(19) OFICINA ESPAÑOLA DE PATENTES Y MARCAS

ESPAÑA

(11) Número de publicación: **2 712 605**

(51) Int. CI.:

G06F 21/10 (2013.01)
G11B 20/00 (2006.01)
G10L 19/018 (2013.01)
H04N 1/32 (2006.01)

(12) TRADUCCIÓN DE PATENTE EUROPEA T3

(96) Fecha de presentación y número de la solicitud europea: **13.01.2009 E 15176595 (5)**

(97) Fecha y número de publicación de la concesión europea: **05.12.2018 EP 2960819**

(54) Título: **Procedimiento y sistema que incorporan una huella digital no detectable en un archivo de medios digitales**

(45) Fecha de publicación y mención en BOPI de la traducción de la patente:
14.05.2019

(73) Titular/es:

CAPRICODE OY (100.0%)
Kiviharjunlenkki 1 E
90220 Oulu, FI

(72) Inventor/es:

LÖYTYNOJA, MIKKO;
BROCKMAN, MARKO;
KOUTANIEMI, JUKKA y
SEPPÄNEN, EERO

(74) Agente/Representante:

CURELL SUÑOL, S.L.P.

DESCRIPCIÓN

Procedimiento y sistema que incorporan una huella digital no detectable en un archivo de medios digitales.

Campo técnico de la invención

La invención se refiere a un soporte legible por ordenador que comprende archivos de medios digitales que comprenden un archivo de medios digitales con marcas de agua.

Antecedentes de la invención

Uno de los elementos habilitadores para la música en línea y móvil ha sido la gestión de derechos digitales (DRM). Esta proporciona los medios para proteger la propiedad de contenido y los derechos de autor limitando la distribución y el uso no autorizados. No obstante, las soluciones tradicionales de DRM han resultado ser controvertidas. Para evitar las copias de CD de audio se probaron diferentes técnicas, pero las mismas provocaron problemas de compatibilidad con tantos reproductores que la DRM ya no se usa en la distribución de CD de audio. En la música móvil, existen grupos independientes de fabricantes de reproductores musicales y minoristas de música en línea que usan diferentes técnicas de DRM, las cuales no tienen capacidad de interfuncionamiento. Desde la perspectiva del consumidor, esta no es una situación ideal, debido a que la música protegida por DRM y comprada de una tienda de música en línea puede que sea reproducible en reproductores de audio digital de solamente un fabricante.

El formato de música digital dominante es en la actualidad la Capa de Audio 3 del MPEG-1 (Grupo de Expertos en Imágenes en Movimiento), conocido más comúnmente como MP3. Esta es también la codificación estándar de facto de música reproducida en reproductores de audio digital. El problema del MP3 en relación con la distribución de música digital es que el mismo no soporta la protección de copias. Esto ha provocado que minoristas de música en línea usen otros formatos de audio privativos habilitados para la DRM. La finalidad consiste en hacer que el uso de los archivos de música de maneras no especificadas y permitidas por las discográficas resulte complicado. La mayoría de las soluciones actuales basadas en un cifrado se puede sortear grabando (*burning*) la música en un CD y, a continuación, extrayendo de nuevo la misma en algún formato no protegido, tal como el MP3.

Con el fin de crear una solución para el problema de la gestión de derechos de audio digital se pueden usar las marcas de agua digitales. La naturaleza de las marcas de agua permite que el audio se presente sin ningún cifrado debido a que la protección del contenido se incorpora en la propia señal de audio.

El uso de un formato de archivo no protegido permite que la música se reproduzca en cualquier productor de audio digital, y la música también se puede grabar (*burn*) fácilmente en un CD. Esto elimina muchos de los ataques usados en otros sistemas de DRM y permite que el consumidor quede más satisfecho debido a la mayor facilidad de utilización. No obstante, el problema es que las marcas de agua digitales pueden ser vulnerables a ataques por procesado de la señal. La señal con marcas de agua se puede modificar de manera que la modificación resulte inaudible para un oyente humano, aunque la señal de marca de agua se puede destruir en el proceso. Esto constituye un desafío fundamental para todas las aplicaciones de marcas de agua.

A un sistema que aplica un modelo de derechos se le denomina sistema de DRM 10. En la figura 1 se representa un ejemplo. Aunque la arquitectura de un sistema de DRM depende fuertemente del escenario de uso específico, existen algunos componentes comunes, los cuales se encuentran en la mayoría de los sistemas. A este tema común se le denomina arquitectura de referencia de DRM. Consiste en tres componentes principales: el servidor de contenido 11, el servidor de licencias 12 y el cliente 13.

El servidor de contenido 11 incluye una base de datos de contenido 111 para todos los archivos de contenido, y la funcionalidad 113 destinada a preparar contenido para una distribución controlada por DRM. Además del propio contenido, la base de datos almacena información de metadatos 112 sobre el contenido, tal como título, autor, formato y precio. Para usuarios finales, el servidor de contenido 11 permite acceso a las descargas de contenido habilitadas por DRM.

Los archivos de contenido habitualmente se manipulan de alguna manera con el fin de prepararlos para una distribución controlada cuando los mismos se importan al repositorio de contenido 111. Esto se realiza con el componente empaquetador de contenido del servidor de contenido. Todos los archivos que son traídos al sistema por los proveedores de contenido son procesados, en primer lugar, por el empaquetador de contenido 113 y, a continuación, se colocan en la base de datos de contenido para su almacenamiento. Otra tarea importante del empaquetador de contenido 113 es la especificación de derechos que desea permitir el proveedor de contenido para el usuario. Se pueden especificar derechos independientes con fines de previsualización, y al usuario se le pueden ofrecer varias opciones de compra. El empaquetador de contenido 113 puede ser, por ejemplo, una interfaz web que se ejecuta por encima del servidor que proporciona acceso a bases de datos para los proveedores de contenido.

Una de las características esenciales del empaquetador de contenido es el procesado por lotes. Puesto que los proveedores de contenido añaden generalmente una gran cantidad de contenido en una única sesión, debe resultar posible introducir múltiples archivos con modelos de derechos personalizables en el sistema.

El servidor de licencias 12 en un sistema de DRM 10 típico crea licencias por medio de un generador de licencias 123 para cada usuario a partir de derechos de contenido 121, identidades de usuario 124 y claves de cifrado de contenido 122. Los derechos 121 y las posibles claves de cifrado 122 son proporcionados por el servidor de contenido, y el cliente proporciona información sobre la identidad del usuario. Puesto que el trayecto de comunicaciones entre el servidor de licencias y el cliente es habitualmente inseguro, las transmisiones de datos se deben proteger con criptografía de clave pública.

Además de generar y transmitir licencias al cliente, el servidor de licencias 12 es responsable de la transacción financiera del proceso de concesión de licencias. El servidor de licencias usa la identidad del usuario para recuperar los detalles necesarios referentes a la transacción, tales como detalles de la tarjeta de crédito o de la cuenta. La identidad del usuario se puede crear a partir de un nombre de usuario, un número de seguridad social, o cualquier otra información que identifique de manera precisa al usuario.

La aplicación del lado de cliente de DRM 13 puede residir en una variedad de plataformas. La funcionalidad principal del cliente 13 está contenida en un controlador de DRM 131, el cual o bien puede ser un elemento de software independiente o bien puede estar integrado en la propia aplicación de reproducción del contenido. Las funciones principales del controlador de DRM son recopilar información de identidad 132 del usuario, obtener licencias 135 que comprenden derechos de usuario y claves de cifrado del servidor de licencias 12, autorizar a la aplicación de reproducción 133 a que tenga acceso al paquete de contenido 134 que comprende el contenido y metadatos, y llevar a cabo el posible descifrado del contenido. Adicionalmente, el controlador entrega las órdenes del usuario al servidor de licencias para solicitar licencias y comprobar las opciones de pago. El controlador de DRM debe soportar una criptografía de clave pública con vistas a una transmisión segura de datos entre el cliente 13 y el servidor de licencias 12.

