistema pueden disponerse para procesar flujos multimedia que no corresponden a un formato conocido. Algunas realizaciones del sistema pueden volver a combinar archivos de audio y video de flujos/archivos separados. Tal funcionalidad surge típicamente del procesamiento de la información de encabezado de las diversas capas del tráfico de red como se describe a continuación.

En un protocolo de transmisión adaptable, cada versión 1408, 1410 de los contenidos multimedia de referencia con marca de agua se codifica de tal manera que se mantienen múltiples versiones y cada versión es adecuada para diferentes velocidades de bits, diferentes tamaños de tramas, etc. Sin embargo, la marca de agua es tal que sobrevivirá la codificación de los flujos con marca de agua 1408, 1410. La figura muestra un depósito 1414, 1416 respectivamente para cada una de la primera y segunda versiones con marca de agua. El experto en la técnica apreciará que, si bien es conveniente mostrar repositorios separados, no es necesario proporcionar repositorios separados. Por tanto, en otras realizaciones puede haber más o menos repositorios.

El experto en la técnica apreciará que para un protocolo de transmisión adaptativo, el flujo multimedia se entrega, no como un flujo continuo, sino como una serie de segmentos de duración limitada (por ejemplo, bloques) generados a partir del flujo multimedia original. Por lo tanto, cada una de las múltiples versiones de los flujos multimedia contenidos dentro de los repositorios 1414, 1416 comprende una serie de segmentos de duración limitada. En la realización que se describe, cada uno de los segmentos de duración determinada es típicamente de 5 a 10 segundos de duración. Todas las tramas dentro de un segmento de duración limitada determinado tendrán la misma marca de agua aplicada. Sin embargo, es concebible que al menos algunas realizaciones utilicen una longitud para cada segmento de duración limitada de menos de 5 segundos, que es el tiempo de segmento de duración limitada mínimo sugerido para algunos protocolos de transmisión adaptativos para garantizar que no haya demasiado tráfico de red.

Las realizaciones que no usan protocolos de transmisión adaptativos, en particular, pero no exclusivamente, pueden usar segmentos de duración limitada de menos de 5 segundos y, por ejemplo, pueden usar segmentos de duración limitada en el rango de 0,1 segundos a 15 segundos. El experto en la técnica apreciará que cuanto más corto sea el segmento de duración limitada, más rápidamente se puede extraer una marca de agua del contenido multimedia/flujo multimedia.

Cada segmento de duración limitada tendrá una hora de inicio y una hora de finalización en relación con el flujo multimedia original, es decir, de referencia. Por ejemplo, en una realización en la que cada segmento de duración limitada tiene una duración de 5 segundos, el tercer segmento de duración limitada comenzará a los 10 segundos en la secuencia y terminará a los 15 segundos en la secuencia. Tanto la hora de inicio como la hora de finalización pueden considerarse información de tiempo.

El sistema 1400 se dispone para almacenar la información de tiempo en relación con cada uno de los flujos multimedia contenidos en los repositorios 1414, 1416. En algunas realizaciones, la información de tiempo se mantiene dentro del almacén 210, pero no es necesario que este sea el caso.

Se dispone un proceso de entrega 1418 para atender la solicitud de un flujo multimedia desde un dispositivo cliente. Tres de estos dispositivos cliente se muestran en la Figura: un teléfono móvil 1420 (tal como un iPhone™); un dispositivo Android™; un móvil Windows™; o una Blackberry™); un ordenador 1422 (tal como un PC que ejecute Windows™, LINUX o similares; un ordenador Apple; o similares) y una tableta 1424 (tal como un iPad™; un Kindle™; una tableta Android™ o similares). El experto en la técnica apreciará que otros dispositivos, tales como televisores; relojes; o similares también pueden formar dispositivos cliente. Como se describe a continuación, el proceso de entrega 1418 se dispone para generar una versión única del flujo multimedia para cada uno de los dispositivos cliente 1420, 1422, 1424, cuya versión única se dispone para identificar el dispositivo cliente al que se entregó ese flujo multimedia. Por lo tanto, cada dispositivo cliente 1420, 1422, 1424 puede considerarse como un dispositivo objetivo al que se puede entregar contenido multimedia.

Específicamente, un usuario 1420-1424 solicita un flujo multimedia (es decir, un componente multimedia) del sistema 1400. Según los protocolos de transmisión adaptativos (como HLS), el cliente, siguiendo la solicitud del flujo multimedia, recibe una lista de texto de URL (localizadores uniformes de recursos) que genera el sistema 1602. Estas URL son para los segmentos de video de duración limitada que componen una versión del flujo multimedia solicitado. Como se analizó anteriormente, cada uno de los repositorios 1414, 1416 contiene múltiples copias del flujo multimedia con el fin de que diferentes condiciones de canal, dispositivos de destino, etc. puedan utilizar el flujo multimedia.

Se apreciará que en otras realizaciones, un usuario puede solicitar un componente multimedia que no sea un flujo multimedia. Por ejemplo, un usuario podría descargar un archivo que proporcione contenido multimedia (por ejemplo, un archivo de video o un archivo de audio).

Además, el sistema 1400 determina la identidad del usuario 1420-1424 que realizó la solicitud del flujo multimedia. Por ejemplo, es probable que el usuario tenga una cuenta con el proveedor del flujo multimedia al que está realizando la solicitud y la identidad del usuario puede obtenerse de los detalles de la cuenta de ese usuario 1420-1424. En otras realizaciones, el usuario 1420-1424 puede necesitar ingresar un identificador, etc. cuando solicitan el flujo multimedia.

El sistema de entrega 1412 se dispone entonces para generar una lista de reproducción para ese usuario 1420-1424 para esa solicitud. Por lo general, esta lista de reproducción será única para esa solicitud particular del flujo multimedia. Las URL que componen la lista de reproducción apuntan a segmentos de duración limitada en cada uno de los repositorios 1414, 1416 dentro del sistema de entrega 1412 de manera que un conjunto de segmentos de duración limitada que cuando se unen por el protocolo generan el flujo multimedia. Como tal, cuando el navegador u otro software utilizado por el usuario 1420-1424 accede a la lista de URL, se entregan segmentos de duración limitada que se generan a partir de cada una de la primera 1408 y segunda 1410 versiones con marca de agua del flujo multimedia. Debido a que la lista de URL ha sido generada por el sistema específicamente para cada uno de los usuarios 1420-1424, cada usuario recibe, en la realización que se describe, una combinación única de segmentos de duración limitada para formar el flujo multimedia entregado que se le entrega. Tales flujos multimedia entregados están ejemplificados por las líneas 1419a, 1419b, 1419c en la Figura 14.

En la realización que se describe, el sistema de entrega 1412 usa un número binario de 32 bits para generar la secuencia de URL. En la realización que se describe, se puede pensar que la elección del símbolo de marca de agua dentro de los segmentos de duración limitada del flujo multimedia enviado a un usuario 1420-1424 proporciona un cero o uno de un número binario. Aquí se apreciará que cada uno del primer 1408 y segundo 1410 flujos con marca de agua tiene un símbolo diferente aplicado a estos. Por ejemplo, la presencia de un segmento de duración limitada de la primera versión con marca de agua 1408 da un cero y la presencia de un segmento de duración limitada de la segunda versión con marca de agua 1410 da un uno. Por lo tanto, en la realización que se describe, se requieren 32 segmentos de duración limitada para identificar al usuario 1420-1424 a quien se transmitió el flujo multimedia. Por tanto, el orden en el que se entregan los segmentos de duración limitada proporciona un código de identidad que se ha asignado a un usuario determinado.

Por tanto, en la realización que se está describiendo, se necesitan 32 segmentos de duración limitada para proporcionar el código de identidad que se ha insertado en el flujo multimedia entregado 1419. Por lo tanto, suponiendo que un segmento de duración limitada tiene una duración de 5 segundos, se necesitan 2 minutos y 40 segundos de flujo multimedia para recuperar el código de identidad. Por lo tanto, la longitud de los segmentos de duración limitada y el número de bits que codifican el código de identidad son un equilibrio entre proporcionar una identificación robusta, permitir indexar una gran población de usuarios y permitir que el código de identidad se recupere rápidamente de un flujo multimedia.

El experto en la técnica apreciará que otras realizaciones pueden utilizar un número diferente de bits. Además, otras realizaciones pueden usar bases numéricas distintas de la base 2 para la generación del código y tales realizaciones comprenderían más de dos versiones de los flujos multimedia con marca de agua: para generar un número de base 3 se requeriría la versión 3 del flujo multimedia con marca de agua, etc.

A medida que el contenido multimedia se entrega a los dispositivos cliente 1420-1424, un sistema de monitoreo 218 monitorea 1604 la red 1402/251 en busca de contenidos multimedia que tengan una huella digital almacenada dentro del almacén 210. Si el sistema de monitoreo 218 detecta, en la red, copias de un flujo multimedia de referencia 204

que tienen una huella digital reconocida (es decir, una huella digital almacenada en el almacén 210), entonces un controlador 1426 se dispone para identificar, usando cualquier marca de agua dentro del flujo multimedia, el usuario 1420-1424 al que se envió el contenido multimedia. El controlador 1426 se dispone entonces, en algunas realizaciones, para detener la entrega del flujo multimedia entregado a ese usuario 1420-1424 controlando 1428 el proceso de entrega 1418.

Aunque solo se muestra un generador de huellas digitales 202 en la Figura 1, el sistema de huellas digitales 200 típicamente tendrá varios generadores de huellas digitales de referencia 202a, 202b, 202c, 202d. La Figura 2 muestra cuatro generadores de huellas digitales 202a-d, pero esto es meramente indicativo y otras realizaciones pueden contener cualquier número.

