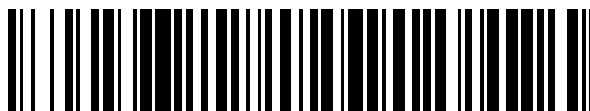


19



OFICINA ESPAÑOLA DE
PATENTES Y MARCAS

ESPAÑA



11 Número de publicación: **2 712 131**

51 Int. Cl.:

G06F 17/00	(2006.01)
G06F 5/00	(2006.01)
G10H 1/00	(2006.01)
G10H 1/06	(2006.01)
G10L 25/00	(2013.01)
G10L 19/00	(2013.01)
H03M 7/30	(2006.01)
G10L 19/032	(2013.01)
G06F 17/30	(2006.01)

12

TRADUCCIÓN DE PATENTE EUROPEA

T3

- 86 Fecha de presentación y número de la solicitud internacional: **16.08.2012 PCT/CA2012/050562**
- 87 Fecha y número de publicación internacional: **28.02.2013 WO13026155**
- 96 Fecha de presentación y número de la solicitud europea: **16.08.2012 E 12825060 (2)**
- 97 Fecha y número de publicación de la concesión europea: **19.12.2018 EP 2745211**

54 Título: **Método de formalización y estructuración de información multinivel y multiestructural y aparato asociado**

30 Prioridad:

19.08.2011 US 201161525629 P
18.01.2012 US 201261588101 P

45 Fecha de publicación y mención en BOPI de la traducción de la patente:
09.05.2019

73 Titular/es:

GENERAL HARMONICS INTERNATIONAL INC.
(100.0%)
First Floor, Hastings House, Balmoral Gap,
Hastings, BB14034
Christ Church, BB

72 Inventor/es:

ZHIRKOV, ALEXANDER;
ORAEVSKY, ALEXEY;
GRICHINE, ANDREI;
BLONDHEIM, GEORGE;
WANDINGER, MAX y
ATTWOOD, WADE

74 Agente/Representante:

CAMPELLO ESTEBARANZ, Reyes

ES 2 712 131 T3

Aviso: En el plazo de nueve meses a contar desde la fecha de publicación en el Boletín Europeo de Patentes, de la mención de concesión de la patente europea, cualquier persona podrá oponerse ante la Oficina Europea de Patentes a la patente concedida. La oposición deberá formularse por escrito y estar motivada; sólo se considerará como formulada una vez que se haya realizado el pago de la tasa de oposición (art. 99.1 del Convenio sobre Concesión de Patentes Europeas).

DESCRIPCIÓN

Método de formalización y estructuración de información multinivel y multiestructural y aparato asociado

5 **SOLICITUDES RELACIONADAS**

Esta solicitud reivindica la prioridad de las solicitudes provisionales de Estados Unidos número 61/525.629 y 61/588.101, presentadas el 19 de 2011 y el 18 de enero de 2012.

10 **ANTECEDENTES DE LA DESCRIPCIÓN**

Campo de la descripción

Los aspectos de la presente descripción se refieren al procesamiento de información para lograr la formalización y
 15 estructuración, incluido el análisis de audio, la manipulación y la representación, y, más particularmente, a sistemas y métodos de análisis estructurado y la determinación de la relación entre el valor de la información y la cantidad de información relacionada con los datos configurados armónicamente, incluidos medios digitales.

Descripción de la técnica relacionada

20 El artículo COIFMAN ET AL: "Entropy-Based Algorithms for Best Basis Selection", IEEE TRANSACTIONS ON INFORMATION THEORY, IEEE PRESS, EE.UU., vol. 38, n.º 2, 1 de marzo de 1992 (01-03-1992), páginas 713-718, XP002316486, ISSN: 0018-9448, DOI: 10.1109/18.119732 describe un método que permite la compresión eficiente de una diversidad de señales tales como sonido e imágenes.

25 Un estándar generalmente reconocido para el concepto de audio unificado u otra formalización de datos digitales puede no estar generalmente disponible en la técnica, aunque pueden implementarse diversas técnicas para lograr lo mismo. Por ejemplo, algunas técnicas emplean representaciones de audio simplificadas de una señal de sonido, que incluyen principalmente el reconocimiento de voz y la síntesis de voz, así como la compresión de datos digitales
 30 que representan música. En un aspecto, las tecnologías del habla han progresado desde la representación de señales de sonido a través de la forma de onda correspondiente, aunque dichas técnicas a menudo funcionan sobre la base de palabras o incluso de frases completas en los datos de voz. Tal base en palabras/frases representa una forma de información que está más cerca de la percepción natural del cerebro humano. En contraste, las tecnologías de formalización implementadas para la representación de música generalmente usan solo una representación de
 35 información físicamente perceptiva, es decir, en una forma cercana a la de la percepción física del sonido por el oído humano.

Hasta la fecha, otros intentos de obtener una representación universal de la música más abstracta siguen siendo en
 40 gran medida infructuosos. Dicha falta de éxito puede atribuirse, por ejemplo, al hecho de que la información del habla incluye una forma de lenguaje primario y la sintaxis de su descripción, diseñada con matemáticas precisas que han sido establecidas y afinadas por muchas generaciones. En contraste, las representaciones de música existentes, por ejemplo, basadas en notas o muestras de sonido, son relativamente primitivas en comparación con el análisis de la información del habla y, por lo general, tales representaciones no son universalmente aplicables.

45 En este sentido, una de las representaciones digitales de sonido relativamente más informativas actualmente disponibles puede ser el formato PCM, generalmente denominado audio sin comprimir. Sin embargo, aunque tal formato puede ser relativamente más informativo, tal informatividad se compensa con un tamaño del archivo de datos relativamente grande. El gran tamaño del archivo de datos, a su vez, puede hacer que tal formato o
 50 representación sea inadecuada o impracticable, por ejemplo, para una entrega/transmisión rápida y/o almacenamiento compacto. Si se desean dichos atributos, se han desarrollado o se están desarrollando representaciones más compactas, aunque probablemente menos informativas, que generalmente emplean un enfoque popular para la reducción de datos, tal como se utiliza, por ejemplo, en MP3, OGG, WMA y otros modelos o representaciones psicoacústicas clásicas. Sin embargo, los sonidos naturales incluyen más redundancia que las representaciones/modelos típicos de señales de audio que son capaces de analizar de manera eficaz. Además, la
 55 percepción humana de la música es generalmente mucho más compleja que cualquier modelo psicoacústico existente. Como tal, existe la necesidad de un enfoque mejorado para la formalización del sonido que sea capaz de representar el sonido, el audio, la música y/o cualquier otro dato digital relacionado con los armónicos de una manera más compacta (es decir, menos intensiva en datos), pero de manera más informativa (es decir, en términos de la integridad de la representación que se puede proporcionar).

En los esfuerzos para satisfacer esta demanda, se han desarrollado representaciones/modelos más progresivos que se están empleando actualmente, por ejemplo, en MP3-Pro, HE AAC, MP3 PlusV, MPEG-4 SSC, audio estructurado MPEG 4 y MIDI. MP3-Pro y HE AAC utilizan esencialmente las peculiaridades de la percepción humana como la base para extraer elementos de estructura en una señal de audio, sin conservar la búsqueda específica de fase y similitud en la señal. Las frecuencias bajas se replican en las frecuencias altas, sin conservar la fase, pero conservando el principio de similitud y los parámetros generales de sonido, tal como la conservación de la energía y la naturaleza caótica de la señal. MP3-PlusV extrae, almacena y genera armónicos, sin conservar la fase, y también se puede aplicar para determinar la parte de alta frecuencia de la señal. MPEG-4 SSC (codificación sinusoidal) es un método para representar la señal como un conjunto de objetos organizados, tales como armónicos, aciertos y ruido. Sin embargo, tal método para extraer estos objetos de la señal es diferente al esquema de percepción que tiene lugar naturalmente con el cerebro humano. Por lo tanto, la reproducción de la señal de esta representación/modelo puede incluir artefactos no deseados. El audio estructurado MPEG-4 intenta representar el sonido mediante un algoritmo unificado que es capaz de generar una diversidad de estructuras de sonido. Si bien este enfoque puede tener cierto potencial, la creación de tal algoritmo puede ser problemática debido a los recursos computacionales requeridos. El formato MIDI generalmente requiere un tamaño de archivo de datos relativamente más pequeño, pero, al igual que el audio estructurado MPEG-4, es una representación/modelo que generalmente es adecuado para la escritura manual de música, y no para la representación de sonidos de origen natural y/o composiciones de audio ya creadas.

Por lo tanto, aunque se hacen esfuerzos más avanzados para reducir o eliminar la redundancia perceptiva inherente en una señal de audio, en comparación con los métodos o modelos psicoacústicos clásicos, las representaciones/modelos de sonido existentes estructurados y orientados a objetos pueden tender a perder el alcance de la informatividad de la señal inicial a una tasa de bits baja (es decir, gran cantidad de información) y, por lo tanto, no pueden conservar la calidad de la señal de audio original con un grado aceptable de tolerancia, ya que la representación/modelo de sonido produce un tamaño de archivo de datos más compacto. Por consiguiente, existe la necesidad de un esquema de formalización y disposición para los medios digitales, tal como el audio, que sea capaz de reducir la cantidad de información o la tasa de bits de un archivo de información digital mediante una estructuración apropiada, mientras que conserva un valor de información dentro de un umbral de o incluso mayor que el archivo de información digital original.