Los escenarios de autorización de uso dependen de los modelos de derechos usados del contenido. El modelo básico autoriza al usuario a que tenga acceso al contenido 134 tantas veces como sea posible por una única cuota. Otros modelos pueden conceder o restringir acceso al contenido temporalmente en relación con las opciones de pago seleccionadas. Otra posibilidad consiste en restringir el número de reproducciones con una solución basada en contadores. La protección del contador de uso en el dispositivo de cliente sigue siendo un problema de implementación, especialmente en casos en los que al usuario no se le requiere que esté en línea cuando accede al contenido. Para obtener un almacenamiento seguro del contador de uso se han propuesto soluciones de informática de confianza y basadas en funciones *hash*.

El actor más importante en la industria de la DRM Móvil es la Open Mobile Alliance (OMA), la cual es un organismo de normalización que desarrolla normas abiertas para la industria de telefonía móvil.

El DRM 1.0 de la OMA fue el primer procedimiento de normalización industrial para proteger contenido móvil. El mismo fue aprobado en 2004, y actualmente es soportado en la mayoría de los teléfonos móviles del mercado. El objetivo del DRM 1.0 de la OMA es seguir prácticas de DRM comunes en el cumplimiento de requisitos y características especiales del dominio móvil, al mismo tiempo que proporcionando una funcionalidad básica con cierto nivel de seguridad. La versión 1.0 proporciona tres procedimientos para la protección y la entrega de contenido: bloqueo de reenvío, entrega combinada y entrega independiente.

En la primera revisión de la DRM, la OMA se centraba en las piezas constitutivas fundamentales de un sistema de DRM. La nueva DRM 2.0 de la OMA afronta los problemas de seguridad con características nuevas basadas en el procedimiento de entrega independiente.

El modelo de seguridad de la DRM 2.0 de la OMA se basa fuertemente en el agente de DRM del dispositivo de usuario. El propio contenido se empaqueta en un contenedor seguro similar cifrado con una clave simétrica de cifrado de contenido, pero utiliza, además, certificados de PKI (Infraestructura de Clave Pública) para una mayor seguridad. Todo dispositivo con soporte de DRM 2.0 de la OMA tiene un certificado de PKI individual con una clave pública y una privada. A continuación, todo objeto de tipo derechos se cifra con la clave pública del receptor antes de que se envíe a través de la red. El objeto de tipo derechos contiene la clave simétrica que se usa para descifrar los archivos de contenido reales.

La aplicación de marcas de agua digitales es un proceso en el que, en una señal anfitriona digital, la cual puede ser, por ejemplo, un vídeo, un audio o una imagen, se incorpora información. La marca de agua puede ser detectable o no detectable en función de la aplicación. La idea de utilizar una marca de agua extraíble, audible, para proteger contenido de audio se presentó en la publicación “Audio scrambling using removable watermarking”, de M. Löytynoja, N. Cvejic, y T. Seppänen, *Sixth International Conference on Information, Communications and Signal Processing (ICICS 2007*), Singapur, del 10 al 13 de diciembre de 2007.

En el documento US 2001/0044899 se divulga un procedimiento de transformación de marcas de agua de una señal de medios con marca de agua para adaptar la marca de agua a las restricciones de robustez y perceptibilidad de un entorno nuevo. El procedimiento representado de transformación de marcas de agua detecta la primera marca de agua digital en la señal de medios. A continuación, incorpora información de mensaje de la primera marca de agua digital a una segunda marca de agua digital diferente en la señal de medios en otro formato antes de que la señal de medios experimente un proceso de transformación. La segunda marca de agua digital está adaptada para sobrevivir al proceso de transformación.

Las marcas de agua digitales tienen tres características importantes que vienen determinadas por el tipo de aplicación: capacidad, robustez e imperceptibilidad. La capacidad es la cantidad de datos que se puede incorporar en la marca de agua, la robustez es la capacidad de la marca de agua para resistir modificaciones en la señal anfitriona, y la imperceptibilidad significa que la marca de agua no se puede detectar a partir de la señal anfitriona con los sentidos humanos. Estas características son parcialmente excluyentes, lo cual significa que algunas áreas se pueden acentuar aunque deteriorando otras.

Las marcas de agua se pueden incorporar en el audio en el dominio del tiempo o en algún dominio de transformada, tal como el dominio de Fourier. La selección del dominio afecta a las propiedades de la marca de agua en relación con la imperceptibilidad y la robustez. Las marcas de agua en el dominio de la frecuencia se consideran en general más inaudibles, pero son especialmente vulnerables contra modificaciones en frecuencia, tales como el desplazamiento de la altura tonal o la compresión dinámica. Las técnicas de aplicación de marcas de agua en el dominio del tiempo en general usan una aplicación de marcas de agua basada en espectros ensanchados. Otros dominios usados para la aplicación de marcas de agua de audio son el dominio de las ondículas y el dominio *cepstrum*, que es, básicamente, la transformada de Fourier del espectro de decibelios de la señal.

La aplicación de marcas de agua por espectro ensanchado significa que la potencia de la información de la marca de agua se ensancha deliberadamente ampliándola en el dominio de la frecuencia con el fin de ocultar la señal de forma más eficiente en la señal encubridora. En la aplicación de marcas de agua digitales se usan en general dos tipos de procedimientos de espectro ensanchado: los procedimientos de espectro ensanchado por saltos de frecuencia y por secuencia directa. El procedimiento por saltos de frecuencia se basa en una conmutación rápida de la frecuencia portadora de acuerdo con una secuencia seudoaleatoria, la cual debe conocerse en las fases tanto de incorporación como de extracción. El procedimiento de secuencia directa ensancha la señal de marca de agua obteniendo una señal de banda más ancha, creada también a partir de una secuencia seudoaleatoria.

En la aplicación de marcas de agua de espectro ensanchado por secuencia directa, la señal de marca de agua construida a partir de las secuencias seudoaleatorias se puede añadir a la señal encubridora simplemente adicionando o sustrayendo las muestras. Puesto que la secuencia seudoaleatoria es en general mucho más corta que la señal anfitriona, la secuencia se repite para cada bloque de la señal anfitriona. Uno de los procedimientos posibles consiste en adicionar la señal seudoaleatoria al bloque si el bit que se debe incorporar es uno, y sustraerla si el bit es cero. Este tipo de planteamiento mantiene la complejidad computacional del algoritmo de incorporación en un nivel muy bajo para facilitar un uso en tiempo real.

Un uso importante de los procedimientos de espectro ensanchado por secuencia directa en la aplicación de marcas de agua de audio es la sincronización. Es un procedimiento para determinar la ubicación exacta de la marca de agua en el proceso de extracción. La sincronización se puede llevar a cabo insertando la señal de sincronización una vez o bien en el comienzo de la secuencia de bloques o bien en el comienzo de cada bloque.

La señal de sincronización es habitualmente una señal de espectro ensanchado seudoaleatoria similar a los procedimientos de secuencia directa, excepto que la señal de sincronización puede ser mucho más larga. En el proceso de extracción, el punto de sincronización se calcula calculando la correlación cruzada de la señal de sincronización original y la señal con marca de agua. Si la marca de agua se incorpora con el procedimiento de saltos de frecuencia deben usarse señales de sincronización independientes.

La naturaleza del procedimiento de saltos de frecuencia es muy diferente a la del procedimiento de secuencia directa. En lugar de ser una señal de banda ancha, la marca de agua por saltos de frecuencia está presente en bandas muy estrechas en cualquier instante de tiempo dado. La frecuencia de la señal cambia rápidamente con el tiempo de acuerdo con una secuencia seudoaleatoria predefinida. La banda de saltos de frecuencia define límites para la secuencia de saltos. La secuencia seudoaleatoria que define la secuencia de saltos de frecuencia se puede usar como clave de marca de agua para proteger la ubicación exacta de la señal de marca de agua en los coeficientes de frecuencia.