Con referencia a los sistemas descritos anteriormente, luego los flujos multimedia candidatos proporcionan contenido multimedia candidato a decodificar del que se realiza un intento de extraer una marca de agua como se describe a continuación.

El proceso de detección de huellas digitales 1606 como se describe en relación con la Figura anterior identifica los flujos multimedia candidatos que se cree que coinciden con los flujos multimedia de referencia. Estos flujos multimedia candidatos identificados pueden considerarse como flujos multimedia coincidentes. El siguiente proceso se realiza en flujos multimedia coincidentes para determinar si ese flujo lleva una marca de agua. Aquí se apreciará que el sistema de monitoreo descrito anteriormente puede usarse para monitorear flujos multimedia tanto con marca de agua como sin marca de agua.

La salida del proceso anterior proporciona lo que puede considerarse una sincronización temporal aproximada entre el flujo multimedia de referencia y el flujo multimedia candidato. En este caso, puede pensarse que la sincronización temporal identifica qué tan remota está una porción del flujo multimedia candidato desde el comienzo del flujo multimedia de referencia original. La referencia de porción aquí podría ser una trama de un flujo de video u otra porción. Tal alineación temporal es útil en realizaciones que detectan marcas de agua, ya que se apreciará que se aplican diferentes marcas de agua a segmentos de duración limitada de 5 segundos del flujo multimedia. Por tanto, para identificar el código de identidad proporcionado por la secuencia de marca de agua, el sistema identifica al menos los puntos de inicio y finalización de los segmentos de duración limitada dentro de los flujos multimedia coincidentes.

La alineación del flujo multimedia coincidente 1500 contra un flujo multimedia de referencia 204 se muestra en la Figura 15a y 15b. La Figura 15a muestra un flujo multimedia coincidente 1500 que está temporalmente alineado con el flujo multimedia de referencia 204 y se puede ver que los segmentos de duración limitada, cada uno representado por un bloque, tal como 1502a, 1502b, 1504a, 1504b. Como en la Figura 141, el sombreado de cada bloque muestra la marca de agua que se ha aplicado a ese bloque. Se apreciará que el flujo multimedia candidato 1500 no puede dividirse en segmentos de duración limitada del flujo multimedia 204 original, de referencia, cuyos segmentos de duración limitada probablemente se hayan eliminado durante cualquier copia, modificación, etc. de uno de los flujos multimedia entregados 1419. Sin embargo, las regiones con marca de agua probablemente retendrán la sincronización del flujo multimedia con marca de agua y, por lo tanto, en la realización que se describe, aparecerán en bloques de 5 segundos dentro de un flujo multimedia candidato 1500.

La Figura 15b muestra un flujo multimedia coincidente 1506 que no está alineado temporalmente con su flujo multimedia de referencia 204 asociado. Como se analizó anteriormente, la alineación temporal significa que el sistema no ha identificado porciones del flujo multimedia coincidente con respecto a un punto, típicamente el comienzo, del flujo multimedia de referencia 204. Como tal, puede verse que el segmento 1502b no está alineado con el bloque 1502a y el bloque 1504a no está alineado con el bloque 1504b. Por tanto, en el ejemplo de la Figura 15b, si se comparan las marcas de agua, la comparación probablemente fallaría ya que el sistema ha identificado incorrectamente los segmentos de duración limitada con respecto al inicio del flujo multimedia de referencia 204.

En el caso de un flujo multimedia de video, la salida del proceso de toma de huellas digitales anterior también identificará si el flujo multimedia coincidente se ha invertido horizontal y/o verticalmente. Las realizaciones pueden procesar el flujo multimedia coincidente para revertir cualquier cambio detectado antes de detectar la marca de agua dentro del flujo multimedia coincidente.

Es concebible que la alineación temporal del flujo multimedia coincidente del proceso de toma de huellas digitales no sea suficiente. Como tal, al menos algunas realizaciones, incluida la que se describe, realizan una alineación adicional. En particular, se apreciará que la tasa de cambio en algunos flujos multimedia puede ser bastante baja y tales flujos multimedia, especialmente, pueden necesitar una alineación más precisa. Se apreciará que cuanto más precisos puedan alinearse los puntos de inicio y final de los segmentos de duración limitada dentro de un flujo multimedia coincidente con el flujo multimedia de referencia, entonces es más probable que el proceso de extracción de marcas de agua descrito a continuación dé el resultado correcto.

La realización que se describe extrae descriptores de características del flujo multimedia coincidente del mismo formato en el flujo multimedia de referencia 204 original (descriptores de características ORB). A continuación, se realiza un registro de video entre las características extraídas del flujo multimedia coincidente y el flujo multimedia de

referencia 204 con el que se ha encontrado una coincidencia. El registro de vídeo identifica una sincronización temporal más precisa entre el flujo multimedia coincidente 1500 y el flujo multimedia de referencia, así como cualquier transformación geométrica (por ejemplo, rotación; escalado; y/o traslación) entre el flujo multimedia coincidente 1500 y el flujo multimedia de referencia 204. El flujo multimedia coincidente 1500 se transforma para revertir cualquier transformación de este tipo, de modo que tenga la misma geometría que el flujo multimedia de referencia 204 original.

A continuación, el proceso identifica las tramas dentro de cada uno de los segmentos de duración limitada (por ejemplo, 1502b, 1504b) y procesa cada una de esas tramas para extraer las marcas de agua que se encuentran dentro de cada trama.

En la realización que se describe, hay dos posibles marcas de agua que podrían haberse aplicado a cualquier trama dada. En otras realizaciones, es posible que se pueda aplicar un número diferente de símbolos de marca de agua, o simplemente la presencia o ausencia de una marca de agua. Como tal, el proceso asigna una creencia a cada una de las posibles marcas de agua y determina que la trama que se está procesando contiene la marca de agua en la que se asigna la creencia más alta.

Para extraer la marca de agua, la trama se transforma en un espacio DCT de bloque como se describe en el proceso de incrustación de la marca de agua. La puntuación de cada bloque DCT viene dada por la correlación sumada entre los componentes DCT y las secuencias aleatorias correspondientes a cada código de marca de agua candidato (como se describió anteriormente). Los puntajes generales para la trama vienen dados por la suma de los puntajes de los bloques DCT individuales.

A continuación, se asigna un valor de marca de agua a cada uno de los segmentos de duración limitada. En la realización que se describe, esta determinación comprende determinar cuál de los dos símbolos de marca de agua tiene la puntuación más alta en cada segmento de duración limitada. Esta determinación se realiza sobre una base de puntuación simple: el símbolo de marca de agua con la puntuación más alta sumada a las tramas que componen el segmento de duración limitada se toma para representar el valor de la marca de agua en ese segmento.

A continuación, se determina 1608 el código de identidad que se ha codificado en el flujo multimedia entregado, del cual se ha derivado el flujo coincidente. Se apreciará desde arriba que el código de identidad está codificado, en esta realización, en 32 segmentos consecutivos de duración limitada. Como tal, la realización coteja las marcas de agua de 32 segmentos consecutivos de duración limitada. La información de tiempo que el sistema ha almacenado se utiliza para ayudar a garantizar que los puntos de inicio y finalización de la secuencia de segmentos de duración limitada se identifiquen correctamente. El resultado de este proceso es la generación de un número de 32 bits proporcionado por la marca de agua que se ha aplicado a los segmentos de duración limitada del flujo multimedia coincidente. En otras realizaciones, es igualmente posible cualquier otra longitud de número.

Finalmente, el número de 32 bits se compara con códigos de identidad válidos que se han asignado a los usuarios 1420-1424. Si el número de 32 bits corresponde a un código válido, entonces la extracción del código de identidad ha sido exitosa, mientras que si este no es el caso, la extracción del código de identidad ha fallado.

Algunas realizaciones pueden disponerse para extraer continuamente códigos de identidad de un flujo multimedia candidato, o al menos realizar extracciones múltiples del código de identidad de un flujo multimedia candidato, y posteriormente promediar o combinar de otro modo los códigos de identidad. Tales realizaciones pueden ayudar a aumentar la creencia en el código de identidad que se extrae del flujo multimedia coincidente.

Una vez que se ha extraído un código de identidad válido de un flujo multimedia coincidente, se determina la identidad del usuario al que se envió el flujo multimedia entregado 1419a-c, típicamente accediendo a una base de datos de usuarios que contiene el código de identidad asignado a cada uno de ellos. Las realizaciones del sistema pueden entonces tomar la acción apropiada, incluyendo pero sin limitarse a terminar automáticamente el flujo multimedia entregado al usuario identificado (por ejemplo, el usuario 1420).

El proceso de determinar la identidad de un usuario (por ejemplo, el usuario 1420) a partir de un flujo multimedia coincidente se realiza en tiempo real de manera que es posible terminar el flujo multimedia mientras se transmite el flujo multimedia entregado. En la realización que se describe, se apreciará que el proceso de identificación de huellas digitales realizado en el flujo multimedia candidato 219 toma el orden de 1 minuto. Además, el proceso de extracción de la marca de agua tomará el orden de 5 minutos (32 bits, con 10 segundos por bit). Por lo tanto, aquí, en tiempo real, puede significar en el orden de 5 a 10 minutos. Por ejemplo, en tiempo real puede significar sustancialmente 8 minutos.

Como se describió anteriormente, un número de 32 bits se codifica en los flujos entregados a través de la primera y segunda marcas de agua para proporcionar un código de identidad, típicamente, único en cada uno de los flujos entregados. En algunas realizaciones, al menos algunos de los bits del código incrustado se utilizan para proporcionar bits de encabezado y/o de verificación. Como tal, es posible que el código de identidad no sea un número de 32 bits. El experto en la técnica apreciará que el número de bits usados para codificar información en los flujos multimedia entregados 1419a-c puede variar entre realizaciones.