BREVE RESUMEN DE LA DESCRIPCIÓN

Las necesidades anteriores y otras se satisfacen mediante aspectos de la presente descripción, en donde uno de dichos aspectos se refiere al uso de estructuras altamente organizadas y parámetros como métricas de señal de audio solamente. Otro aspecto incluye un método de crecimiento sucesivo de una diversidad de estructuras de sonido a partir de primotrones de información u otros elementos de datos para crear estructuras complejas, pero más compactas (con respecto al tamaño del archivo de datos) y ricas en información. Las estructuras de nivel de orden superior consisten en primotrones de información de sonido que pueden ser localizados en el tiempo, localizados en la frecuencia o tener una localización híbrida (es decir, tiempo-frecuencia), en donde la informatividad en combinación con el tamaño relativamente bajo del archivo de datos proporcionado por aspectos de la tecnología descrita.

Como se usa en el presente documento, un "primotró" se puede definir como elementos o estructuras de datos representativos primarios que incluyen objetos o entidades multinivel y multiestructurales (MSML) para la formalización digital de medios tales como sonido o audio, que pueden utilizarse para crear estructuras complejas, pero más compactas (con respecto al tamaño del archivo de datos), ricas en información. Las estructuras de nivel de orden superior de los primotrones de información de sonido pueden ser localizadas en el tiempo, localizadas en la frecuencia o tener una localización híbrida (es decir, tiempo-frecuencia). Los primotrones localizados en el tiempo pueden generar ataques y aciertos, y un ritmo de nivel superior (con respecto al audio), mientras que los primotrones localizados en la frecuencia son el origen de los armónicos, un modelo de sobretono de nivel superior, y melodías recurrentes de nivel superior. Generalmente, un primotró, determinado por aspectos de un método de formalización como se describe en el presente documento, representa un patrón o combinación de uno o más bits o elementos de datos en los diversos niveles y/o dimensiones de la matriz de bits del archivo de audio digital, y se deduce que el archivo de audio digital puede representarse posteriormente mediante una combinación de dichos patrones identificados y/o determinados a incluir en el mismo. La evolución o "historia de vida" de dichos patrones puede caracterizarse adicionalmente por los principios de evolución de la información en un sistema caótico dinámico definido por un número finito de estructuras, concretamente, los primotrones tienen cada uno una "vida útil" que puede variar desde una fracción de segundo hasta varios minutos dentro del tiempo o la duración del tiempo-

- frecuencia del archivo de audio digital. Debido a que el archivo de audio digital ahora está representado por tal combinación o sistema vivo de primotrones, en donde cada primotón representa un subconjunto de bits de la matriz de bits para u otra estructura o entidad asociada con ese archivo de audio digital, el archivo de audio MSML que incluye la combinación identificada de primotrones puede presentar, por ejemplo, un tamaño del archivo de datos
- 5 reducido (cantidad de información) en comparación con el archivo de audio digitalizado original (al mismo tiempo que conserva la calidad de la información del archivo de audio digital dentro de un umbral perceptivo del archivo de audio digital original). Dichos primotrones pueden facilitar la reducción en la cantidad de información o la tasa de bits para el archivo de audio digital en particular, al tiempo que conservan la informatividad del archivo de audio digital dentro de un umbral perceptivo del archivo de audio digital original. Por consiguiente, tal formato de datos digitales
- 10 multifuncional y altamente estructurado puede configurarse para un almacenamiento y reproducción altamente compactos y eficaces de cualquier material representado en forma de una señal de audio de alta fidelidad de base o construcción armónica (incluyendo música, voz y otros medios basados en un sonido rico), dentro de un umbral de la señal de datos digital original.
- 15 Desde otra perspectiva, los aspectos de la presente descripción proporcionan sistemas y métodos para representar una composición de sonido como una formalización multinivel y multiestructural (MSML), en donde cualquier nivel de estructura relativamente superior de la composición de sonido puede configurarse como una integración de los elementos de un nivel relativamente inferior de la composición del sonido utilizando, por ejemplo, los criterios de información perceptiva (PIC). Los PIC se consideran satisfechos, por ejemplo, si la información perceptiva de un nivel es menor que la suma de los valores de información de los elementos constituyentes independientes de ese nivel. Un nivel cero se define como un nivel en el que la cantidad de información es equivalente a la "entropía perceptiva". Los niveles de orden superior incluyen primotrones de información de sonido (o estructuras de sonido primarias) que pueden estar localizadas en el tiempo, localizadas en la frecuencia y/o tener una localización híbrida. Las privaciones de información pueden generar una estructura de sonido de orden superior, que puede incluir
- 20 fonemas para señales de voz. Los primotrones localizados en el tiempo pueden generar ataques y aciertos, y un ritmo de nivel superior. Los primotrones localizados en la frecuencia pueden asociarse con el origen de los armónicos, después los modelos de sobretono de nivel superior, y después las melodías recurrentes de nivel superior. Una vez construida, la formalización MSML de cada composición o señal de sonido puede estar compuesta por, transformarse con o en, o combinarse de otra manera con la formalización MSML de otra composición o señal
- 25 de sonido, formando de este modo estructuras semánticas MSML de nivel superior más complejas.
- 30

- Más particularmente, los aspectos de la presente descripción proporcionan métodos y sistemas para estructurar la información. Dichos métodos y sistemas generalmente implican el análisis de un archivo de información digital original para determinar una cantidad de información original y un valor de información original asociado con la
- 35 misma. Después, se aplica un proceso de manipulación inicial al archivo de información digital original para formar un primer archivo de información digital resultante, y después se puede aplicar un proceso de manipulación posterior al primer archivo de información digital resultante para formar un segundo archivo de información digital resultante. Cada proceso de manipulación está configurado para eliminar al menos un elemento del archivo de información digital procesado, y/o representar una combinación de una pluralidad de elementos del archivo de información digital
- 40 procesado con un elemento representativo y una primera indicación asociada con una interrelación entre el elemento representante y al menos uno de la pluralidad de elementos en la combinación, a fin de reducir la cantidad de información y estructurar el archivo de información digital procesado. El elemento eliminado y/o el elemento representativo se determinan para reducir el valor de información del archivo de información digital procesado en no más de un umbral seleccionado con respecto al archivo de información digital original. Dichos procesos de
- 45 manipulación se aplican sucesivamente al archivo de información digital previamente resultante hasta que las aplicaciones de proceso de manipulación sucesivas no logran una reducción de umbral en la cantidad de información en el archivo de información digital resultante posterior sobre el archivo de información digital previamente resultante. Por lo tanto, el último archivo de información digital resultante se designa con una estructura primaria con una cantidad de información reducida con respecto a la cantidad de información original y un valor de
- 50 información resultante dentro del umbral seleccionado del valor de información original.

- De acuerdo con otro aspecto, pueden proporcionarse sistemas y métodos para el ancho de banda y la reducción de la velocidad de bits de los datos de audio utilizando una interfaz de codificación sin fases (proceso de manipulación) para una formalización de composición de sonido MSML, en donde la interfaz está configurada para considerar una
- 55 jerarquía del valor de información perceptiva. El método de reducción del ancho de banda y la tasa de bits utiliza la interfaz para realizar tareas dirigidas a la reducción de datos seleccionados, al tiempo que conserva la informatividad o el valor de información de todas las estructuras de sonido asociadas con la composición del sonido, por ejemplo, a través de la comparación con métricas de sonido conocidas. Tal método realiza una comparación de objetos de sonido de diferente nivel en el espacio, centrándose en determinar la integridad de las estructuras y el grado de

degradación de las mismas, dentro de una representación de cantidad reducida de la señal después de la filtración o la aplicación de un proceso de manipulación, lo que facilita la selección de un parámetro adecuado para cada elemento de sonido. La interfaz de codificación sin fases está configurada de manera apropiada de tal forma que la dependencia del tiempo de la señal no se conserve, y las estructuras de sonido se conserven en lugar de la forma de onda de la señal, dando como resultado de este modo una disminución del ancho de banda esencial y la tasa de bits. Para lograr una mayor velocidad de ancho de banda de audio y reducción de la tasa de bits, las estructuras de nivel relativamente más informativas de la formalización de composición de sonido MSML están configuradas para mantener parámetros constantes. Los parámetros de las estructuras de nivel relativamente menos informativas se guardan con menos precisión o se generan durante la decodificación con el uso de valores aleatorios.