En la figura 2 se presenta un ejemplo del procedimiento de saltos de frecuencia. El mismo divide el audio anfitrión en bloques de 1024 coeficientes de FFT, y selecciona dos coeficientes de acuerdo con la secuencia de saltos de frecuencia seudoaleatoria. El procedimiento cambia los valores de estos coeficientes por la media de

subbanda, la cual se calcula a partir de los coeficientes alrededor de los dos coeficientes. Si se incorpora un bit “uno”, la magnitud del coeficiente inferior 21 se fija a K decibelios mayor y el coeficiente superior 22 se fija a K decibelios menor. Si se incorpora un bit “cero”, el procedimiento es el opuesto. La intensidad de la marca de agua viene determinada directamente por el valor de K usado. Por lo tanto, K no puede ser mayor que la distancia desde el valor de media de subbanda al umbral de enmascaramiento de frecuencia con el fin de que la marca de agua permanezca por debajo del nivel de JND (Diferencia Apenas Perceptible).

Sumario de algunos ejemplos de la invención

El objetivo de la presente invención es proporcionar un procedimiento para incorporar una huella digital no detectable en un archivo de medios digitales. El archivo de medios digitales puede ser un archivo de audio, un archivo de vídeo o un archivo de imagen. Es también un objetivo de la invención proporcionar un sistema de entrega para los archivos de medios digitales con huella digital.

El objetivo de la presente invención se cumple proporcionando un soporte legible por ordenador de acuerdo con la reivindicación 1.

La idea básica de la invención es básicamente la siguiente: como ejemplo en un archivo de audio, la invención se puede utilizar de la manera siguiente. El esquema de protección según la invención combina tanto la marca de agua extraíble audible como huellas digitales inaudibles robustas, que se incorporan en el archivo de audio anfitrión. En primer lugar, en el archivo de audio se incorpora la marca de agua audible y extraíble, y, a continuación, el archivo se publica en Internet, desde donde puede ser descargado por los usuarios y, posiblemente, los mismos pueden compartirlo con otros usuarios.

Los usuarios pueden escuchar libremente el archivo de audio con marca de agua, el cual sirve como avance para el contenido real. La marca de agua se incorpora de tal manera que es claramente audible y reduce significativamente la calidad del audio, aunque permitiendo, al mismo tiempo, que el usuario obtenga una muestra de cómo sonaría el contenido sin la marca de agua.

Si al usuario le gusta la canción del archivo de audio, el mismo puede comprar la versión original simplemente descargando la clave de marca de agua que se usa para extraer la marca de agua audible del archivo de audio. El software de reproductor de acuerdo con la invención soporta el procedimiento de aplicación de marcas de agua usado, con el fin de poder extraer la marca de agua audible mientras se reproduce el contenido para el usuario por primera vez.

Cuando se extrae una señal de ruido (es decir, la marca de agua audible) del archivo de vista previa, al contenido del archivo de audio se le añade ventajosamente una huella digital individual del usuario. Esta huella digital del usuario individual se puede usar de manera ventajosa posteriormente para averiguar quién es el responsable de filtrar el contenido del archivo de audio para una distribución ilegal.

A partir de la descripción detallada que se ofrece posteriormente en la presente se pondrá de manifiesto un ámbito adicional de aplicabilidad de la presente invención. No obstante, deberá entenderse que la descripción detallada y los ejemplos específicos, aunque indican formas de realización preferidas de la invención, se proporcionan únicamente a título ilustrativo, ya que, para aquellos versados en la materia, se pondrán de manifiesto, a partir de esta descripción detallada, diversos cambios y modificaciones dentro del espíritu y del alcance de la invención.

Breve descripción de los dibujos

La presente invención se llegará a entender más minuciosamente a partir de la descripción detallada que se ofrece a continuación en la presente memoria y de los dibujos adjuntos que se aportan únicamente a título ilustrativo, y, por lo tanto, no limitan la presente invención, y en la que

la figura 1 muestra esquemáticamente una representación de una entrega DRM conocida en la técnica anterior;

la figura 2 muestra un ejemplo de una marca de agua audible incorporada en un archivo de audio usando el procedimiento de saltos de frecuencia;

la figura 3 muestra en forma de un diagrama de flujo a modo de ejemplo las etapas principales del procedimiento según la invención cuando se crea un archivo de audio con marca de agua en un servidor de contenido;

la figura 4 muestra en forma de un diagrama de flujo a modo de ejemplo las etapas principales del procedimiento según la invención cuando se crea un archivo de audio con huella digital en una aplicación de cliente;

la figura 5 muestra en forma de un diagrama de flujo a modo de ejemplo las etapas principales del procedimiento según la invención para hallar una fuente de una distribución ilegal de un archivo de audio; y

la figura 6 muestra un ejemplo de una huella digital no audible incorporada en un archivo de audio usando el procedimiento de saltos de frecuencia.

Descripción detallada

En la siguiente descripción, las formas de realización consideradas son meramente a modo de ejemplo, y alguien versado en la materia puede encontrar otras formas de implementar la invención. Aunque la memoria puede referirse a “cierta”, “una” o “alguna” forma(s) de realización en diversos lugares, esto no significa necesariamente que cada una de estas referencias se haga con respecto a la(s) misma(s) forma(s) de realización, o que la característica se aplique solamente a una única forma de realización. Características individuales de formas de realización diferentes también se pueden combinar para proporcionar otras formas de realización.

Las figuras 1 y 2 se describieron en combinación con la descripción de la técnica anterior.

El algoritmo de aplicación de huellas digitales de la presente invención se puede dividir en tres fases principales: incorporación, transformación de ruido y detección de huella digital. La figura 3 representa la fase de incorporación, la figura 4 representa la fase de transformación de ruido (es decir, extraer una marca de agua detectable e insertar una huella digital de usuario no detectable) y la figura 5 representa cómo un propietario de derechos puede averiguar quién está distribuyendo ilegalmente un archivo de audio.

Las etapas principales del procedimiento para incorporar una marca de agua detectable en un archivo de audio se representan en un diagrama de flujo a modo de ejemplo de la figura 3. En la fase de incorporación, una marca de agua extraíble se inserta en el audio original con el fin de producir la versión de vista previa distribuible. El algoritmo de incorporación puede combinar varias técnicas digitales de aplicación de marcas de agua, tales como aplicación de marcas de agua de espectro ensanchado por saltos de frecuencia y por secuencia directa.

Las entradas del proceso son el archivo de audio original no comprimido 301 y la clave seudoaleatoria 304 para mejorar la seguridad de la marca de agua. En primer lugar, el archivo original 301 se divide en bloques de 1024 muestras, etapa 302, y cada bloque se procesa por separado a partir de este momento en adelante. Una muestra de bloque de audio que comprende 1024 muestras se representa con la referencia 311.

Para el bloque de audio 311 en cuestión se materializa una Transformada Rápida de Fourier 312. La FFT 312 proporciona una matriz de coeficientes de FFT complejos 313. Tomando valores absolutos 314 de los coeficientes de FFT complejos, también se pueden expresar ventajosamente magnitudes absolutas 315 de los coeficientes de FFT en decibelios 316.

Ventajosamente una incorporación de una marca de agua 317 se puede realizar modificando, de manera ventajosa, dos coeficientes de frecuencia de la muestra de archivo de audio que se pueden definir por medio de una secuencia de saltos de frecuencia seudoaleatoria 306. La secuencia de saltos seudoaleatoria se logra con un Generador Congruencial Lineal (LCG) 305 que usa, como entradas, parámetros de banda de frecuencia 303 y una clave seudoaleatoria 304. La banda de frecuencia de saltos de frecuencia seudoaleatoria puede comprender, por ejemplo, 512 coeficientes de frecuencia.

Un par de coeficientes de frecuencia modificados se puede seleccionar, ventajosamente, de manera que sea cinco coeficientes superior al coeficiente seleccionado por la secuencia de saltos de frecuencia. El coeficiente inferior se puede modificar con un modificador de – K y el coeficiente superior se puede modificar con un modificador de + (K/2). El valor de K es ventajosamente el valor del valor de K aleatorio 333.

Para modificar las magnitudes de los coeficientes de FFT extraídos 316, se selecciona un valor de K aleatorio, referencia 333, usando un generador aleatorio 332 de entre un intervalo [min_k, max_k], referencia 331, con etapas de 0.1. Este parámetro define la cantidad de ruido en dB que se adicionará ventajosamente a un bloque de audio actual. Para cada bloque de audio se usa un valor de K aleatorio diferente 333. Los valores de K usados se pueden almacenar, ventajosamente, para un uso posterior en una matriz específica 351.