- 5 En algunas realizaciones, los bits de verificación pueden proporcionar un código de corrección de errores. Por ejemplo, los bits de verificación pueden permitir cualquiera de los siguientes códigos: un código de Hamming; un código de Reed-Solomon; una verificación de paridad de baja densidad; o similares. El experto en la técnica apreciará que el uso de dichos códigos aumentará la robustez, para las realizaciones que usan el código, con el que el código de identidad puede extraerse del flujo multimedia coincidente.
- En algunas realizaciones, el generador de huellas digitales de referencia puede denominarse aparato de monitoreo.
- 10 En algunas realizaciones, los mecanismos de procesamiento de identificación 216a-d pueden considerarse como un nodo.
- En lo anterior, el o cada componente multimedia puede considerarse como un conjunto de datos.
- 15 Como lo apreciará el experto en la técnica, las conexiones 208a-d, 212, 214a-d, o cualquier otra de las conexiones de red a las que se hace referencia en la presente memoria puede ser cualquier forma de conexión de red inalámbrica o por cable. Típicamente, estas conexiones se utilizarán en protocolos TCP/IP, pero no es necesario que sea así.

REIVINDICACIONES

1. Un método implementado por ordenador para derivar una secuencia de huellas digitales (FPn) a partir de uno o más componentes multimedia (204a-d, 219), comprendiendo el método;
  - a) recibir el o cada componente multimedia (204a-d, 219) en un dispositivo de procesamiento (112) y hacer que ese dispositivo de procesamiento (112):
    - i) mueva una ventana (305) a través del o de cada uno de los componentes multimedia (204a-d, 219) a una pluralidad de posiciones (305a, 305b);
    - ii) seleccione una parte del componente multimedia o de cada uno de ellos (204a-d, 219);
    - iii) procese la parte del o cada componente multimedia (204a-d, 219) para generar una huella digital (FPn) indicativa de la parte del o cada componente multimedia (204a-d, 219);
    - iv) almacene la huella digital (FPn);
    - v) procese la parte para generar una medida de calidad indicativa de la unicidad de la huella digital (FPn), en donde la medida de calidad se determina según un cambio en una propiedad entre tramas contiguas (300a-n, 306a-n) de uno o más componentes multimedia (204a-d, 219); y
    - vi) almacene la medida de calidad en asociación con la huella digital (FPn).
2. El método de la reivindicación 1, en el que la medida de calidad se genera para cada una de las posiciones (305a, 305b) de la ventana (305).
3. El método de la reivindicación 1 o 2, en el que el componente multimedia (204a-d, 219) es un vídeo y la parte del componente multimedia (204a-d, 219) seleccionada por cada posición (305a, 305b) de la ventana (305) cubre una pluralidad de tramas (300a-n, 306a-n) del vídeo y la medida de calidad se determina a partir del cambio en la propiedad entre al menos algunas de las tramas (300a-n, 306a-n) seleccionada por una posición (305a, 305b) de la ventana (305).
4. El método de cualquier reivindicación 3 anterior, en el que la propiedad es el cambio de intensidad entre tramas contiguas (300a-n, 306a-n) para generar un cambio de intensidad entre tramas.
5. El método de la reivindicación 4, en el que el cambio de intensidad entre tramas se ordena según la magnitud.
6. El método de la reivindicación 4 o 5, que toma un percentil predeterminado de la intensidad ordenada entre tramas cambia según la medida de calidad.
7. El método de cualquiera de las reivindicaciones 3 a 6, en donde la medida de calidad está dispuesta para dar una indicación de la cantidad de movimiento entre tramas.
8. Un sistema informático (100, 202a-d) dispuesto para generar una secuencia de huellas digitales (FPn) a partir de un componente multimedia (204a-d, 219), comprendiendo el sistema:
  - un dispositivo de procesamiento (112) dispuesto para recibir el o cada componente multimedia (204a-d, 219) y comprendiendo el dispositivo de procesamiento (112) además:
  - un generador de ventanas (126) dispuesto para seleccionar una parte del o de cada componente multimedia (204a-d, 219);
  - un generador de huellas digitales (130) dispuesto para seleccionar una primera parte del o cada componente multimedia (204a-d, 219) seleccionado por el generador de ventanas (126) y para generar una primera huella digital (FPn) indicativa de la primera parte del componente multimedia (204a-d, 219) y para almacenar esa huella digital (FPn); y
  - un evaluador de calidad (132) dispuesto para procesar la primera parte del componente multimedia (204a-d, 219) para generar una medida de calidad indicativa de la singularidad de la primera huella digital (FPn) y almacenar la medida de calidad en asociación con la primera huella digital (FPn), en donde la medida de calidad se determina según un cambio en una propiedad entre tramas contiguas (300a-n, 306a-n) de uno o más componentes multimedia (204a-d, 219).
9. El sistema de la reivindicación 8, donde el dispositivo de procesamiento (112) está dispuesto para generar la medida de calidad para cada una de las posiciones (305a, 305b) de la ventana (305).
10. El sistema de las reivindicaciones 8 o 9, en el que el componente multimedia (204a-d, 219) es un vídeo y la parte del componente multimedia (204a-d, 219) seleccionada por cada posición (305a, 305b) de la ventana (305) cubre una pluralidad de tramas (300a-n, 306a-n) del vídeo y el dispositivo de procesamiento (112) está dispuesto para determinar la medida de calidad a partir del cambio en la propiedad entre al menos algunas de las tramas (300a-n, 306a-n) seleccionadas por una posición (305a, 305b) de la ventana (305).

11. El sistema de cualquiera de las reivindicaciones 8 a 10, en donde el dispositivo de procesamiento (112) está dispuesto para usar, como cambio de propiedad, el cambio de intensidad entre tramas contiguas (300a-n, 306a-n) para generar un cambio de intensidad entre tramas.
- 5 12. El sistema de la reivindicación 11, en el que el dispositivo de procesamiento (112) está dispuesto para ordenar el cambio de intensidad entre tramas según la magnitud y en donde, opcionalmente, el dispositivo de procesamiento (112) está dispuesto para tomar un percentil predeterminado de la intensidad ordenada entre tramas cambia como medida de calidad.
- 10 13. Un medio legible por ordenador no transitorio que almacena un código de programa informático ejecutable para derivar una secuencia de huellas digitales (FPn) a partir de uno o más componentes multimedia (204a-d, 219), comprendiendo las etapas ejecutables del código de programa;
- 15 a)recibir el o cada componente multimedia (204a-d, 219) en un dispositivo de procesamiento (112) y hacer que ese dispositivo de procesamiento (112):
- 20 i)mueva una ventana (305) a través del o de cada uno de los componentes multimedia (204a-d, 219) a una pluralidad de posiciones (305a, 305b);
- ii)seleccione una parte del componente multimedia o de cada uno de ellos (204a-d, 219);
- 25 iii)procese la parte del o cada componente multimedia (204a-d, 219) para generar una huella digital (FPn) indicativa de la parte del o cada componente multimedia (204a-d, 219);
- iv)almacene la huella digital (FPn);
- v)procese la parte para generar una medida de calidad indicativa de la unicidad de la huella digital (FPn), en la que la medida de calidad se determina según un cambio en una propiedad entre tramas contiguas (300a-n, 306a-n) de uno o más componentes multimedia (204a-d, 219); y
- vi)almacene la medida de calidad en asociación con la huella digital (FPn).
- 30 14. El medio legible por ordenador no transitorio según la reivindicación 13, cuyo medio contiene además un código de programa informático ejecutable para hacer que un ordenador lleve a cabo el método de cualquiera de las reivindicaciones 2 a 7.