10

De acuerdo con otro aspecto, pueden proporcionarse sistemas y métodos para el análisis de la estructura de datos de audio, utilizando una interfaz de análisis de estructura para una formalización de composición de sonido MSML, para la extracción de elementos repetitivos difusos de dichas composiciones musicales. Tal interfaz de análisis de estructura está configurada para procesar datos de audio recibidos de la formalización MSML de la señal original e identificar elementos de bucle repetitivos difusos dentro de un nivel relativamente bajo de la formalización MSML, para su uso en una mayor reducción de datos y sin disminuir la informatividad general de la señal. Tal interfaz comprende tres módulos. El primer módulo proporciona la búsqueda, el posicionamiento y el marcado de longitud de todos los elementos repetitivos difusos a lo largo de todo el nivel inicial de la formalización MSML. El segundo módulo extrae un componente de sonido general para cualquier conjunto de fragmentos repetitivos difusos y define el nivel de cohesividad de los datos formando una señal residual correspondiente. El tercer módulo realiza una construcción predictiva del fragmento de audio resultante integrando los elementos repetitivos extraídos, los datos de marcado correspondientes y la información residual. Se utiliza un método de filtrado para el modelado de fases para minimizar la información de la fase inicial requerida durante el proceso de construcción de sonido. Esta interfaz está configurada para ser utilizada como un elemento para una diversidad de aplicaciones, incluyendo, pero sin limitación, motores de búsqueda de música, visualización de música, y tecnologías de compresión profunda de audio.

Otro aspecto está dirigido a sistemas y métodos para la compresión de audio utilizando estructuras repetitivas de jerarquía en una formalización de composición de sonido MSML. La interfaz de codificación de repetición sin fases avanzada utiliza estructuras formadas por primotrones de sonido de macro-repetición del nivel superior dentro de una jerarquía de sonido de formalización MSML significativa. Se utilizan dos tipos de datos para el proceso de decodificación: datos a gran escala que definen la estructura general dentro de la formalización MSML y datos que comprenden fragmentos de señales originales descritos por un modelo de nivel relativamente inferior. Para comprimir los fragmentos de señal, se puede utilizar un codificador psicoacústico general, en algunos casos, posiblemente utilizando un enmascaramiento de señal externo. La reducción de la velocidad de bits se produce debido a la eliminación de fragmentos repetitivos construidos en el nivel superior de formalización MSML, mientras se almacenan las ubicaciones de estos fragmentos. La equivalencia perceptiva de fragmentos repetitivos permite una reducción en el número de muestras de referencia.

Otro aspecto más está dirigido a los sistemas y métodos para la reducción del ancho de banda de audio y la tasa de bits utilizando una interfaz de compresión sin pérdida para una formalización de composición de sonido MSML al reducir el número de matrices de valores enteros asociadas con la misma a través del uso de un codificador de entropía avanzado que realiza solo operaciones enteras, sin divisiones, y estimaciones de probabilidad secundarias. Dado que el codificador de rango y el modelo contextual de la interfaz no utilizan la operación matemática de división, permite la aplicación de esta interfaz a/la ejecución de la interfaz mediante microordenadores de bajo coste que no tienen una operación de división. Para mejorar la fiabilidad del modelo contextual de la interfaz, se emplea una estimación de probabilidad secundaria. La estimación de probabilidad secundaria es un modelo contextual complejo que utiliza la probabilidad predicha como un contexto con respecto a otro modelo. La estimación de probabilidad secundaria bidimensional también se puede utilizar para mezclar dos modelos diferentes. La estimación de probabilidad secundaria bidimensional es una modificación de la estimación de probabilidad secundaria, que tiene dos probabilidades de entrada y las utiliza como contexto. También se pueden utilizar contextos enteros adicionales.

Otro aspecto más está dirigido a sistemas y métodos para mejorar la compresión de señales de audio. También se proporcionan sistemas y métodos para mejorar la codificación aritmética y el modelado contextual para aumentar las relaciones de compresión y las velocidades de procesamiento, y así permitir que los datos espectrales se compriman con alta eficacia y mayor velocidad. El método de compresión puede ser completamente automatizado y no requiere inicialización previa para diferentes tipos de datos de audio. Es lo suficientemente flexible para ajustarse a diferentes tamaños de datos de audio espectral, lo que permite su uso con diferentes transformaciones espectrales. En lugar

de un codificador aritmético estándar, se utiliza un codificador de rango más eficiente. El modelado de contexto se aplica al flujo de datos, los modelos algorítmicos construidos y la optimización algorítmica de una función de codificador y/o de decodificador. Este aspecto también se basa, al menos parcialmente, en el uso de técnicas de codificador de rango adaptativo que implican aumentar la probabilidad del valor codificado. Para mejorar la fiabilidad del modelo contextual, se puede emplear una estimación de probabilidad secundaria.

Otros aspectos incluyen sistemas y métodos para la transmisión eficiente de datos de audio utilizando una interfaz de acelerador avanzado de datos de audio para una formalización de composición de sonido MSML, y utilizando un proceso de pérdida de calidad nula (NQL) para reducir el número de coeficientes espectrales de tiempo-frecuencia cuantificados (QMDCT), para proporcionar una entrega más rápida del audio original sin una degradación esencial de la calidad en el sonido resultante. El proceso NQL realiza una clasificación de las estructuras de sonido dentro de la formalización MSML dependiendo de su importancia en función de los límites de la percepción humana. La avanzada interfaz del acelerador reconstituye los coeficientes espectrales cuantificados de tiempo-frecuencia del formato inicial, sin la descuantificación de los datos ni la transcodificación inversa a través del formato PCM. La avanzada interfaz del acelerador divide el archivo codificado inicialmente de coeficientes espectrales de tiempo-frecuencia en tres grupos, donde los coeficientes de un grupo espectral de baja frecuencia permanecen inalterados, mientras que algunos coeficientes de un grupo de frecuencia intermedia se ponen a cero utilizando el proceso NQL, y los coeficientes del tercer grupo (frecuencias más altas) se eliminan reemplazándolos con los subgrupos de coeficientes más similares, uniéndolos en grupos bajos e intermedios, y ahorrando una referencia a este subgrupo, así como factores integrales, sin afectar a los elementos de sonido estructurales (es decir, armónicos y aciertos).

También se proporcionan sistemas y métodos, de acuerdo con otro aspecto más, para reducir el tamaño de los datos de audio, utilizando una interfaz de codificación de transformadas afines para una formalización de composición de sonido MSML, mientras se conserva la integridad total de la reproducción de sonido original en forma de transformada afín de los primotrones de sonido similares, lo que conduce a una disminución en la entropía de la señal (es decir, una estructuración más alta). La interfaz afín asociada con la formalización MSML se usa como una adición o complemento de cualquier procedimiento de codificación de audio para aumentar la relación de compresión, así como el proceso de compresión del núcleo, para la clase de señales donde los microfragmentos similares afines de tiempo-frecuencia contienen una parte esencial de la información de sonido. Esta interfaz de codificación de transformadas afines, mientras que se utiliza junto con la formalización MSML, está configurada para utilizar microfragmentos espectrales similares de la señal de audio digital para aumentar la relación de compresión de la señal. Durante el procesamiento, el rango espectral completo se divide en sub-bandas. Se realiza una búsqueda de fragmentos espectrales similares en la parte anterior de la señal de sonido de forma independiente y utilizando una etapa diferente en cada sub-banda. Se usa una transformada afín de cambio de amplitud, tiempo y espectral durante la búsqueda de fragmentos similares. Cada fragmento encontrado se resta del fragmento espectral original y el residuo se procesa adicionalmente como una diferencia de componentes espectrales, eliminando de este modo una redundancia significativa de los datos espectrales.

En resumen, los aspectos de la presente descripción proporcionan sistemas y métodos dirigidos a una formulación de múltiples niveles y estructura múltiple de un sonido u otra composición de medios digitales, así como diversos sistemas y métodos para determinar un patrón y/o combinación de objetos primarios o entidades en la formalización para proporcionar un tamaño del archivo de datos reducido en comparación con el archivo de audio digitalizado original (es decir, una cantidad reducida de información o tasa de bits), al mismo tiempo que se conserva la máxima informatividad (es decir, la calidad de la información) de los elementos determinados, y sistemas y métodos para reconstruir o decodificar el patrón/combinación de objetos primarios o entidades (elementos) para sintetizar el sonido original u otra composición de medios digitales. Por lo tanto, los aspectos de la presente descripción abordan las necesidades identificadas y proporcionan otras ventajas como se detallada de otro modo en el presente documento.