Usando el valor aleatorio K 333 y los coeficientes de FFT seleccionados por la secuencia de saltos de frecuencia seudoaleatoria 306, en la fase 318 pueden definirse valores de escalado concretos para el bloque de audio en cuestión. Los valores concretos de los valores de escalado k_1 y k_2 dependen del valor aleatorio K de la muestra de audio, referencia 333.

En la etapa 320, los valores de escalado definidos k_1 y k_2, referencia 319, se usan para modificar los dos coeficientes de FFT definidos de la matriz de FFT compleja original 313. Los dos coeficientes definidos en la

matriz de FFT compleja se escalan de acuerdo con los valores de escalado definidos k_1 y k_2 con el fin de producir una matriz de FFT compleja 321 con ruido detectable añadido. La matriz de FFT modificada 321 es similar al ejemplo representado en la figura 2 en donde dos coeficientes de FFT, números 36 y 41 de entre 512 coeficientes de FFT, se transforman para añadir una marca de agua en una muestra de audio.

A continuación, el bloque de audio con marca de agua, con ruido, se transforma al dominio del tiempo usando una IFFT (Transformada Rápida Inversa de Fourier) en la etapa de 322. El resultado es un bloque de audio 323 en el dominio del tiempo, que comprende un archivo de audio con una señal de ruido detectable.

Las etapas 311 a 333 se repiten para todos los bloques de audio que comprenden, cada uno de ellos, 1024 muestras. El valor aleatorio usado K 333 y la secuencia de saltos seudoaleatoria 306 se pueden cambiar después de cada bloque de audio procesado. Esto significa que las posiciones de los coeficientes de FFT con ruido no son iguales en todos los bloques de audio, y que los valores de escalado k_1 y k_2 también pueden variar de una muestra de audio a otra.

En la etapa 341, todos los bloques de audio modificados se juntan y se realiza un escalado de nivel final para el archivo de audio completo con el fin de evitar problemas de recorte de la señal. El resultado es un archivo de audio distribuible 342.

La etapa final 343 consiste en añadir una señal de sincronización de espectro ensanchado 309 por medio de un generador de señales de sincronismo 308. El generador de señales de sincronismo 308 construye una señal de sincronización 309 usando parámetros de sincronización definidos 307. La señal de sincronización 309 se incorpora ventajosamente en el comienzo de la secuencia de bloques para facilitar el proceso de sincronización en la fase en la que el ruido se extrae del archivo de audio. La señal de sincronización 309 se puede añadir al comienzo de cada muestra de audio o se puede usar solamente una señal de sincronización en el comienzo del archivo de audio 342. Por ejemplo, como señal de sincronización se puede usar una señal de espectro ensanchado de 16 384 muestras limitadas a una banda de frecuencia de 10 a 20 kHz. La misma se puede incorporar en el comienzo de la señal de audio con una intensidad de 0.03.

El proceso de aplicación de marcas de agua finaliza en una etapa en la que un archivo de audio 361 con una marca de agua está preparado para su publicación en Internet. Para extraer el ruido posteriormente (es decir, la marca de agua), se deben almacenar la clave seudoaleatoria 304 y los cambios definidos de los coeficientes de FFT en dB (una matriz de K valores 351). Estos parámetros forman la clave de marca de agua para el archivo de audio. Adicionalmente, se debe almacenar la señal de sincronización de espectro ensanchado 309 usada.

La figura 4 representa la fase de transformación de ruido de la presente invención. La fase de transformación de ruido comprende la transformación de una marca de agua detectable del archivo de audio en una huella digital de usuario no detectable. Las etapas principales del procedimiento para transformar una marca de agua detectable en una huella digital no detectable en un archivo de audio se representan en un diagrama de flujo a modo de ejemplo de la figura 4. Una transformación de una marca de agua a huella digital de usuario se puede lograr en un aparato eléctrico de varios tipos. La invención se puede materializar en cualquier tipo de aparato que comprenda una unidad de procesador y suficiente memoria para guardar un programa de ordenador utilizado en la transformación. El aparato puede ser, por ejemplo, un ordenador, un teléfono celular, un asistente personal digital (PDA), un receptor digital de televisión, un receptor digital de radio, un reproductor de MP3, etc.

Los parámetros requeridos para crear una licencia para un usuario y modificar el archivo de audio con marca de agua y distribuible en un archivo de audio con una huella digital exclusiva son: clave seudoaleatoria exclusiva del archivo de audio 304, banda de frecuencia 303 para el ruido de marca de agua (por ejemplo, banda de frecuencia 1-512 de la figura 2), una matriz de cambios de dB realizados en el archivo de audio 351 durante la aplicación de las marcas de agua, intensidad pretendida de la huella digital en dB, identificación de usuario del comprador y señal de sincronización 309 y su escala.

La clave seudoaleatoria 333 y los parámetros de banda de frecuencia 303 deben presentar los mismos valores que se usaron en la adición de la marca de agua en el archivo de audio. La matriz de cambios de dB 351 también se toma de los datos almacenados en la operación de adición de la marca de agua. La intensidad de la huella digital determina directamente la calidad del archivo de audio resultante. Es la cantidad de ruido que queda en la canción después de extraer el ruido de marca de agua de la muestra distribuible. Este ruido sobrante forma la huella digital de usuario individual, que contiene la identificación de usuario del comprador en el sistema.

Cuando el usuario contacta con un servidor de una tienda musical, en primer lugar debe identificarse con una identificación de usuario exclusivo. A continuación, esta identificación de usuario, durante la transformación de ruido, se codifica en la matriz de cambios de dB (una matriz de K valores) del archivo de audio con huella digital.

La incorporación de la huella digital se puede realizar incrementando o decrementando valores de escalado k_1 y k_2 usados en la aplicación de la marca de agua del archivo de audio. El parámetro de intensidad de la huella digital define la magnitud en la que se cambian los valores de dB. En una forma de realización ventajosa de la

invención, los valores de dB se incrementan si el bit incorporado es “uno”, y se decrementan si el bit es “cero”.

Antes de incorporar la huella digital se puede usar una corrección de errores hacia delante para una mayor fiabilidad. Además de la matriz de cambios de dB, a los datos de licencia se les añade la clave seudoaleatoria del archivo de audio. Estos dos elementos forman la licencia exclusiva del usuario.

El proceso de transformación se puede dividir en tres etapas principales: sincronización, procesado de bloques y combinación del audio resultante. La señal de audio con marca de agua se debe sincronizar antes de que se pueda extraer el ruido de ella. La sincronización se realiza adoptando una correlación cruzada entre el audio y la señal de sincronización original. El valor máximo de la correlación es la deriva de sincronización. Después de que se haya encontrado la deriva de sincronización, la señal de sincronización ya no es necesaria, y la misma se puede extraer de la señal de audio. Se puede extraer sustrayendo la señal de sincronización original, escalada, a partir del punto de deriva de sincronización en el archivo de audio distribuible 361.

La sincronización determina también el punto inicial de la secuencia de marca de agua. El procedimiento de sincronización puede utilizar técnicas de aplicación de marcas de agua de espectro ensanchado por secuencia directa. La sincronización puede resultar necesaria debido a que diferentes codificadores de compresión con pérdidas, por ejemplo la codificación MP3, pueden añadir algunas muestras adicionales en el comienzo del archivo de audio en la fase de codificación. La señal de sincronización se extrae, ventajosamente, del archivo de audio después de que se haya localizado el punto inicial con el fin de lograr una mayor calidad de audio.

En la sincronización, la etapa 402, una aplicación de cliente sincroniza un archivo de audio con marca de agua usando una señal de sincronismo 309. El resultado es un archivo de audio sincronizado 403. El archivo de audio sincronizado se puede dividir en bloques de audio de 1024 muestras. Ventajosamente, cada bloque de audio se procesa por separado.