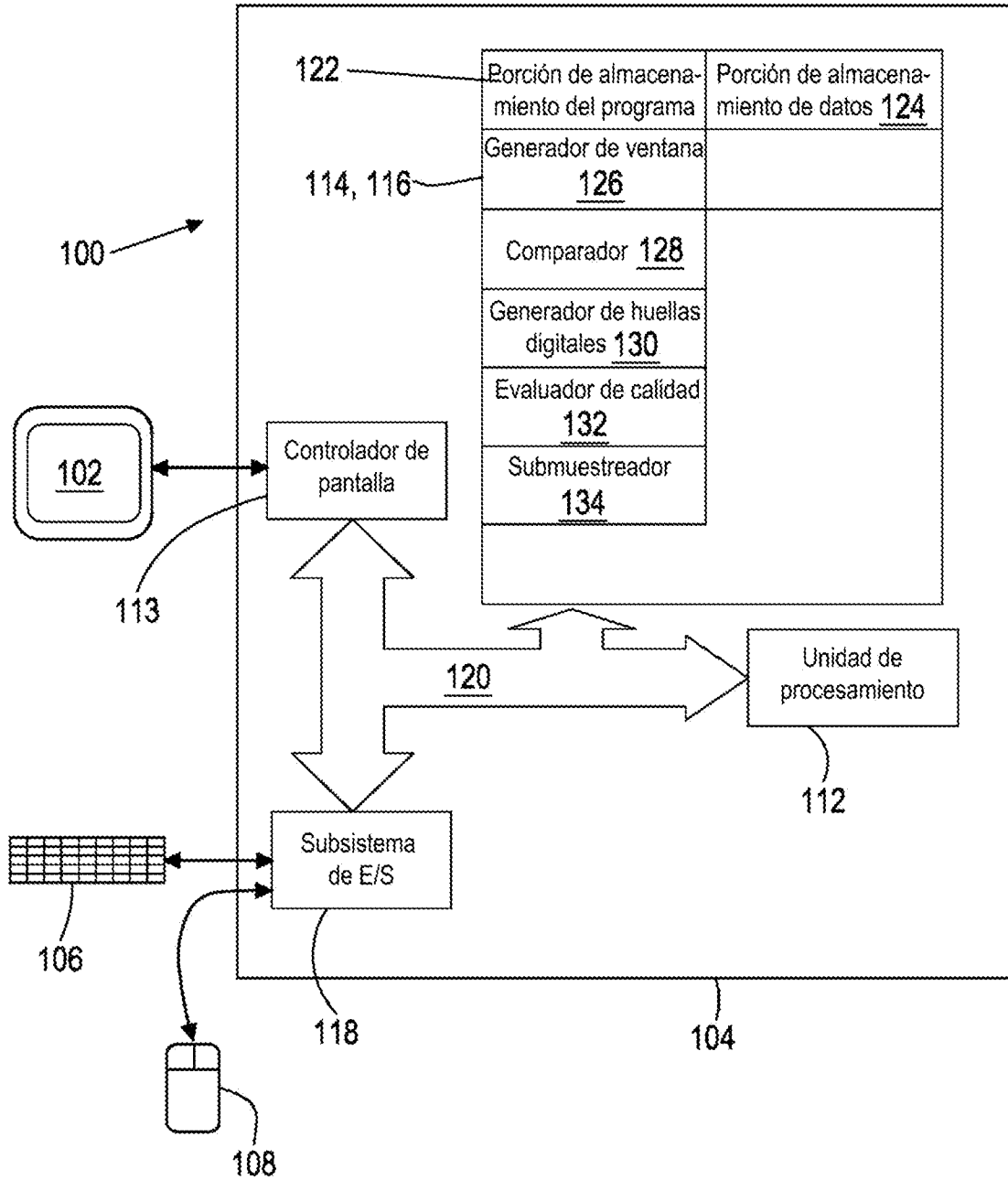


Figura 1

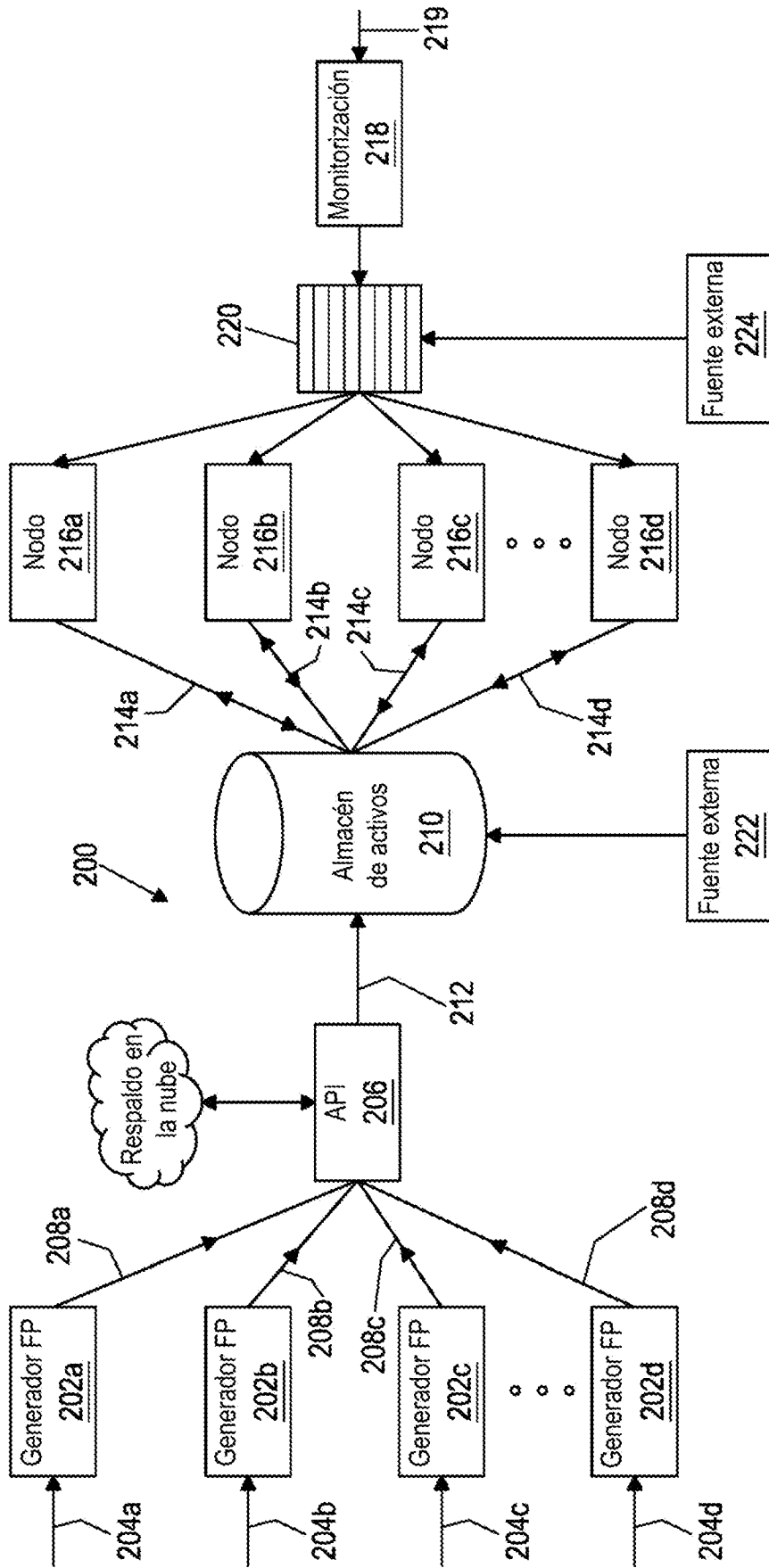


Figura 2

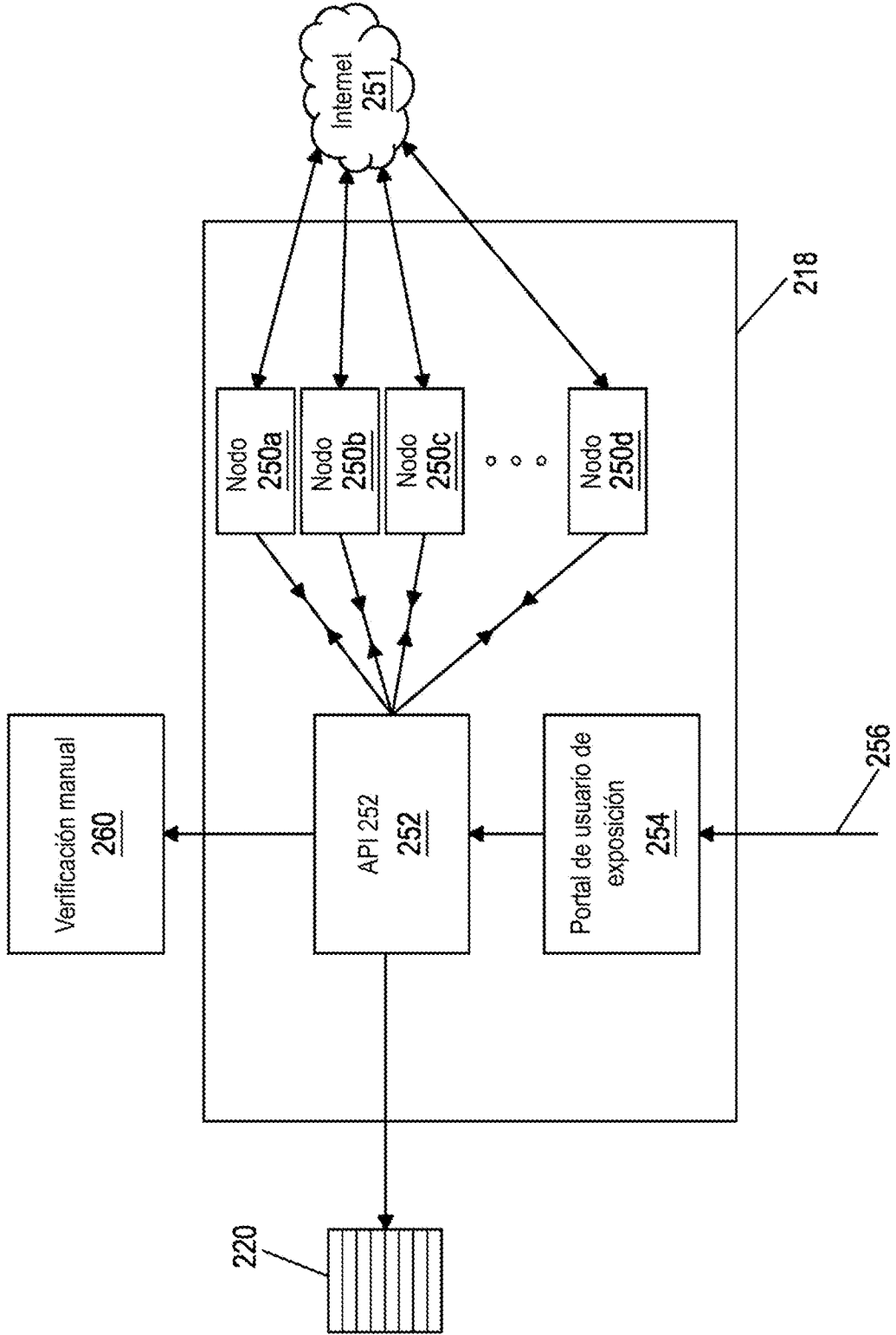


Figura 3

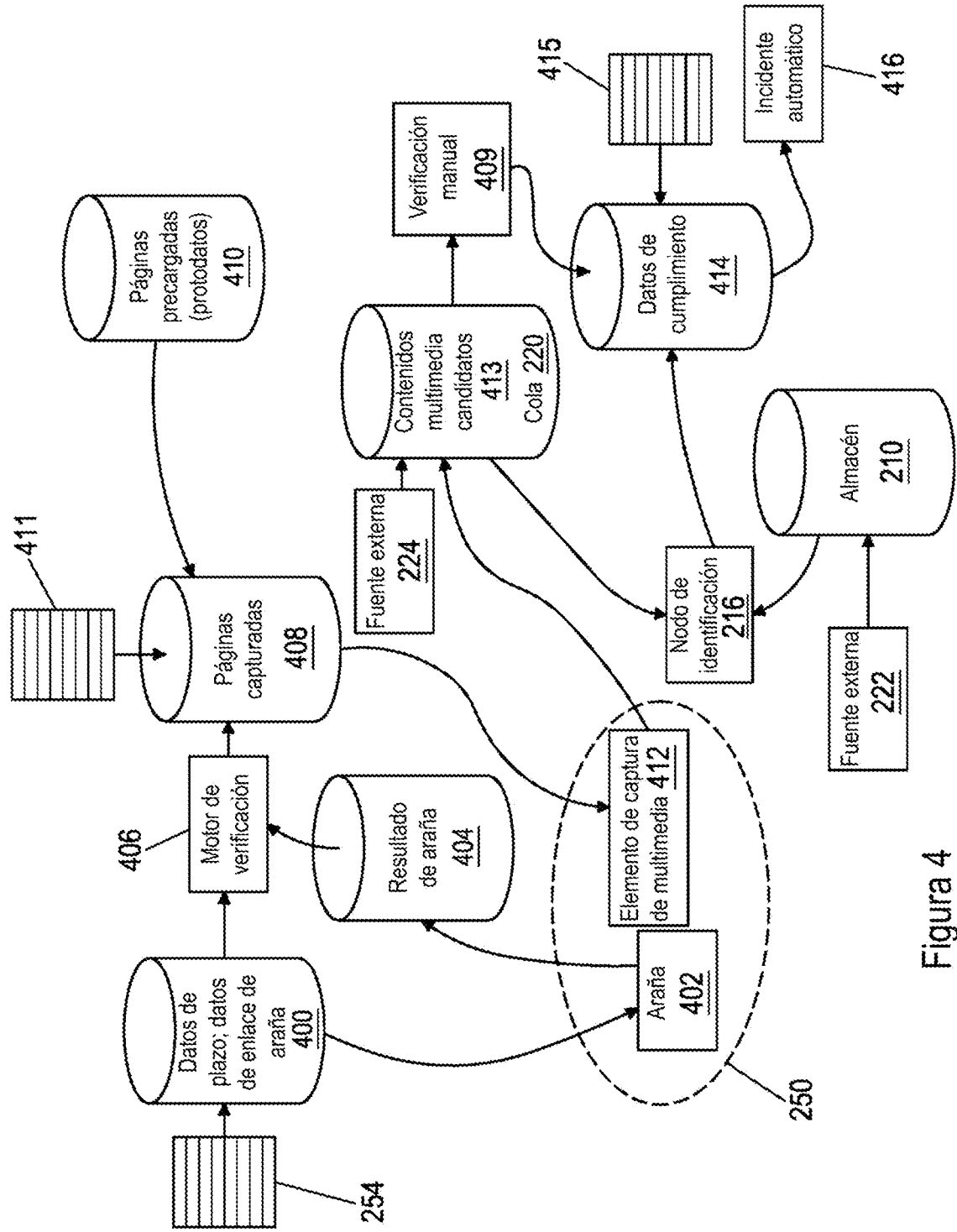