BREVE DESCRIPCIÓN DE LAS VARIAS VISTAS DE LOS DIBUJOS

Habiendo descrito así la descripción en términos generales, ahora se hará referencia a los dibujos adjuntos, que no están necesariamente dibujados a escala, y en los que:

La **Figura 1** ilustra esquemáticamente una formalización multinivel y multiestructura de datos de audio digital, como un ejemplo de estructuración de información de acuerdo con diversos aspectos de la presente descripción;

Las **Figuras 2A - 2C** ilustran esquemáticamente una visualización de primotrones de información de sonido informativos correspondientes a una formalización MSML de un archivo de audio digital, como un ejemplo de estructuración de información de acuerdo con diversos aspectos de la presente descripción;

La **Figura 3A** ilustra esquemáticamente un proceso de manipulación para la codificación sin fases de datos de audio digital, como un ejemplo de estructuración de información de acuerdo con diversos aspectos de la presente descripción;

5

La **Figura 3B** ilustra esquemáticamente un proceso de manipulación para la decodificación sin fases de datos de audio digital, como un ejemplo de estructuración de información de acuerdo con diversos aspectos de la presente descripción;

La **Figura 4A** ilustra esquemáticamente un proceso de manipulación para descomponer una composición musical en una parte residual y una parte repetitiva (es decir, extraer elementos repetitivos difusos), basados en 3 señales "base" iguales, como un ejemplo de estructuración de información de acuerdo con diversos aspectos de la presente descripción;

10

La **Figura 4B** ilustra esquemáticamente un proceso de manipulación para descomponer una composición musical para extraer elementos repetitivos difusos, como un ejemplo de estructuración de información de acuerdo con diversos aspectos de la presente descripción;

La **Figura 5** ilustra esquemáticamente un proceso de manipulación para la compresión de audio utilizando estructuras repetitivas de jerarquía, como un ejemplo de estructuración de información de acuerdo con diversos aspectos de la presente descripción;

15

La **Figura 6A** ilustra esquemáticamente un proceso de manipulación para la compresión sin pérdida utilizando un modelo contextual que implica un método de mezcla de probabilidad, como un ejemplo de estructuración de información de acuerdo con diversos aspectos de la presente descripción;

20

La **Figura 6B** ilustra esquemáticamente un proceso de manipulación para la compresión sin pérdida utilizando un método de estimación de probabilidad secundario, como un ejemplo de estructuración de información de acuerdo con diversos aspectos de la presente descripción;

La **Figura 7A** ilustra esquemáticamente un proceso de manipulación para una transmisión eficiente de datos de audio utilizando un esquema de compresión con pérdida que implica un acelerador de datos de audio y un proceso de pérdida de calidad nula (NQL) para reducir el número de coeficientes espectrales de tiempo-frecuencia cuantificados (QMDCT), como un ejemplo de estructuración de información de acuerdo con diversos aspectos de la presente descripción;

25

La **Figura 7B** ilustra esquemáticamente los fragmentos de coeficiente QMDCT que son resultado de un proceso de manipulación que implica un esquema de compresión con pérdida, como un ejemplo de estructuración de información de acuerdo con diversos aspectos de la presente descripción;

30

La **Figura 8A** ilustra esquemáticamente un esquema de codificación de un proceso de manipulación de codificación de transformadas afines, como un ejemplo de estructuración de información de acuerdo con diversos aspectos de la presente descripción;

La **Figura 8B** ilustra esquemáticamente un esquema de decodificación de un proceso de manipulación de codificación de transformadas afines, como un ejemplo de estructuración de información de acuerdo con diversos aspectos de la presente descripción;

35

La **Figura 8C** ilustra esquemáticamente un esquema de decodificación rápida de un proceso de manipulación de codificación de transformadas afines, como un ejemplo de estructuración de información de acuerdo con diversos aspectos de la presente descripción;

40

La **Figura 9** ilustra esquemáticamente cambios en los datos entre los niveles de una formalización MSML de un archivo de información digital y procesos particulares de codificación y decodificación asociados con el mismo, como un ejemplo de estructuración de información de acuerdo con varios aspectos de la presente descripción;

La **Figura 10** ilustra esquemáticamente cambios en los datos entre los niveles de una formalización MSML de un archivo de audio digital y procesos particulares de codificación y decodificación asociados con el mismo, como un ejemplo de estructuración de información de acuerdo con varios aspectos de la presente descripción; y

45

La **Figura 11** ilustra esquemáticamente los procesos aplicados a los datos entre los niveles de una formalización MSML de un archivo de información digital, como un ejemplo de estructuración de información de acuerdo con varios aspectos de la presente descripción.

50

DESCRIPCIÓN DETALLADA DE LA DESCRIPCIÓN

La presente descripción se describirá ahora más detalladamente en lo sucesivo en el presente documento con referencia a los dibujos adjuntos, en los que se muestran algunos, pero no todos, los aspectos de la descripción. De hecho, la descripción puede realizarse de muchas formas diferentes y no debe interpretarse como limitada a los aspectos expuestos en el presente documento; más bien, estos aspectos se proporcionan para que esta descripción satisfaga los requisitos legales aplicables. Los números similares se refieren a elementos similares en todo el documento.

55

- Varias teorías principales contemporáneas de ciertos fenómenos se basan en principios de optimalidad relativamente sencillos formulados en una sola frase. Sin embargo, llegar a dichas formulaciones sencillas generalmente requiere una evolución significativa y lenta con respecto a la teoría subyacente. Por ejemplo, la óptica geométrica se basa en el principio de un mínimo de movimiento a lo largo del haz de luz, la mecánica explota el principio de un mínimo de acción, la termodinámica se basa en el principio de un máximo de entropía, la teoría de la evolución se basa en el principio de supervivencia de las especies más adaptadas, y la biología implementa el principio de un mínimo de energía. Los aspectos de la presente descripción, concretamente, los aspectos dirigidos a una teoría de formalización de sonido (basada en una disposición de estructura múltiple/nivel múltiple (MSML)), como se describe, generalmente se basan y se dirigen al principio de maximización de la informatividad con respecto al tamaño del archivo de datos, en donde, en aspectos particulares, tal principio se extiende a la maximización de la informatividad o el valor de información de los elementos de datos junto con la minimización del tamaño del archivo de datos.
- 15 Como premisa inicial, puede darse el caso de que la cantidad de información incluida dentro de las ondas de sonido enviadas a los oídos humanos sea apreciablemente mayor que la que recibe o se perciba de otro modo por el cerebro humano. Como tal, determinar y cuantificar de manera sucinta el alcance de la información que recibe y puede ser percibida por el cerebro humano puede ayudar a facilitar un esquema de formalización de sonido que se corresponda más estrechamente con el mecanismo de percepción humana para el sonido/audio, que puede estar entre los más eficientes dentro de un sistema vivo. Por consiguiente, los aspectos de los esquemas de formalización MSML presentados en el presente documento pueden abordar los problemas mencionados anteriormente y pueden servir para crear una base de formalización universal para representar el sonido, tanto con respecto al habla como a la música (es decir, el audio).
- 25 Cuando se presenta con un evento de audio de espectro completo, puede darse el caso de que al menos una parte de la información de audio se rechace a nivel fisiológico de la estructura del oído humano. Es decir, la fisiología del oído humano puede tener limitaciones que no facilitan la recepción de toda la información de audio (es decir, limitaciones de frecuencia). Sin embargo, de la parte de la información de audio que recibe la fisiología del oído humano, la red de neuronas humanas asociada con el oído humano es responsable entonces de extraer y seleccionar naturalmente los objetos de sonido más significativos de la información de audio restante. En esos casos, los objetos de sonido extraídos y seleccionados pueden caracterizarse por datos específicos locales al evento de audio, así como, y posiblemente incluso principalmente, por estructuras de sonido más globales y generalizadas.
- 35 A este respecto, los aspectos de la presente descripción se dirigen a tales estructuras de sonido globales/generalizadas que pueden asociarse generalmente con la percepción humana del sonido/audio. En algunos casos, dichas estructuras de sonido pueden tener una premisa adicional o estar asociadas con, por ejemplo, tres niveles de pérdida de calidad nula (NQL), en donde estos niveles pueden depender, por ejemplo, de un momento de percepción indistinguible de la información recibida. Más particularmente, los tres niveles de NQL pueden caracterizarse de la siguiente manera:

1. Eventos de sonido indistinguibles perceptiblemente si un evento de sonido está, por ejemplo, separado en aproximadamente 20 ms de duración en el tiempo de otro evento de sonido.

- 45 Para tales eventos de sonido, la NQL puede satisfacerse solo si ambos eventos de sonido tienen formas de onda equivalentes. Los modelos de representación de sonido que cumplen con este criterio se utilizan generalmente en códecs de audio estándar o convencionales. Sin embargo, una formalización MSML basada en esta escala/duración de tiempo puede incluir un número sustancial o una proporción de estructuras o elementos MSML de baja información. Como tal, un códec asociado con una formalización MSML que se basa únicamente en tal criterio puede no ser capaz de maximizar o ni siquiera conservar la informatividad o calidad necesarias de las estructuras o elementos MSML, particularmente, por ejemplo, a bajas tasas de bits (o poca cantidad de información).

2. Eventos de sonido indistinguibles perceptiblemente si un evento de sonido está, por ejemplo, separado en aproximadamente 1 min de duración en el tiempo de otro evento de sonido.

- 55 En esta escala/duración de tiempo, el requisito de equivalencia matemática (es decir, formas de onda equivalentes) no es necesario, y las características de percepción de los eventos de sonido se definen por estructuras o elementos MSML comparativamente mucho menos esenciales, ya que una cantidad relativamente menor de información sobre el evento de sonido inicial escuchado/experimentado previamente permanece en o asociada con el cerebro. Sin

embargo, se puede realizar un análisis de sonido más complejo para determinar si las estructuras o elementos asociados con el segundo nivel o niveles superiores de formalización MSML pueden tener calidad suficiente, aunque la correspondencia exacta puede no ser necesaria para lograr la equivalencia de percepción.