La secuencia de saltos de frecuencia se genera a partir de la clave de inicialización seudoaleatoria 304. La secuencia se limita con los mismos parámetros que los usados en la aplicación de las marcas de agua. La secuencia resultante es igual a la secuencia generada en el proceso de incorporación de marcas de agua de la figura 3.

El audio sincronizado se divide en bloques de 1024 muestras 410 que comienzan desde el punto de deriva de sincronización. Cada bloque de audio 410 se procesa ventajosamente por separado a partir de este momento en adelante. Un proceso de FFT 411 transforma la muestra de audio en una matriz de FFT compleja 412. A continuación, en la etapa 413, se toman valores absolutos de cada coeficiente de FFT. Este proceso da como resultado las magnitudes 414 de los coeficientes de FFT. Las magnitudes de los coeficientes de FFT se transforman a dB en la etapa 415.

A continuación, el valor de K para el bloque de audio actual se lee de la matriz de cambios de dB 352. La matriz 352 comprende versiones modificadas de la matriz de K valores 351 usada en la aplicación de las marcas de agua. Este elemento de matriz 352 contiene, ventajosamente, valores de escalado modificados k_1 y k_2. Utilizando estos valores de escalado modificados, el ruido de aplicación de marcas de agua se transforma ventajosamente en una huella digital de usuario.

A continuación, el valor de K correspondiente al bloque de audio actual se lee de la matriz de cambios de dB 351. Este elemento de matriz contiene las modificaciones realizadas en el bloque respectivo del audio original, que dan como resultado el ruido de aplicación de marcas de agua.

En las etapas 416 a 419, el ruido de aplicación de marcas de agua se extrae en primer lugar modificando aquellas magnitudes de coeficientes de FFT en el dominio de los decibelios, que se usaron en la aplicación de marcas de agua de la muestra de audio. Después de esto, los mismos coeficientes de FFT se modifican con valores de escalado nuevos que generan menos ruido que los correspondientes usados en la aplicación de la marca de agua. Los valores de escalado usados no dejan ruido audible en el archivo de audio. Los valores de escalado nuevos de la aplicación de la huella digital se modifican también para que contengan la huella digital del usuario.

La figura 6 representa un ejemplo de un bloque de audio en el que los coeficientes de FFT 36 y 41 se transfieren de la marca de agua a una huella digital. Las diferencias entre los coeficientes de FFT originales de la muestra de audio y la muestra de audio con huella digital, referencias 61 y 62, son menores que las diferencias de los coeficientes de FFT originales de la muestra de audio y la muestra de audio con marca de agua.

En la etapa 421 se lleva a cabo una IFFT sobre la matriz de FFT con huella digital 420. La transformada da como resultado un bloque de audio con huella digital de 1024 muestras. A continuación, el bloque de audio 422 se concatena con los otros bloques de audio del mismo archivo de audio.

Cada bloque de audio del archivo de audio se procesa, ventajosamente, por separado. Cuando se han

transformado todos los bloques de audio del archivo de audio, un archivo de audio con huella digital 432 está listo para ser escuchado.

La transformación de ruido concreta desde ruido a una huella digital se realiza cuando los coeficientes de FFT se modifican con los valores de matriz K 352. Esto es posible porque los valores de matriz K no son exactamente los mismos en la fase de aplicación de la huella digital en comparación con los valores que se almacenaron en el servidor en la fase de incorporación de la marca de agua. Los mismos son modificados ligeramente por el servidor de tal manera que los valores de matriz K contienen una huella digital no detectable del usuario. La identificación del usuario en la tienda musical se puede usar en calidad de los datos de huella digital. Esto significa que el servidor debe generar una matriz de K exclusiva cada vez que un cliente nuevo compra una licencia correspondiente a un archivo de audio, debido a los diferentes datos de huella digital.

Una ventaja de este tipo de proceso es que el archivo de audio no se encuentra nunca en un estado desprotegido, ya que se transforma directamente desde la versión de vista previa gratuita, con marca de agua, a la versión de usuario con huella digital sin ninguna etapa adicional en medio. Resulta también cómodo para el usuario ya que el mismo no tiene que descargar la canción nuevamente después de realizar la compra. Por el contrario, solamente necesita adquirir la licencia y esperar a que se complete el proceso de transformación de ruido local.

Las etapas principales del procedimiento para leer una huella digital de un archivo de audio se representan en un diagrama de flujo a modo de ejemplo de la figura 5.

Antes de leer una huella digital de un archivo de audio, el archivo de audio debe ser identificado. Después de eso, se puede extraer una clave seudoaleatoria correcta K a partir de la matriz de valores K 351.

La sincronización 501 del archivo de audio con huella digital 432 se puede realizar con respecto a un archivo de audio original 301. Se calcula una correlación cruzada entre la señal de audio con huella digital y la señal de audio original. El valor máximo de la correlación es la deriva de sincronización. Si el archivo de audio con huella digital tiene muestras adicionales en el comienzo, las mismas se recortan de manera que el audio original y el audio con huella digital estén sincronizados cuando el propietario de los derechos digitales comience a leerlos ambos en la primera muestra.

La secuencia de saltos seudoaleatoria usada en la modificación de los coeficientes de FFT se genera en primera instancia a partir de la clave de inicialización seudoaleatoria 333 y de los parámetros de banda de frecuencia 303.

A continuación, tanto el archivo de audio con huella digital sincronizado 502 como el archivo de audio original 301 se dividen en bloques que comprenden 1024 muestras de audio (referencias 503 y 511). Los bloques se transforman 512 con una FFT que da como resultado una matriz de FFT compleja 513. Las magnitudes de los coeficientes de FFT se calculan tomando los valores absolutos 514 de los coeficientes de FFT complejos. A continuación, las magnitudes de la FFT se transforman, ventajosamente, en el dominio de los dB, referencia 515.

La lectura de la huella digital se puede realizar comparando los pares de coeficientes de FFT del archivo de audio original 301 y el archivo de audio con huella digital 432, etapa 516. El coeficiente de FFT inferior del par se lee de la secuencia de saltos de frecuencia y el coeficiente superior es, ventajosamente, cinco coeficientes mayor.

La integración sobre todos los valores de bits e intensidades en la etapa 517 se puede lograr de la siguiente manera. A partir de estos pares de FFT se pueden calcular, ventajosamente, dos valores de comparación. El primer valor es una magnitud de coeficiente de FFT inferior del archivo de audio con huella digital, a la que se ha sustraído una magnitud de coeficiente de FFT inferior del archivo de audio original. El segundo valor es una magnitud de coeficiente de FFT superior del archivo de audio con huella digital, a la que se ha sustraído una magnitud de coeficiente de FFT superior del archivo de audio original. El bit de huella digital extraído de este bloque de 1024 muestras es 1 si el primer valor es mayor que el segundo valor, y 0 si el segundo valor es mayor que el primer valor. Este proceso se repite con todos los bloques de audio correspondientes de 1024 muestras del archivo de audio con huella digital y el archivo de audio original.

La matriz de bits de huella digital resultante 518 se divide en bloques del tamaño del bloque de corrección de errores hacia adelante 519 utilizado. Por ejemplo, si se usa el código Hamming más simple (7, 4), el tamaño de bloque es 7. Después de la descodificación, la matriz de bits con errores corregidos se divide ventajosamente en bloques de 32 bits. Estos bloques son las matrices de bits de huella digital 520 reales que presentan la identificación de usuario. Si se requiere una corrección de errores adicional, el elevado número de huellas digitales nos permite seleccionar la matriz de bits de huella digital, más común, o bien bit a bit o bien palabra a palabra.

Aunque el procedimiento de aplicación de huellas digitales en las figuras 3, 4 y 5 se representa en el contexto de

un archivo de audio, resultará evidente para alguien versado en la materia que la invención se puede usar también en el contexto de un archivo de vídeo o un archivo de imágenes.