Figura 4

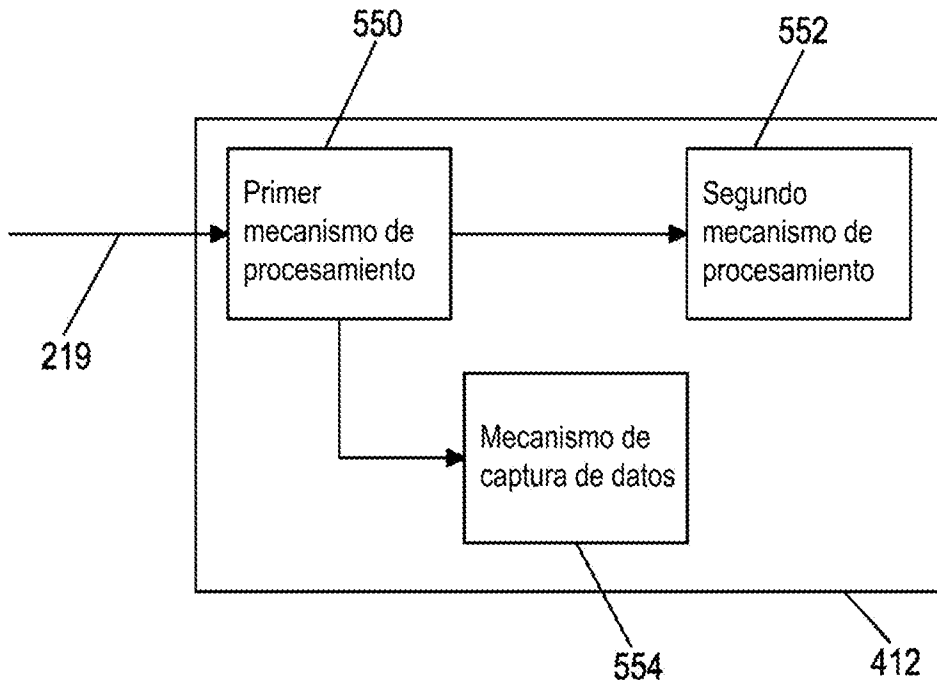


Figura 5

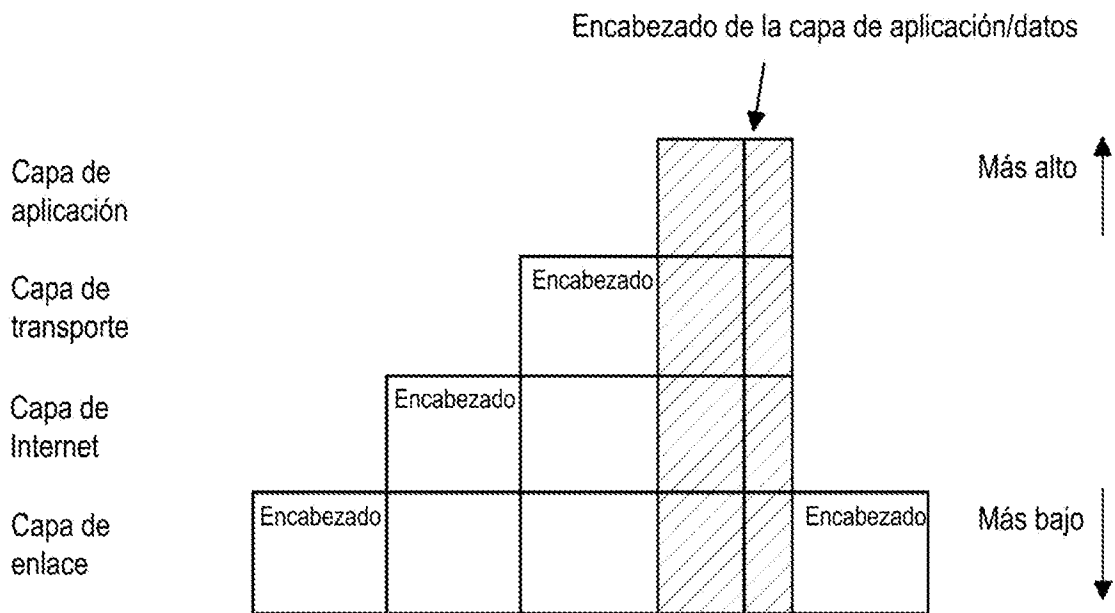


Figura 6

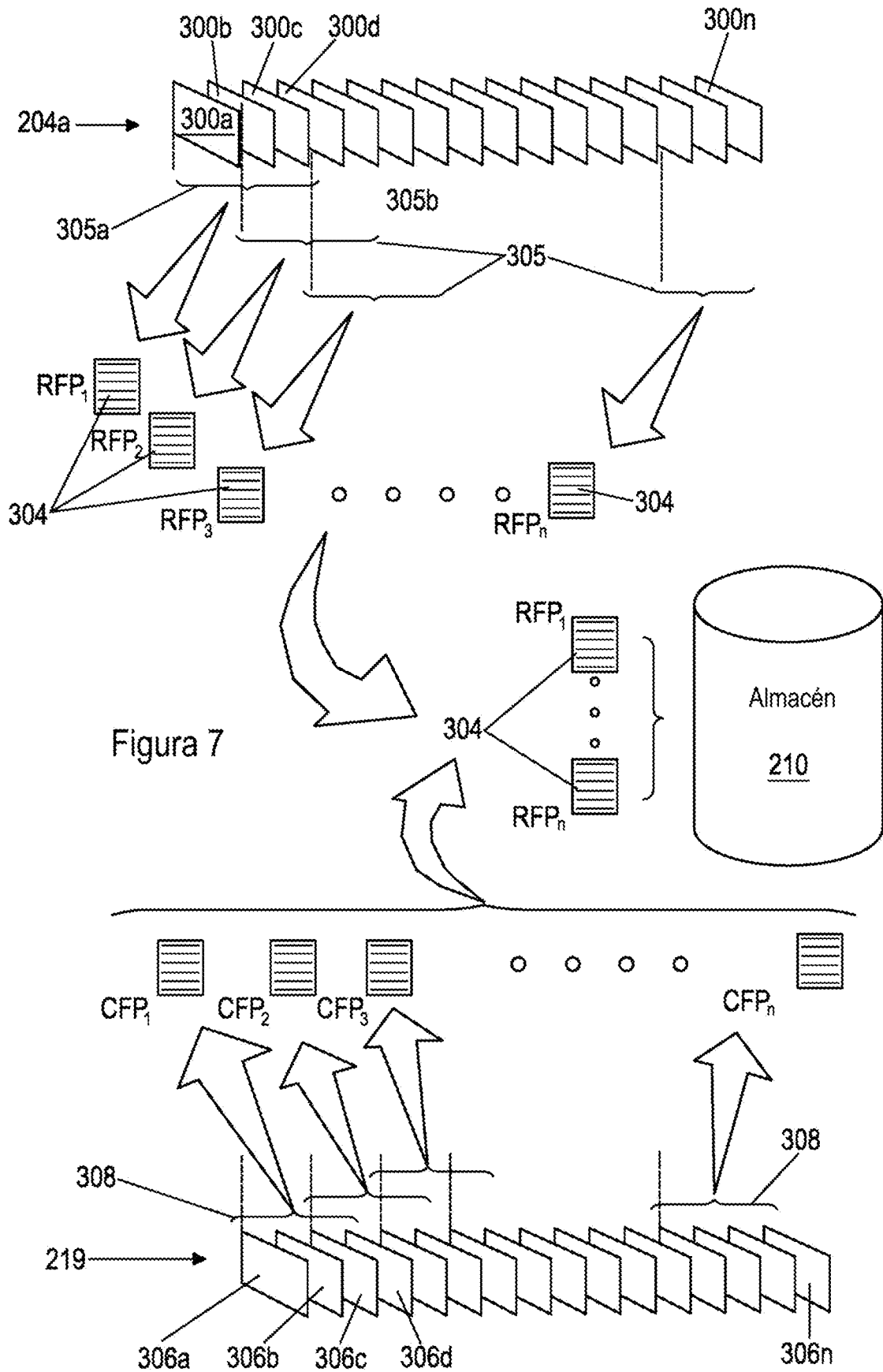


Figura 7