5 3. Eventos de sonido perceptivamente equivalentes si un evento de sonido está, por ejemplo, separado en aproximadamente 1 día o más de duración en el tiempo de otro evento de sonido.

En este periodo/duración de tiempo, el cerebro humano generalmente retiene solo estructuras de información de sonido altamente concentradas, que generalmente representan/corresponden a niveles superiores del esquema de formalización MSML. Dichas estructuras desempeñan un papel importante en los aspectos del análisis de sonido presentados en el presente documento y son importantes para la equivalencia perceptiva, aunque la cantidad de información/informatividad de las estructuras/elementos es relativamente menor que en el periodo/duración de tiempo anterior. Generalmente, los profesionales de audio cualificados pueden distinguir la estructura de sonido de nivel inferior (es decir, asociados con los periodos de tiempo/duraciones más cortas anteriores, incluso en largos periodos de tiempo entre eventos de sonido, aunque la transición/progresión de estructuras de sonido de nivel inferior a superior (es decir, desde el periodo de tiempo/duración más corta hasta el periodo de tiempo/duración más larga) es generalmente el proceso común para la mayoría de las personas que experimentan los eventos de sonido.

De acuerdo con la teoría física, en un sistema que tiene varios estados permisibles en ciertas condiciones, cuanto más ordenado está el sistema, menos entropía está presente. Además, a partir de la segunda ley de la termodinámica, la entropía generalmente aumenta en sistemas aislados. Sin embargo, de acuerdo con Schroedinger, una persona se esfuerza continuamente por reducir la entropía en las experiencias de la vida con el fin de apoyar la organización interna, o de otro modo para guardar información estructurada. El concepto de entropía también puede equipararse a veces con la cantidad de información, así como con la propia información, al menos en parte debido a la falta de una determinación científica de la información y numerosas definiciones de la información. De acuerdo con Shannon, la cantidad de información es inversamente proporcional a la entropía. Es decir, cuanto menos probable es el evento (más entropía), menos información se incluye en los datos sobre ese evento. De otro modo, la cantidad de información puede representarse como un cambio cuantitativo en la incertidumbre como resultado de la comunicación. Sin embargo, la cantidad de información es solo un atributo de dicha información (que puede tener muchos atributos diferentes), y el valor de la información puede ser el más importante de esos otros atributos. En cualquier caso, a partir de muchas definiciones de información, los aspectos de la presente descripción pueden basarse en una dada por Casteler: "la información es una elección memorizada de una serie de posibles elecciones".

De acuerdo con Chernavsky, la macroinformación es una elección memorizada, concretamente, la duración en el tiempo de la existencia del conocimiento que es mayor que la duración en el tiempo del uso de la información dada asociada con ese conocimiento. La microinformación es una elección no memorizada de un aspecto de una serie de aspectos posibles. El alcance de la entropía física (o la entropía de Boltzman) se limita a los sistemas ergódicos que olvidan rápidamente su estado anterior (por lo tanto, en gas este tiempo es de aproximadamente 10^{-13} s a una temperatura de ~ 300 K). A este respecto, la cantidad de macroinformación es proporcional al logaritmo de una cantidad de estados estacionarios, y la microinformación es proporcional al logaritmo de una cantidad de todos los estados, no necesariamente estados estacionarios. Por ejemplo, la cantidad de microinformación en un recipiente lleno de gas es:

$I_{\text{micro}} = \log_2(\text{el número de todos los estados}) = \log_2(\exp(S/k))$, donde S es la entropía igual a $3/2kN$ (es decir, la información es $2,16 \cdot N$, que es comparable a la constante de Avogadro en el orden de 10^{23})
 $I_{\text{macro}} = \log_2(\text{el número de estados estacionarios}) = \log_2(1) = 0$ (ya que solo hay un estado estacionario, es decir, el equilibrio termodinámico).

Por lo tanto, el concepto de contenido de información depende de si esa información se puede o no usar. Más particularmente, el ruido blanco en un sonido puede considerarse el equivalente del equilibrio termodinámico, ya que ambos estados corresponden al de máxima entropía del sistema respectivo. Se puede extraer un corolario con respecto a un recipiente lleno de gas, concretamente, donde la macroinformación es mínima e igual a cero, mientras que la microinformación es máxima.

De acuerdo con un aspecto, el valor de la información se puede definir como la diferencia entre la entropía del sistema con y sin el efecto de la información recibida. Matemáticamente, dicho valor de información se puede definir como: $V = \log_2(P_1/P_0)$, donde P_0 es la probabilidad de cumplir el propósito sin la recepción de información, y P_1 es la probabilidad de cumplir el propósito con la información recibida, aunque la noción de propósito puede ser diferente

para el sistema de captura. En este sentido, el propósito de cada elemento es almacenar (conservar) la información asociada con el mismo. Una vez que se indique o se defina de otro modo el propósito, se puede hacer un seguimiento y analizar la evolución del valor de la información obtenida asociada con ese propósito. Por consiguiente, los aspectos de la presente descripción incluyen el propósito de analizar y extraer primotrones de audio digitalizado como base para el esquema de síntesis de sonido. Un experto en la técnica apreciará, sin embargo, que la referencia en el presente documento a sonido o audio es simplemente una ilustración ejemplar de la aplicabilidad y capacidades de los sistemas y métodos descritos en el presente documento, y que los sistemas y métodos descritos en el presente documento pueden ser más generalmente aplicables a cualquier información capaz de analizarse en forma digital tal como, por ejemplo, medios digitales que incluyen audio, vídeo y similares.

10

Por lo tanto, los aspectos de la descripción pueden proporcionar, por ejemplo, sistemas y métodos para el análisis, identificación y estructuración de primotrones (es decir, elementos o estructuras de datos representativos primarios) organizados entre diferentes niveles de la jerarquía MSML, y asociados con los archivos de audio digital u otro archivo multimedia digital convertible a un formato MSML. Además, dichos sistemas y métodos de análisis, identificación y estructuración pueden ser diferentes de la síntesis o reconstrucción del audio u otros medios digitales dentro de un umbral de los medios digitales originales, a partir de dichos primotrones. Es decir, la codificación del archivo multimedia digital original puede ser un proceso diferente a la decodificación del archivo codificado para sintetizar o reconstruir el archivo multimedia digital original. En cualquier caso, el enfoque general de aspectos particulares de la presente descripción dirigida a la descripción o representación del sonido u otro medio basándose en primotrones puede permitir la representación (es decir, la determinación de las combinaciones de elementos de datos que forman los primotrones) y la síntesis ("reconstrucción") de los primotrones para formar un sonido u otro archivo multimedia puede basarse en las interacciones de las interfaces (también referidas en el presente documento como "procesos de manipulación") con la formalización MSML, en donde dichos procesos de manipulación pueden ser independientemente funcionales o cooperativos para una función particular o en casos particulares. Dado que la naturaleza del sonido (especialmente la música y el habla), pero también en cierta medida otros medios tal como vídeo, puede basarse en la repetitividad particular del sonido básico u otros elementos básicos (es decir, letras individuales, ortografía de palabras, símbolos de melodía, armónicos), entonces los primotrones particulares o combinaciones de los mismos, se pueden usar como una representación compacta del archivo procesado por MSML en términos de disposición de estos elementos básicos o primarios identificados de una manera altamente estructurada (en donde los datos estructurados pueden tener asociados con los mismos parámetros caóticos particulares que pueden facilitar la interpretación de los datos, así como la estructura de los mismos) que reduce la cantidad de información, por ejemplo, reduciendo o minimizando la redundancia de estos elementos, garantizando al mismo tiempo que estos elementos básicos estructurados conservan dentro de un umbral perceptivo tanto o más del valor de la información que el archivo original. Además, de esta manera, diferentes niveles de primotrones o elementos de datos en la jerarquía MSML pueden servir como semilla o semillas a partir de las cuales la señal general (es decir, el archivo de audio digitalizado original) se puede sintetizar o restaurar de otra manera, si las instrucciones o reglas correspondientes para dicha síntesis se definen, por ejemplo, a través de los parámetros caóticos o de otra manera adecuada (es decir, que implica una evaluación estadística de las probabilidades de reconstrucción).