Cualquiera de las etapas del proceso que se han descrito o ilustrado anteriormente se puede implementar usando instrucciones ejecutables en un procesador de propósito general o de propósito especial y almacenadas en un soporte de almacenamiento legible por ordenador (por ejemplo, disco, memoria o similares) para que sean ejecutadas por dicho procesador. Debe entenderse que las referencias a “soporte de almacenamiento legible por ordenador” y “ordenador” abarcan circuitos especializados tales como matrices de puertas programables in situ, circuitos integrados de aplicación específica (ASIC), unidades de almacenamiento *flash* USB, dispositivos de procesado de señales y otros dispositivos.

Habiéndose descrito la invención según la manera anterior, resultará evidente que la misma se puede variar de muchas formas. Por ejemplo, en la aplicación de marcas de agua y la aplicación de huellas digitales se pueden utilizar más coeficientes de frecuencia que el ejemplo representado de dos coeficientes de frecuencia.

REIVINDICACIONES

1. Soporte legible por ordenador que comprende un archivo de medios digitales con marca de agua detectable por los sentidos humanos, en el que:

- la aplicación de marcas de agua detectable por los sentidos humanos se logra mediante una aplicación de marcas de agua por saltos de frecuencia modificándose por lo menos dos coeficientes de frecuencia (21, 22) para generar ruido detectable por los sentidos humanos en el archivo de medios digitales (361), y que

- la aplicación de marcas de agua detectable por los sentidos humanos está configurada para transformarse por medio de información específica (304, 351) del archivo de medios emitida por un propietario de derechos de medios digitales en un dispositivo de cliente, directamente en una huella digital de usuario, individual, no detectable por los sentidos humanos (61, 62) transformando los por lo menos dos coeficientes de frecuencia (21, 22) de la marca de agua detectable por los sentidos humanos directamente en una huella digital de usuario no detectable por los sentidos humanos en el archivo de medios digitales durante el primer uso del archivo de medios digitales por medio de la información específica (304, 351) del archivo de medios digitales,

caracterizado por que la información específica (304, 351) del archivo de medios digitales comprende una clave seudoaleatoria (304) que define los coeficientes de frecuencia utilizados, una matriz de una variable aleatoria (352), basada en una identificación de usuario, que define una intensidad de escalado de los coeficientes de frecuencia utilizados (61, 62) en la aplicación de la huella digital e información de señal de sincronización (309).

2. Soporte legible por ordenador según la reivindicación 1, caracterizado por que un punto inicial de una marca de agua se incorpora en el archivo de medios digitales en una señal de espectro ensanchado (309).

3. Soporte legible por ordenador según la reivindicación 1, caracterizado por que el archivo de medios digitales (361) es uno de los siguientes: un archivo de audio, un archivo de vídeo o un archivo de imágenes.