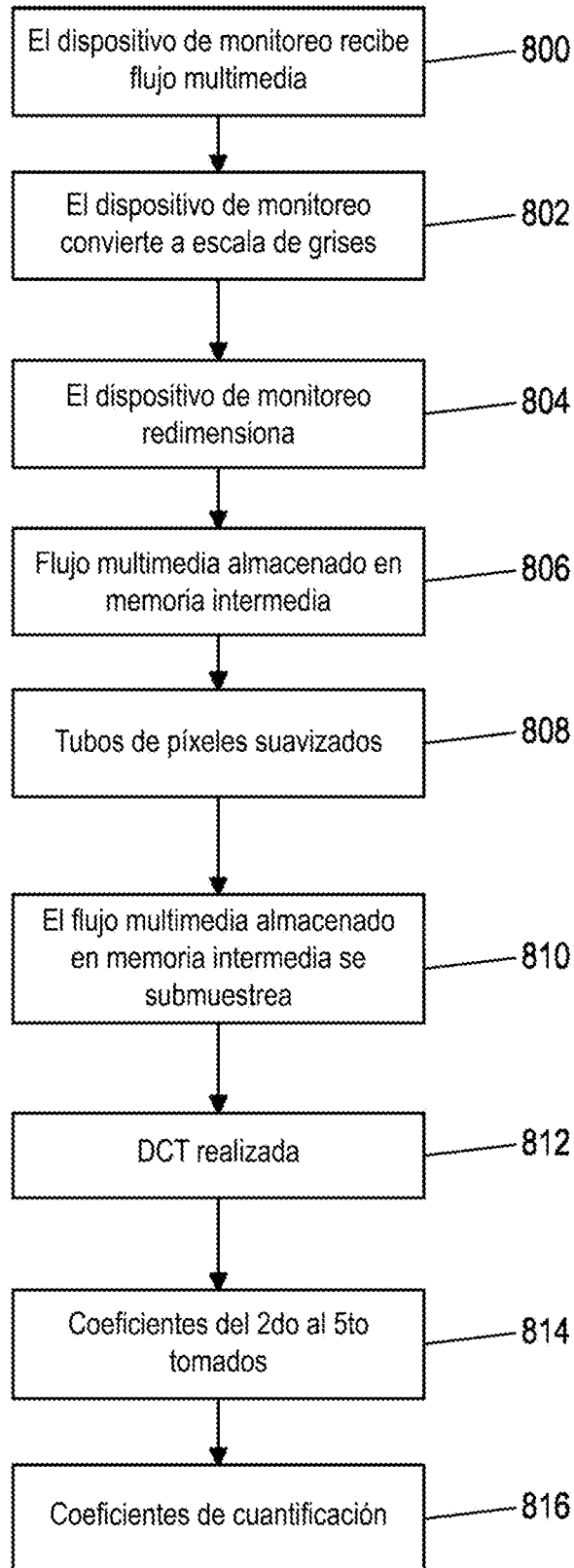


Figura 8

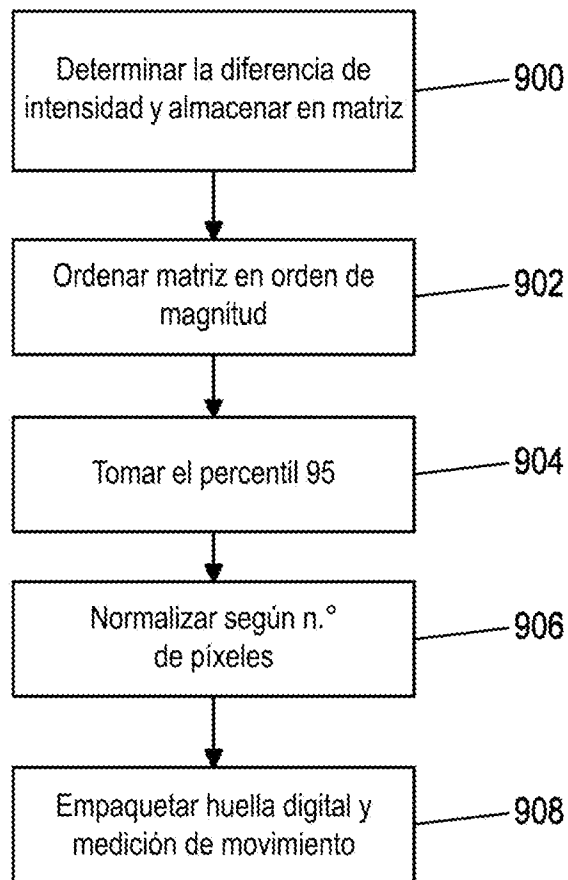


Figura 9

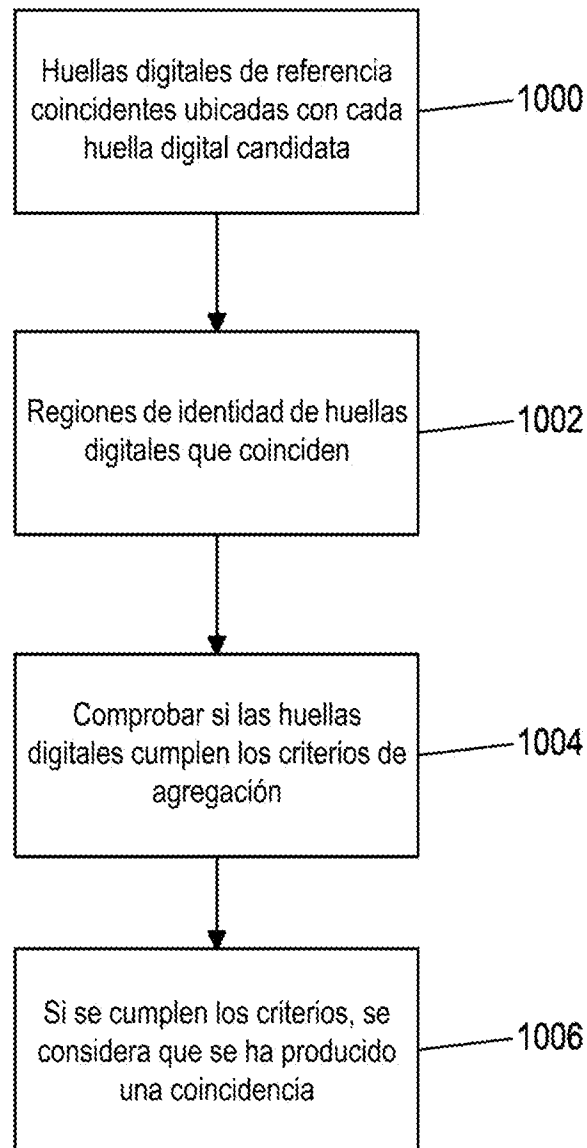


Figura 10

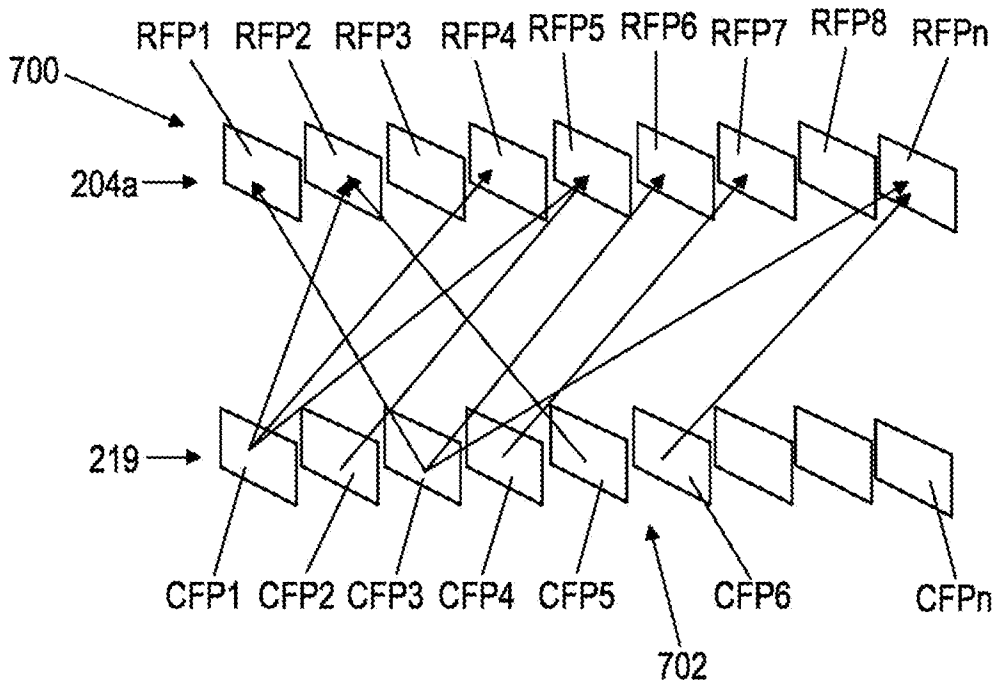