40

Como tal, en la revisión, los aspectos de la presente descripción proporcionan métodos de procesamiento de audio u otros medios digitales dirigidos a formar una formalización digital compacta multiestructural, multinivel (MSML) de tal sonido u otros medios (véanse, por ejemplo, las Figuras 1, 9, 10 y 11), donde dichos métodos se basan en definiciones particulares de un primotrón de información, criterios de información perceptiva (PIC) asociados con la existencia de dichos primotrones (es decir, asociados con los límites de la percepción humana), y la determinación de combinaciones o conjuntos de elementos de datos que forman dichos primotrones. Dicha formalización MSML también puede involucrar métodos de extracción y clasificación de primotrones de audio digital u otros archivos multimedia digitales en los dominios tanto del tiempo como de tiempo-frecuencia. También asociada con los procedimientos de formalización MSML como se describe en el presente documento puede estar la definición de un nivel secundario o conjunto de estructuras de sonido MSML y los PIC asociados con las mismas, que también puede implicar al menos algún análisis de primotrón para construir tales estructuras MSML secundarias. Los procedimientos de formalización MSML también pueden asociarse con la definición de un alto nivel o conjunto de estructuras de sonido MSML y los PIC asociados con las mismas en la jerarquía, lo que también puede implicar el análisis de estructuras MSML de nivel inferior para seleccionar o definir de otro modo estructuras MSML de nivel superior particulares. Estos principios y métodos cooperan para permitir la construcción de un modelo estructural paramétrico de sonido u otro medio con el propósito de formalizarlo, particularmente la formalización MSML que se describe en el presente documento, que puede promover y facilitar la organización de los datos en estructuras y conjuntos de estructuras particulares. También se pueden asociar con los procedimientos de formalización MSML los métodos y aparatos para establecer una biblioteca de primotrones de dominio del tiempo-frecuencia que se

55

puede usar para la síntesis o reconstrucción de sonido, donde la disposición o disposiciones de primotrones particulares dentro de la biblioteca pueden definirse por las estructuras de niveles superiores de la formalización MSML.

- 5 Los aspectos de la presente descripción también proporcionan métodos para determinar dichos criterios de información perceptiva (PIC) para definir las probabilidades de existencia del primotrón realizando una comparación de dos valores de energía, en donde se puede determinar un valor de energía para una estructura forma o está asociada de otro modo con un primotrón, y en donde el segundo valor de energía puede calcularse para una fracción de una señal digital, dentro del dominio del tiempo-frecuencia, que incluye el primotrón señalado. Si la relación entre el primer valor de energía con respecto al segundo valor de energía es mayor que un valor de umbral predeterminado, entonces se cumplen los PIC para el primotrón señalado, afirmando la existencia de los mismos.

- 15 Otro aspecto de la presente descripción está dirigido a un método para definir un valor de energía de umbral como una probabilidad aceptable de un evento en el que la fluctuación aleatoria del caos se puede reconocer como una estructura de sonido valiosa, en donde tal probabilidad se puede determinar realizando un experimento computacional o mediante la aplicación directa de una fórmula derivada teóricamente para definir una función de distribución de probabilidad utilizando ruido aleatorio como señal de entrada.

- 20 Otro aspecto más de la presente descripción se dirige a una formulación teórica de una definición de valor de umbral de energía de primotrón a través de un proceso que incluye el cálculo del área superficial de la esfera n-dimensional (S_n) de un radio de unidad única en un dominio del tiempo-frecuencia, utilizando la siguiente función implícita:

$$P(N, M, k) = \frac{S_{M-1} S_{N-M-1}}{S_{N-1}} \sum_{l=0}^{N-M-1} \sum_{q=0}^{M-1} \left| \frac{C_{N-M-1}^l C_{M-1}^q}{2^{N-2} i^M (2l+2q-N+2)} \right| (-1)^{M-1-q} e^{i(2l+2q-N+2) \arcsin(\sqrt{k})}$$

donde K es el umbral de energía objetivo de los primeros M componentes de un vector N dimensional dividido por la energía total del vector. Los índices en los términos de suma incluyen los valores $2l+2q-N+2 = 0$ y

$$\frac{e^{i(2l+2q-N+2) \arcsin(\sqrt{k})} - 1}{i(2l+2q-N+2)} \otimes \arcsin(\sqrt{k}),$$

- 25 de manera que:
 S_n se determina como:

$$S_{2l} = \frac{(4\pi)^l (l-1)!}{(2l-1)!}, S_{2l+1} = \frac{2\pi^{l+1}}{l!}$$

- 30 Otro aspecto más de la presente descripción está dirigido a un método para definir un criterio de pérdida de calidad nula (NQL) mediante análisis comparativo formalización MSML descritos en el presente documento.

- 35 Otro aspecto más de la presente descripción está dirigido a un método para definir un PIC como un factor de probabilidad para construir una nueva estructura de nivel superior de la jerarquía de formalización MSML, solo cuando el establecimiento o el origen de tal estructura de nivel superior disminuye la entropía general de las estructuras de nivel inferior de la formalización MSML que cooperan para formar la estructura de nivel superior.

- 40 Un aspecto adicional de la presente descripción está dirigido a un método para determinar o designar una estructura de primotrón como un número de primotrones correspondientes u otros elementos de datos dentro del primer nivel de la jerarquía MSML de acuerdo con la coordenada de tiempo asociada con la misma dentro del dominio del tiempo. Además, tal aspecto también puede incluir la definición de un acierto como primotrones de segundo nivel localizados en el tiempo, estructuras de primotrón, u otros elementos de datos que tengan las mismas coordenadas de inicio en el tiempo, en donde se proporciona una disminución de la entropía en una estructura de nivel superior (es decir, una descripción más concisa) al eliminar la redundancia de los valores de componente de tiempo de primotrón aparentes en la estructura de nivel inferior. Un aspecto de este tipo también puede incluir la definición de armónicos como estructuras MSML de segundo nivel localizadas en la frecuencia, en donde un inicio o encuentro de primotrón posterior coincide con la coordenada final de primotrón anterior. En tales casos, la disminución de la entropía para una estructura de nivel superior también se proporciona al eliminar la redundancia de los valores de los componentes de tiempo de primotrón. Tal aspecto puede incluir además la definición de un conjunto de sobretonos como estructuras MSML de segundo nivel localizadas en la frecuencia que incluyen una frecuencia base

única de las estructuras MSML de primer nivel y una cantidad de copias de las mismas desplazadas por un valor de frecuencia fija desde la frecuencia base y entre sí. La entropía general de tales estructuras puede tender a disminuir como resultado de una descripción compacta del conjunto de estructuras de primer nivel en un solo parámetro asociado con el cambio de frecuencia.

5

Más particularmente, cuando el archivo de información digital original es un archivo de audio digital, tal aspecto de la presente descripción implica asociar una cantidad de información del archivo de audio digital con una tasa de bits proporcional a un tamaño y una duración del archivo de audio digital, y asociar un valor de información del archivo de audio digital con una calidad del archivo de audio digital, en el que la calidad se determina a partir de al menos uno de un valor de umbral de energía de primotróon, un criterio de información perceptiva (PIC), una estructura de sonido identificada a partir de un evento de fluctuación de caos aleatorio determinado de acuerdo con una función de distribución de probabilidad que usa ruido aleatorio como señal de entrada, un criterio de pérdida de calidad nula, y un área de superficie de una esfera n-dimensional de un radio de unidad única en un dominio de tiempo-frecuencia determinado a partir de función implícita.

10

En aspectos particulares, por ejemplo, en el caso de un archivo de audio digital como se muestra, por ejemplo, en las Figuras 2A-2C, las interfaces estructurantes o los procesos de manipulación pueden aplicarse sucesivamente a un archivo de audio digital previamente resultante para formar un último archivo de audio digital estructurado resultante configurado de acuerdo con una jerarquía de información que tiene una pluralidad de niveles de información, en donde un nivel de información inferior en la jerarquía comprende una representación del audio como un archivo de información digital de alta resolución, y un nivel de información superior en la jerarquía comprende, por ejemplo, una representación en ficha que incluye una puntuación de nota escrita con las letras correspondientes. Cuando está tan estructurado, el nivel de información superior puede tener una cantidad de información más pequeña que el nivel de información inferior en la jerarquía, en donde el nivel de información inferior se puede formar, sintetizar o reconstruir a partir del nivel de información superior y los parámetros caóticos asociados con el rendimiento y la digitalización del audio de acuerdo con la representación en ficha.

Otro aspecto de la presente descripción reconoce que la aplicación sucesiva de procesos de manipulación al archivo de audio digital previamente resultante forma el último archivo de audio digital estructurado resultante, en donde el último archivo de audio digital estructurado resultante comprende al menos un primotróon definido en uno de un dominio del tiempo y un dominio del tiempo-frecuencia, y por lo tanto, representa una formalización digital compacta de una señal armónica, en donde la cantidad de información asociada con la tasa de bits de audio se reduce en comparación con el archivo de audio digital correspondiente al archivo de información digital original, mientras que el valor de la información asociado con la calidad de audio se mantiene dentro de un umbral del archivo de audio digital correspondiente al archivo de información digital original.

Aún otro aspecto de la presente descripción reconoce que cada archivo de audio digital representa datos relacionados con armónicos, y el análisis de los mismos implica la definición de aciertos como estructuras jerárquicas de nivel superior en el dominio del tiempo; la definición de armónicos como estructuras jerárquicas de nivel superior en el dominio de la frecuencia que tienen un origen de primotróon armónico posterior que coincide con una terminación de primotróon armónico anterior; la eliminación de la redundancia en las estructuras jerárquicas de nivel inferior de dominio del tiempo para obtener una disminución de la entropía en las estructuras jerárquicas de nivel superior; y la definición de sobretonos como estructuras jerárquicas de nivel superior de dominio de la frecuencia que tienen una frecuencia base única correspondiente a las estructuras jerárquicas de nivel inferior y una pluralidad de copias de las estructuras jerárquicas de nivel superior desplazadas una frecuencia fija de la frecuencia base única y entre sí.