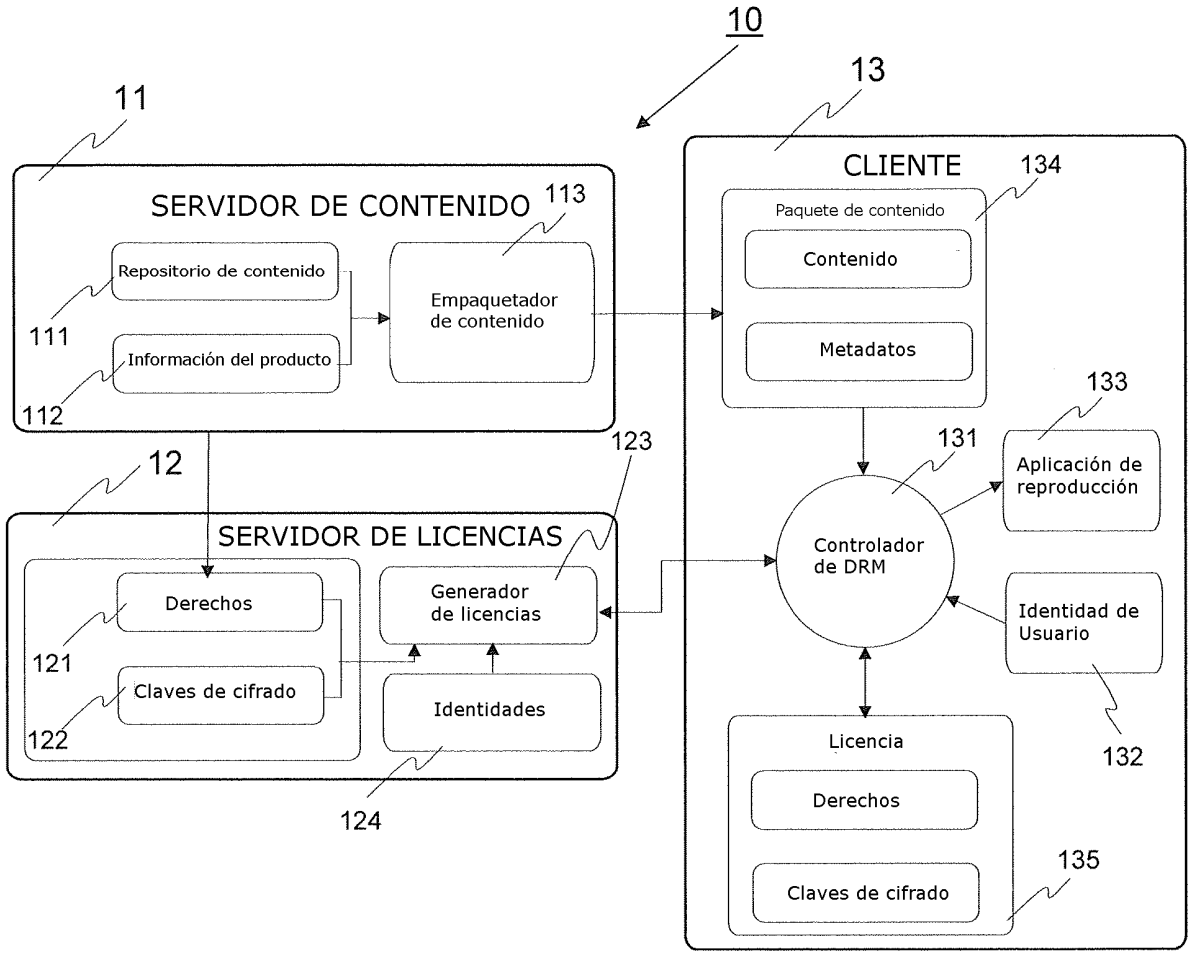

Fig. 1 TÉCNICA ANTERIOR: sistema de DRM

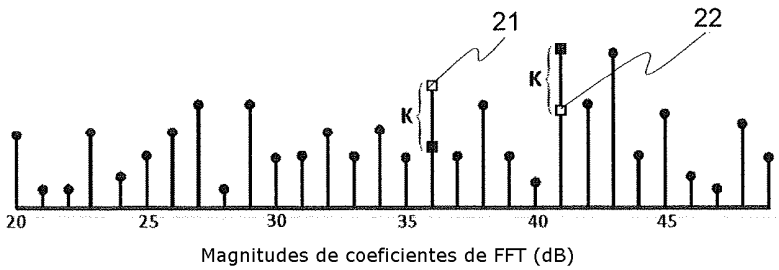

Fig. 2

Magnitud del coeficiente mapeado antes de la incorporación de la marca de agua

Magnitud del coeficiente mapeado después de la incorporación de la marca de agua

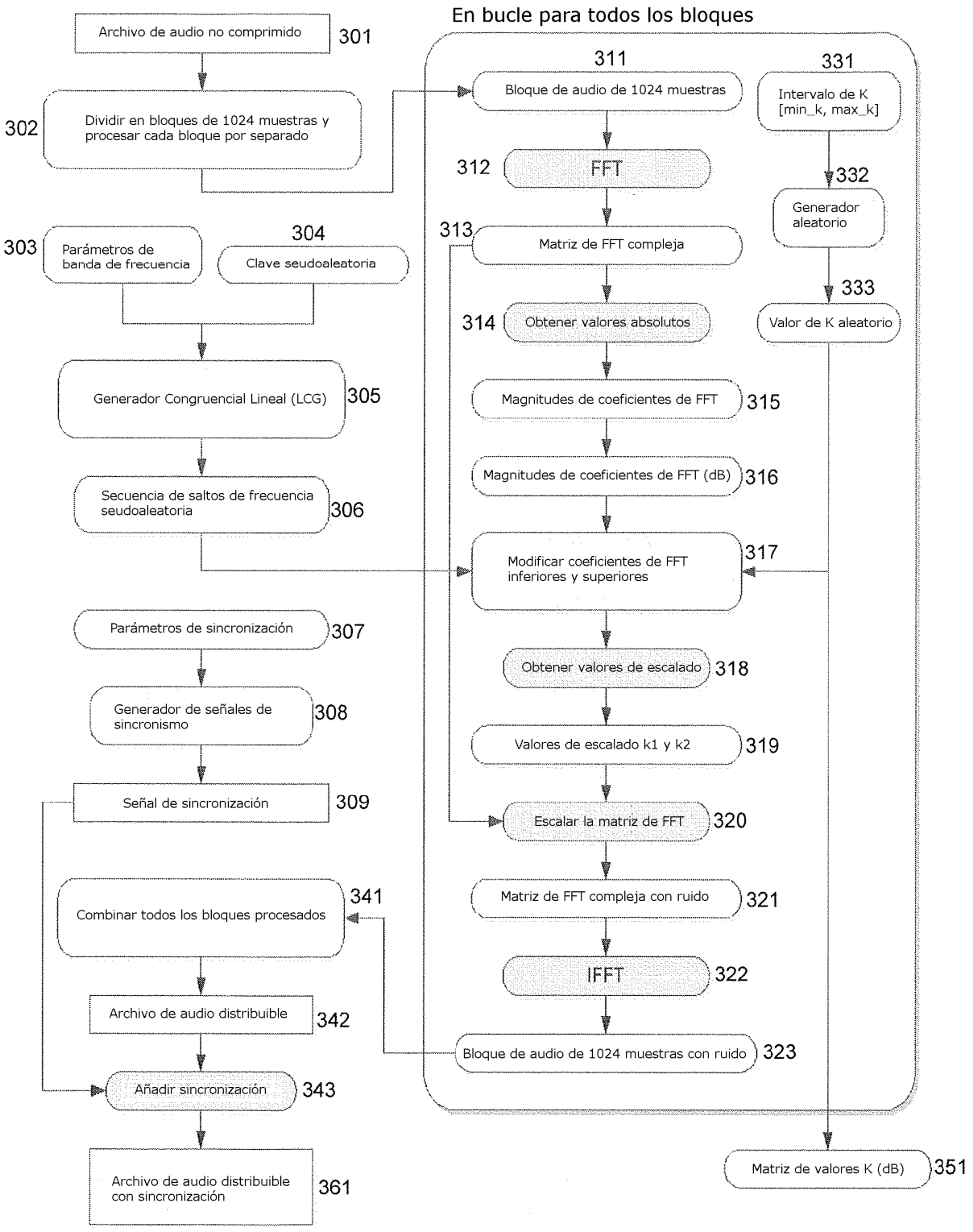
En bucle para todos los bloques
Archivo de audio no comprimido
301
Dividir en bloques de 1024 muestras y procesar cada bloque por separado
302
Parámetros de banda de frecuencia
303
Clave seudoaleatoria
304
Generador Congruencial Lineal (LCG)
305
Secuencia de saltos de frecuencia seudoaleatoria
306
Parámetros de sincronización
307
Generador de señales de sincronismo
308
Señal de sincronización
309
Combinar todos los bloques procesados
341
Archivo de audio distribuible
342
Añadir sincronización
343
Archivo de audio distribuible con sincronización
361
311
Bloque de audio de 1024 muestras
312
FFT
313
Matriz de FFT compleja
314
Obtener valores absolutos
Magnitudes de coeficientes de FFT
315
Magnitudes de coeficientes de FFT (dB)
316
Modificar coeficientes de FFT inferiores y superiores
317
Obtener valores de escalado
318
Valores de escalado k1 y k2
319
Escalar la matriz de FFT
320
Matriz de FFT compleja con ruido
321
IFFT
322
Bloque de audio de 1024 muestras con ruido
323
331
Intervalo de K [min_k, max_k]
332
Generador aleatorio
333
Valor de K aleatorio
Matriz de valores K (dB)
351

Fig. 3

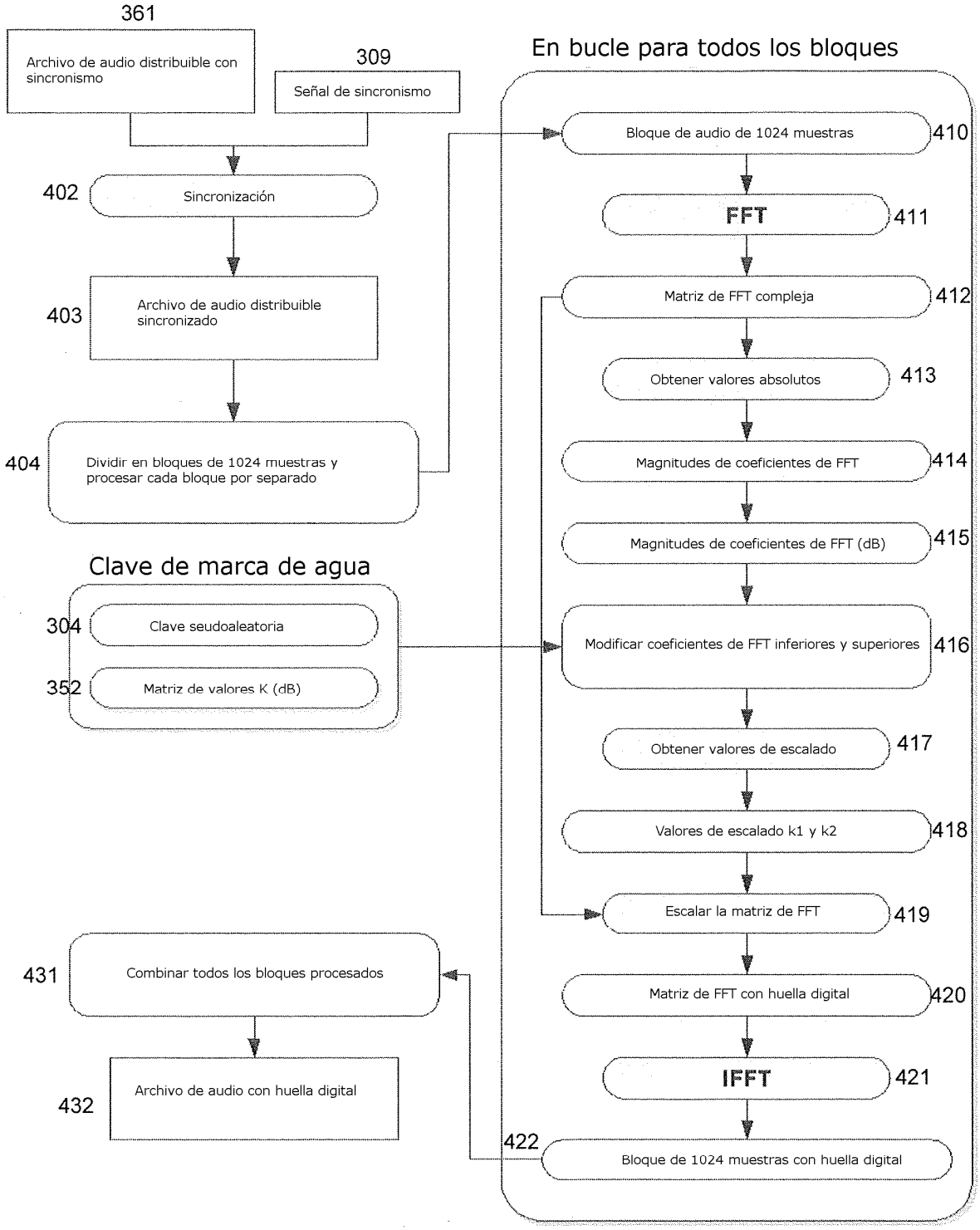
361
Archivo de audio distribuible con sincronismo
309
Señal de sincronismo
En bucle para todos los bloques
Bloque de audio de 1024 muestras
410
402
Sincronización
FFT
411
Matriz de FFT compleja
412
403
Archivo de audio distribuible sincronizado
Obtener valores absolutos
413
Magnitudes de coeficientes de FFT
414
404
Dividir en bloques de 1024 muestras y procesar cada bloque por separado
Magnitudes de coeficientes de FFT (dB)
415
Clave de marca de agua
304
Clave seudoaleatoria
Modificar coeficientes de FFT inferiores y superiores
416
352
Matriz de valores K (dB)
Obtener valores de escalado
417
Valores de escalado k1 y k2
418
Escalar la matriz de FFT
419
431
Combinar todos los bloques procesados
Matriz de FFT con huella digital
420
IFFT
421
432
Archivo de audio con huella digital
422
Bloque de 1024 muestras con huella digital