Figura 11a

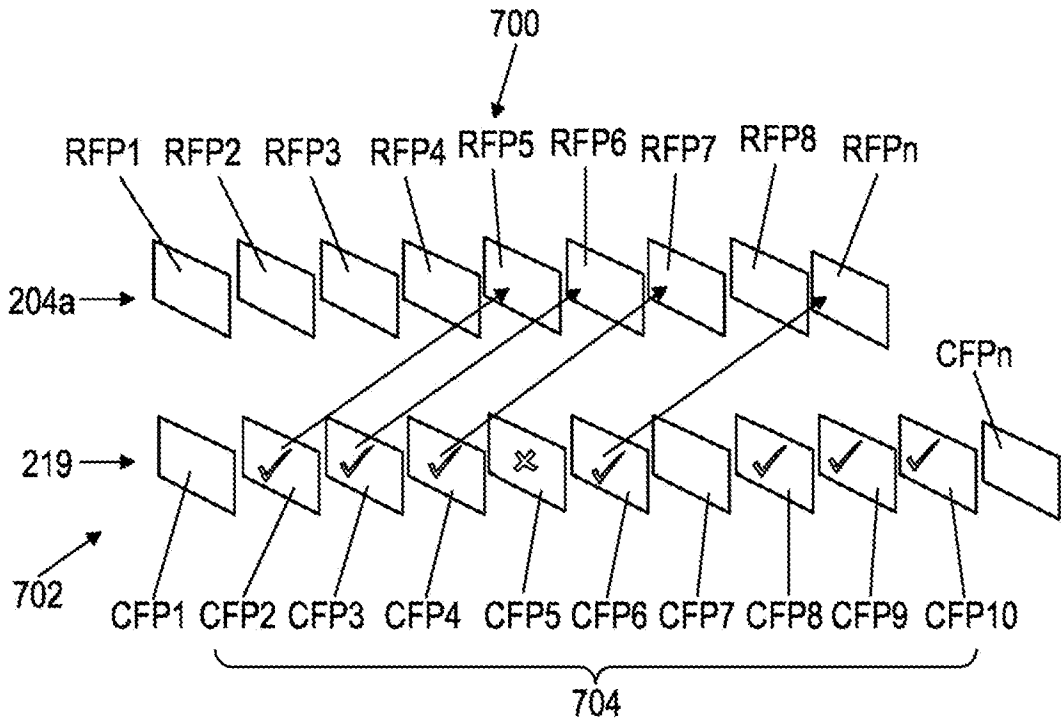


Figura 11b

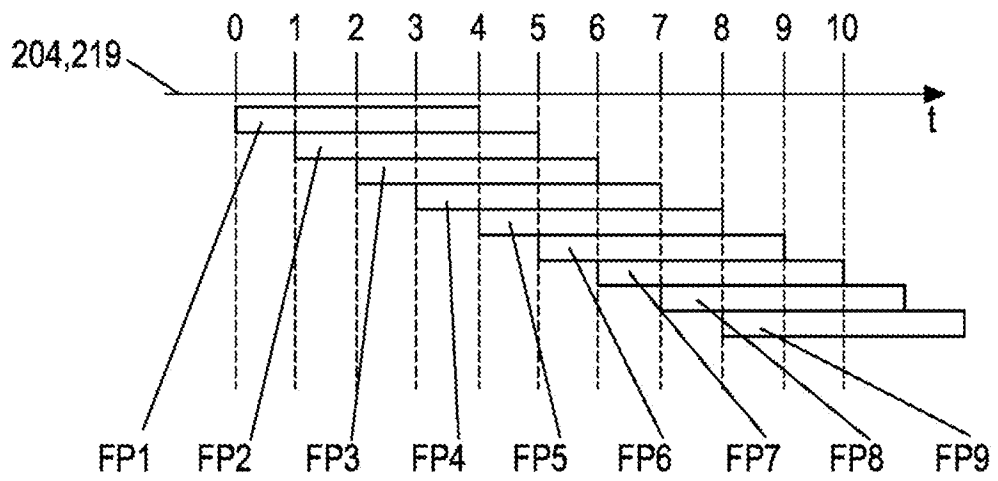


Figura 12

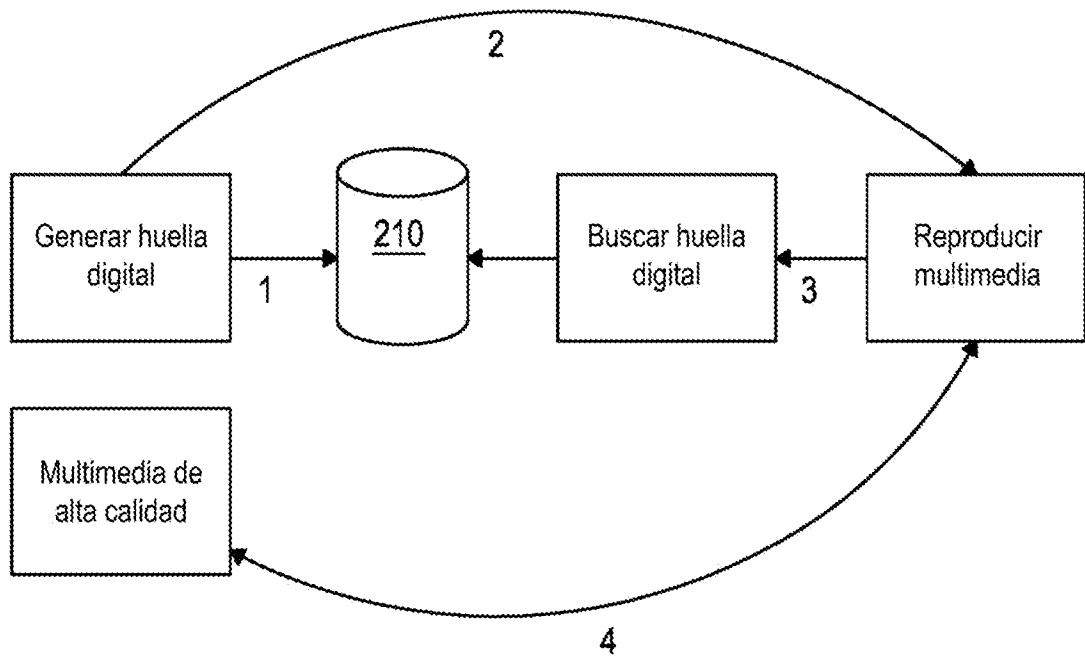


Figura 13

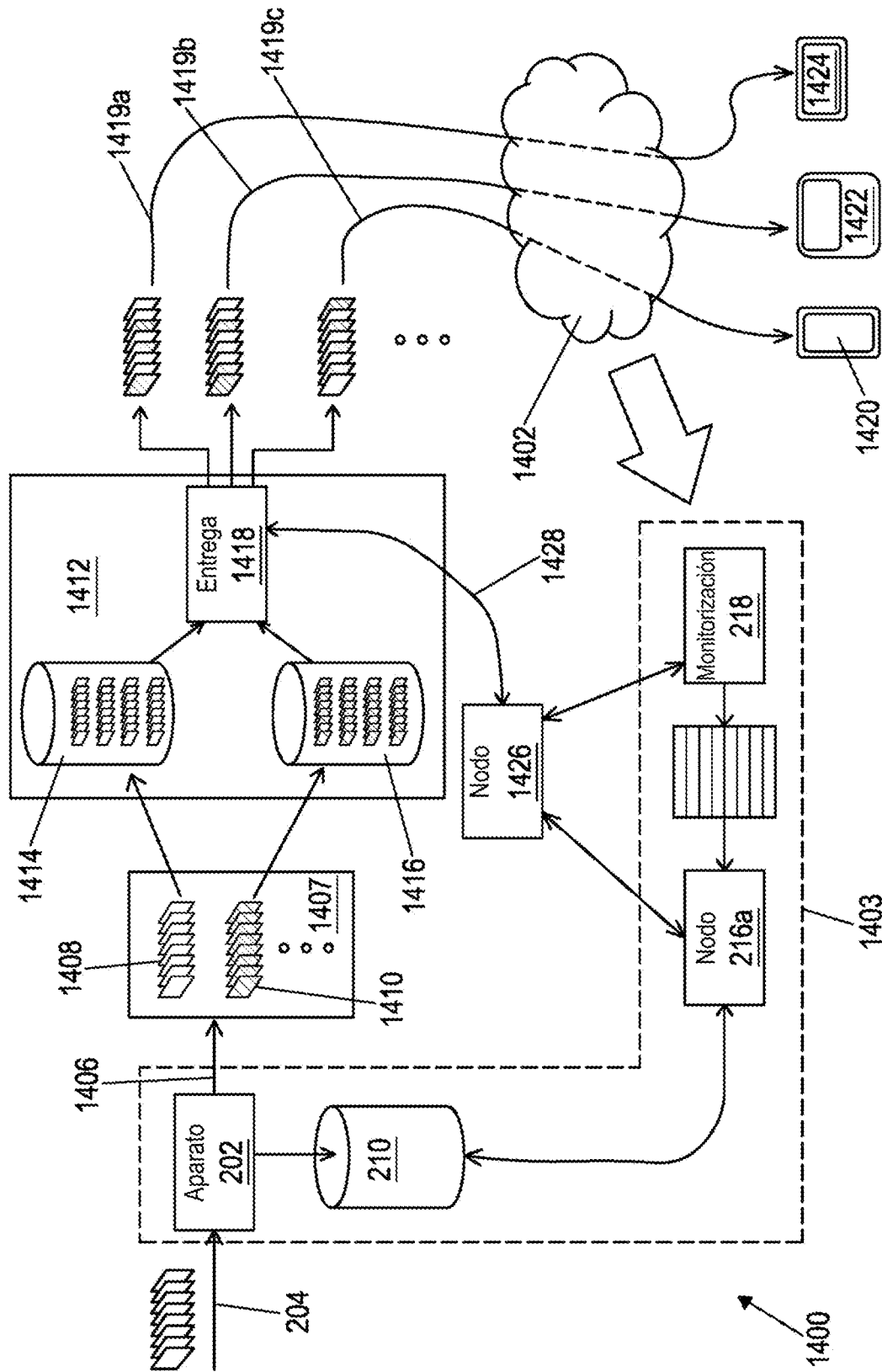