En un sentido más general, la formalización MSML puede referirse a un método de estructuración de información, en el que dicha información puede comprender, por ejemplo, información digital tal como datos de audio, como se ha hecho referencia anteriormente y de otro modo en el presente documento. De este modo, primero se analiza un archivo de información digital original para determinar una cantidad de información original y un valor de información original asociado con la misma. Es decir, primero se determina cuántos elementos de datos se incluyen en el archivo de información digital, y el nivel asociado de informatividad asociado con cada uno de estos elementos de datos. En la estructuración de la información, se puede aplicar un proceso de manipulación inicial al archivo de información digital original para formar un primer archivo de información digital resultante, y aplicarse un proceso de manipulación posterior al primer archivo de información digital resultante para formar un segundo archivo de información digital resultante. En tales casos, cada proceso de manipulación puede configurarse para eliminar al menos un elemento del archivo de información digital procesado, en el que tal elemento eliminado puede, por ejemplo, identificarse como redundante, con ruido o, de otro modo, innecesario. En otros casos, cada proceso de

manipulación puede estar configurado para representar una combinación de una pluralidad de elementos del archivo de información digital procesado con un elemento representativo y una primera indicación asociada con una interrelación entre el elemento representante y al menos uno de la pluralidad de elementos en la combinación, a fin de reducir la cantidad de información y estructurar el archivo de información digital procesado. Es decir, los elementos pueden combinarse y representarse por un único elemento representativo, en donde el único elemento representativo puede tener asociado con el mismo una primera indicación que indique, por ejemplo, una razón, un resultado de análisis, o cualquier otro indicador adecuado de por qué los elementos particulares se combinaron y la naturaleza de la relación entre esa combinación particular de elementos y el elemento representativo designado (es decir, el elemento representativo representa que los elementos combinados forman un patrón o melodía repetida, definen un armónico, están asociados con la letra, etc.). Al eliminar uno o más elementos y/o combinar elementos, cada proceso de manipulación puede configurarse de tal manera que al menos uno de los elementos eliminados y/o el elemento representativo se determine para reducir el valor de información del archivo de información digital procesado en no más de un umbral seleccionado. Es decir, mientras que la cantidad de información previa puede reducirse, cada proceso de manipulación se configura preferiblemente de modo que el valor de información del archivo de información digital resultante esté dentro de una cantidad particular u otra medida adecuada del archivo de información digital original o anterior (es decir, dentro de un umbral particular). De esta manera, los procesos de manipulación pueden aplicarse sucesivamente al archivo de información digital previamente resultante (es decir, para cada archivo de información digital resultante, pueden aplicarse al mismo otro proceso de manipulación), hasta que las aplicaciones de proceso de manipulación sucesivas no logran una reducción de umbral en la cantidad de información en el archivo de información digital resultante posterior sobre el archivo de información digital previamente resultante. Es decir, el proceso de aplicar un proceso de manipulación a cada archivo de información digital resultante puede continuar hasta que no haya una disminución apreciable (es decir, dentro de un umbral) en la cantidad de información en un archivo de información digital resultante posterior sobre el archivo de información digital resultante anterior. En otras palabras, la aplicación del último proceso de manipulación aplicada no da como resultado una eliminación apreciable del elemento o elementos y/o una combinación apreciable de elementos que están representados por un elemento representativo. Como tal, el último archivo de información digital resultante puede tener de este modo una estructura primaria con una cantidad de información reducida con respecto a la cantidad de información original y un valor de información resultante dentro del umbral seleccionado del valor de información original. Es decir, la cantidad de información incluida en el último archivo de información digital resultante es menor que la cantidad de información en el archivo de información digital original, pero el valor de información o informatividad de los datos en el último archivo de información digital resultante está dentro de una cierta medida aceptable (es decir, dentro de un umbral) del archivo de información digital original (véanse, por ejemplo, las Figuras 9-11).

35 Durante el proceso de aplicar sucesivamente diversos procesos de manipulación a cada archivo de información digital resultante, al menos el último archivo de información digital estructurado resultante puede configurarse de acuerdo con una jerarquía de información que tiene una pluralidad de niveles de información. Es decir, el último archivo de información digital resultante incluye información relativamente altamente estructurada (en comparación, por ejemplo, con el archivo de información digital original), en el que la cantidad de datos o información dentro del archivo es menor que el archivo de información digital original, pero el valor de esos datos o información se encuentra dentro de una cantidad o umbral particular del archivo de información digital original (es decir, un archivo más pequeño en términos de cantidad de datos, pero en el que el archivo tiene aproximadamente el mismo valor de información, o posiblemente más). Además, al estar estructurado de este modo, el último archivo de información digital resultante puede tener una configuración jerárquica, en la que los datos en los niveles de información superiores de la jerarquía pueden estar relacionados con los datos en los niveles inferiores de la jerarquía por medio de diversos parámetros caóticos. Es decir, un nivel de información superior de la jerarquía de información puede incluir una cantidad de información más pequeña que un nivel de información inferior, en donde el nivel de información superior se forma a partir de una estructura (es decir, combinaciones de datos o combinaciones de elementos particulares) del nivel de información inferior y parámetros caóticos (es decir, razones u otros parámetros que conducen a la combinación particular) asociados con el mismo. En algunos casos, el nivel de información superior puede formarse de tal manera que el nivel de información superior incluya un valor de información más elevado que el valor de información original, por ejemplo, debido a la información asociada con los parámetros caóticos. En algunos casos, si es necesario, se puede formar una pluralidad de niveles de información inferiores a partir de una estructura de un nivel de información superior, y al menos uno de la pluralidad de niveles de información inferiores puede seleccionarse para asociarse con ese nivel de información superior. En tales casos, la relación entre los mismos puede definirse por parámetros caóticos particulares.

En algunos aspectos, una pluralidad de procesos de manipulación puede aplicarse simultáneamente a un archivo de información digital anterior para formar un archivo de información digital estructurado secundario, además del

archivo de información digital estructurado primario. En tales aspectos, los archivos de información digital estructurados primarios y secundarios se pueden analizar para determinar al menos una de las interrelaciones entre los niveles de información dentro de uno de los archivos de información digital estructurados primarios y secundarios, y una interrelación entre los archivos de información digital estructurados primarios y secundarios.

5 Dichas interrelaciones pueden, en algunos casos, incluir o estar asociadas de otro modo con los parámetros caóticos descritos previamente.

En algunos aspectos, al menos uno de una pluralidad de procesos de manipulación pueden aplicarse a un archivo de información digital anterior, en el que los procesos de manipulación pueden seleccionarse del grupo que consiste
10 en, por ejemplo, un proceso de manipulación de codificación sin fases, un proceso de manipulación de análisis estructurado para estructuras repetitivas, un proceso de manipulación de codificación de repetición sin fases, un método de manipulación de compresión de entropía sin pérdidas, un proceso de manipulación de acelerador sin pérdidas para formar una capa de compresión adicional para un archivo de datos de audio previamente reducido, un proceso de manipulación de acelerador para facilitar la transmisión eficiente de datos de audio, y un proceso de
15 manipulación de codificación de transformada afín.

Más particularmente, en un aspecto de un proceso de manipulación, la reducción del ancho de banda y de la tasa de bits de los datos de audio puede realizarse utilizando una interfaz de codificación sin fases para una formalización de composición de sonido MSML, en donde la interfaz está configurada para considerar una jerarquía de valores de
20 información perceptiva (véase, por ejemplo, la Figura 3A). La reducción del ancho de banda y de la tasa de bits se pueden lograr utilizando una interfaz de este tipo para realizar tareas dirigidas a la reducción de datos seleccionados, al tiempo que se mantiene la informatividad de todas las estructuras de sonido asociadas con la composición del sonido, por ejemplo, mediante el uso de métricas de sonido. De este modo, se realiza una comparación de objetos de sonido de diferente nivel en el espacio, centrándose en determinar la integridad de las
25 estructuras y el grado de daño en los mismos, dentro de una representación reducida de la señal después de la filtración, lo que facilita la selección de un parámetro adecuado para cada elemento de sonido. La interfaz de codificación sin fases está configurada de manera apropiada de tal forma que la dependencia del tiempo de la señal no se conserve, y las estructuras de sonido se conserven en lugar de la forma de onda de la señal, dando como resultado de este modo una disminución del ancho de banda esencial y la tasa de bits. Para lograr una mayor
30 velocidad de ancho de banda de audio y reducción de la tasa de bits, las estructuras de nivel relativamente más informativas de la formalización de composición de sonido MSML están configuradas para mantener parámetros constantes. Los parámetros de las estructuras de nivel relativamente menos informativas dentro de la jerarquía se pueden guardar con menos precisión o se generan durante la decodificación con el uso de valores aleatorios, en donde se muestra un proceso de decodificación ejemplar asociado con tal interfaz de codificación sin fases, por
35 ejemplo, en la Figura 3B.