Fig. 4

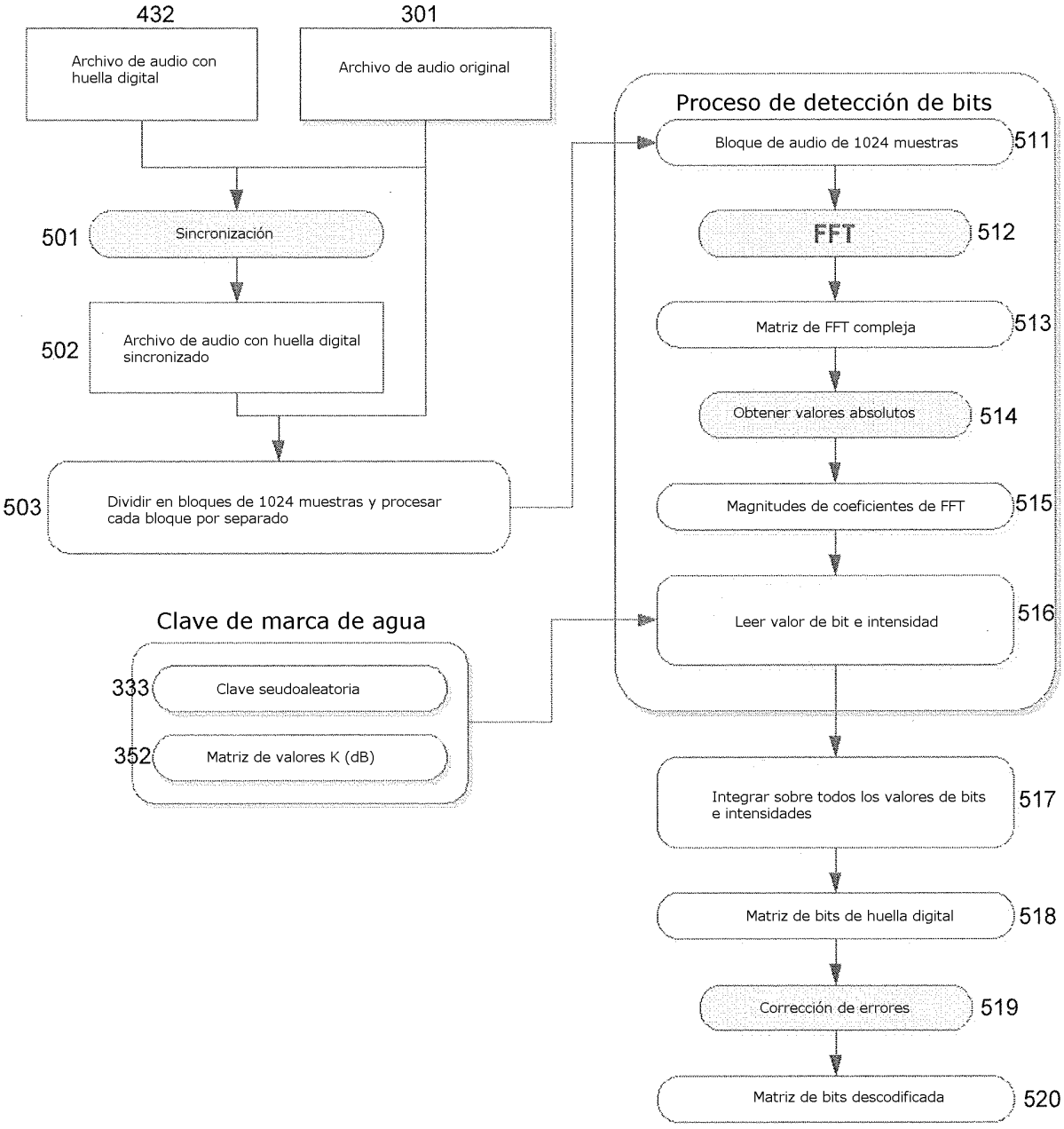
432
Archivo de audio con huella digital
301
Archivo de audio original
501
Sincronización
502
Archivo de audio con huella digital sincronizado
503
Dividir en bloques de 1024 muestras y procesar cada bloque por separado
Clave de marca de agua
333
Clave seudoaleatoria
352
Matriz de valores K (dB)
Proceso de detección de bits
Bloque de audio de 1024 muestras
511
FFT
512
Matriz de FFT compleja
513
Obtener valores absolutos
514
Magnitudes de coeficientes de FFT
515
Leer valor de bit e intensidad
516
Integrar sobre todos los valores de bits e intensidades
517
Matriz de bits de huella digital
518
Corrección de errores
519
Matriz de bits descodificada
520

Fig. 5

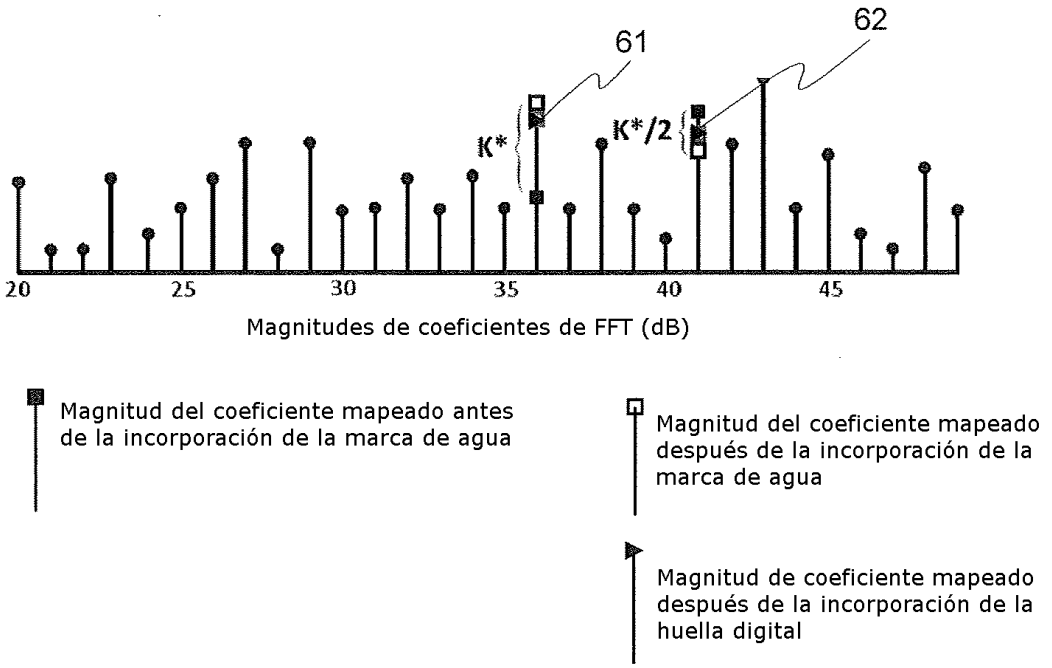
61
62
K*
K*/2
20
25
30
35
40
45
Magnitudes de coeficientes de FFT (dB)
Magnitud del coeficiente mapeado antes de la incorporación de la marca de agua
Magnitud del coeficiente mapeado después de la incorporación de la marca de agua
Magnitud de coeficiente mapeado después de la incorporación de la huella digital

Fig. 6