Figura 14

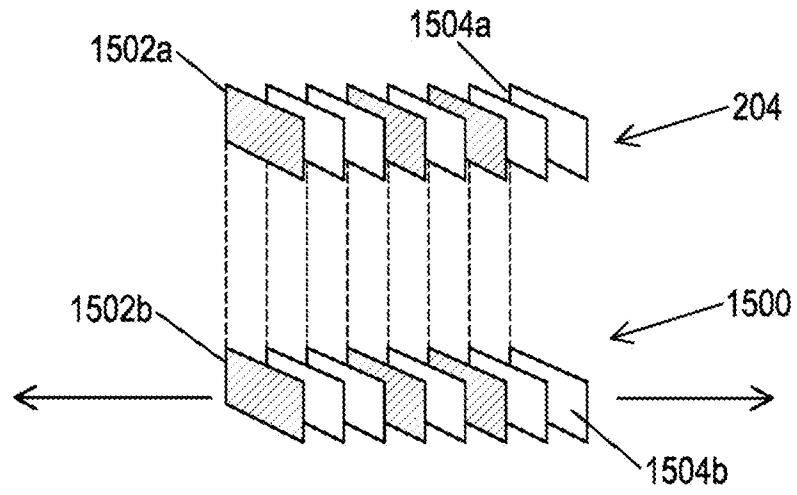


Figura 15a

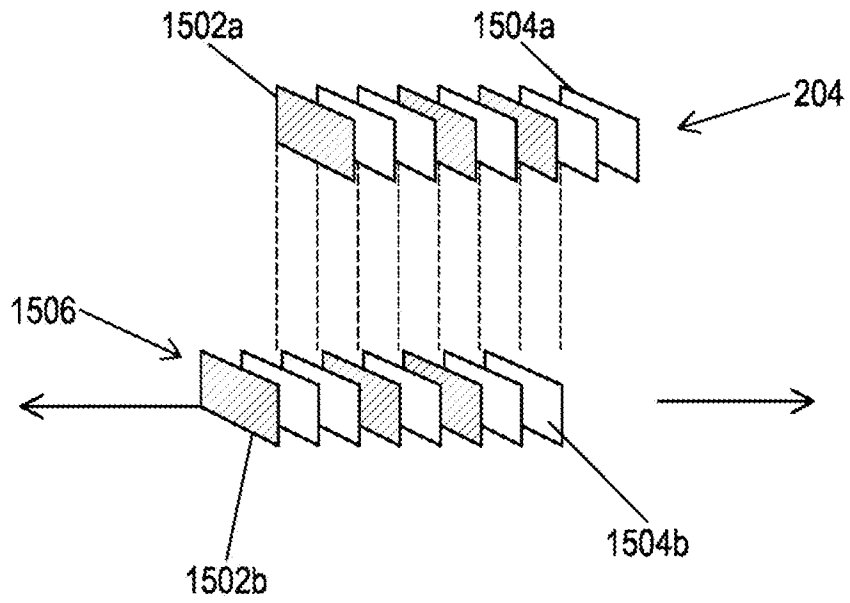


Figura 15b

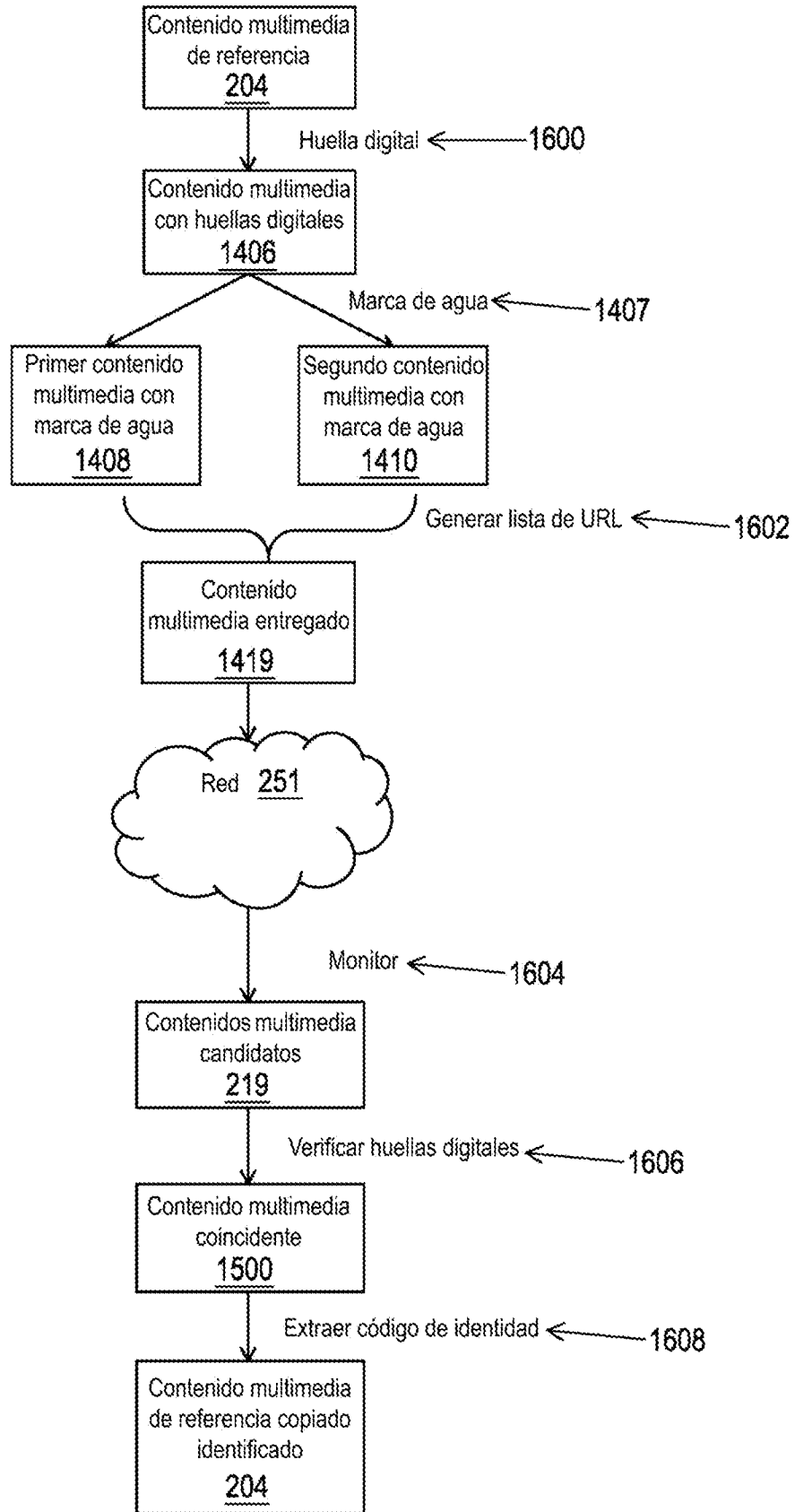


Figura 16