De acuerdo con otro aspecto de un proceso de manipulación, puede realizarse un análisis de la estructura de datos de audio, utilizando una interfaz de análisis de estructura para una formalización de composición de sonido MSML, para la extracción de elementos repetitivos difusos de dichas composiciones musicales como se muestra, por
40 ejemplo, en las Figuras 4A y 4B. Tal interfaz de análisis de estructura puede configurarse para procesar datos de audio recibidos de la formalización MSML de la señal original e identificar elementos de bucle repetitivos difusos dentro de un nivel relativamente bajo de la formalización MSML, para su uso en una mayor reducción de datos y sin disminuir la informatividad general de la señal. Tal interfaz puede comprender tres módulos, en la que el primer módulo puede proporcionar la búsqueda, el posicionamiento y el marcado de longitud de todos los elementos
45 repetitivos difusos a lo largo de todo el nivel inicial de la formalización MSML. El segundo módulo puede configurarse para extraer un componente de sonido general para cualquier conjunto de fragmentos repetitivos difusos y para definir el nivel de cohesividad de los datos formando una señal residual correspondiente. El tercer módulo puede configurarse para realizar una construcción predictiva del fragmento de audio resultante integrando los elementos repetitivos extraídos, los datos de marcado correspondientes y la información residual. Puede aplicarse un método
50 de filtrado para el modelado de fases para minimizar la información de la fase inicial requerida durante el proceso de construcción de sonido. Tal interfaz puede configurarse para ser utilizada como un elemento para una diversidad de aplicaciones, incluyendo, pero sin limitación, por ejemplo, motores de búsqueda de música, visualización de música, y tecnologías de compresión profunda de audio.

55 Otro aspecto de un proceso de manipulación implica la compresión de audio utilizando estructuras repetitivas de jerarquía en una formalización de composición de sonido MSML como se muestra, por ejemplo, en la Figura 5. Tal interfaz de codificación de repetición sin fases avanzada utiliza estructuras formadas por primotrones de sonido de macro-repetición del nivel superior dentro de una jerarquía de sonido de formalización MSML significativa. Se pueden utilizar dos tipos de datos para el proceso de decodificación: datos a gran escala que definen la estructura

general dentro de la formalización MSML y datos que comprenden fragmentos de señales originales descritos por un modelo de nivel relativamente inferior en la jerarquía. Para comprimir los fragmentos de señal, se puede utilizar un codificador psicoacústico general, en algunos casos, posiblemente utilizando un enmascaramiento de señal externo. La reducción de la velocidad de bits se produce debido a la eliminación de fragmentos repetitivos construidos en el nivel superior de formalización MSML, mientras se almacenan las ubicaciones de estos fragmentos. La equivalencia perceptiva de fragmentos repetitivos permite una reducción en el número de muestras de referencia.

Otro aspecto más de un proceso de manipulación implica la reducción del ancho de banda de audio y de la tasa de bits utilizando una interfaz de compresión sin pérdida para una formalización de composición de sonido MSML (véase, por ejemplo, la Figura 6A) que, en algunos casos, reduce el número de matrices de valores enteros asociado con la formalización de composición de sonido MSML a través del uso de un codificador de entropía avanzado que realiza solo operaciones de enteros, sin divisiones y estimaciones de probabilidad secundarias. Dado que el codificador de rango y el modelo contextual de la interfaz no utilizan la operación matemática de división, permite la aplicación de tal interfaz a la ejecución de la interfaz mediante la mayor parte de microordenadores de bajo coste que no tienen una operación de división. Para mejorar la fiabilidad del modelo contextual de la interfaz, puede emplearse una estimación de probabilidad secundaria (véase, por ejemplo, la Figura 6B). La estimación de probabilidad secundaria es un modelo contextual complejo que utiliza la probabilidad predicha como un contexto con respecto a otro modelo. La estimación de probabilidad secundaria bidimensional también se puede utilizar para mezclar dos modelos diferentes. La estimación de probabilidad secundaria bidimensional es una modificación de la estimación de probabilidad secundaria, que tiene dos probabilidades de entrada y las utiliza como contexto. También se pueden utilizar contextos enteros adicionales. Tal aspecto de un proceso de manipulación también puede implicar la compresión de señales de audio, y/o mejorar la codificación aritmética y el modelado contextual para aumentar las relaciones de compresión y las velocidades de procesamiento, y así permitir que los datos espectrales se compriman con alta eficacia y mayor velocidad. El procedimiento de compresión puede ser completamente automatizado y no requiere necesariamente inicialización previa para diferentes tipos de datos de audio u otros datos digitales. Puede ser lo suficientemente flexible para ajustarse a diferentes tamaños o cantidades de datos de audio espectral u otros datos digitales, lo que permite su uso con diferentes transformaciones espectrales. En lugar de un codificador aritmético estándar, puede utilizar un codificador de rango más eficiente. El modelado de contexto se aplica al flujo de datos, los modelos algorítmicos construidos y la optimización algorítmica de una función de decodificador. Este aspecto también puede basarse, al menos parcialmente, en el uso de técnicas de codificador de rango adaptativo que implican aumentar la probabilidad del valor codificado. Para mejorar la fiabilidad del modelo contextual, se puede emplear una estimación de probabilidad secundaria.

Otros aspectos de procesos de manipulación implican la transmisión eficiente de datos de audio utilizando una interfaz de acelerador avanzado de datos de audio para una formalización de composición de sonido MSML, utilizando un proceso de pérdida de calidad nula (NQL) para reducir el número de coeficientes espectrales de tiempo-frecuencia cuantificados (QMDCT), para proporcionar una entrega más rápida del audio original u otra información digital sin una degradación esencial de la calidad en el sonido o datos resultantes (véanse, por ejemplo, las Figuras 7A y 7B). El proceso NQL realiza una clasificación de las estructuras de sonido dentro de la formalización MSML dependiendo de su importancia en función de la percepción humana. La avanzada interfaz del acelerador reconstituye los coeficientes espectrales cuantificados de tiempo-frecuencia del formato inicial, sin la descuantificación de los datos ni la transcodificación inversa a través del formato PCM. Tal interfaz avanzada del acelerador divide el archivo codificado inicialmente de coeficientes espectrales de tiempo-frecuencia en tres grupos, donde los coeficientes de un grupo espectral de baja frecuencia permanecen inalterados, mientras que algunos coeficientes de un grupo de frecuencia intermedia se ponen a cero utilizando el proceso NQL, y los coeficientes del tercer grupo (frecuencias más altas) se eliminan reemplazándolos con los subgrupos de coeficientes más similares, uniéndolos en grupos bajos e intermedios, y ahorrando una referencia a este subgrupo similar, así como, por ejemplo, un factor integral asociado con el mismo, sin afectar a los elementos de sonido estructurales u otros elementos de datos digitales (es decir, armónicos y aciertos).

Otro aspecto más de un proceso de manipulación implica reducir el tamaño o la cantidad del archivo de audio u otros datos digitales, utilizando una interfaz de codificación de transformadas afines para una formalización de composición de sonido MSML (véase, por ejemplo, la Figura 8A), mientras que (en el caso de audio) se conserva la integridad total de la reproducción del sonido original en forma de una transformada afín de los primotrones de sonido similares, lo que conduce a una disminución en la entropía de la señal. La interfaz afín asociada con la formalización MSML puede usarse como una adición o complemento de cualquier procedimiento de codificación de audio u otros datos digitales para aumentar la relación de compresión, así como el proceso de compresión del núcleo, para la clase de señales donde los microfragmentos similares afines de tiempo-frecuencia contienen una parte esencial de la información de sonido (u otros datos digitales). Esta interfaz de codificación de transformadas

afines, mientras que se utiliza junto con la formalización MSML, está configurada para utilizar microfragmentos espectrales similares de la señal digital (audio u otra) para aumentar la relación de compresión de la señal. Durante el procesamiento, el rango espectral completo se divide en sub-bandas. Se realiza una búsqueda de fragmentos espectrales similares en la parte anterior de la señal de sonido u otra señal de datos digital de forma independiente y utilizando una etapa diferente en cada sub-banda. Se usa una transformada afín de cambio de amplitud, tiempo y espectral durante la búsqueda de fragmentos similares. Cada fragmento encontrado se resta del fragmento espectral original y el residuo se procesa adicionalmente como una diferencia de componentes espectrales, eliminando de este modo una redundancia significativa de los datos espectrales. Un esquema de decodificación asociado para tal interfaz de codificación de transformadas afines se muestra, por ejemplo, en la Figura 8B, con un esquema de decodificación rápida opcional para tal interfaz de codificación de transformadas afines que se muestra, por ejemplo, en la Figura 8C.

Un experto en la técnica apreciará que los diversos aspectos de los procesos de manipulación descritos en el presente documento tienen únicamente fines ejemplares y no deben considerarse de ninguna manera como limitantes con respecto a los diversos procesos de manipulación que pueden aplicarse, ya sea en el presente o en el futuro, que puedan ser aplicable a la estructuración de datos de información digital para reducir el tamaño del archivo de datos digitales (es decir, reducir esa cantidad de datos reales o información que comprende el archivo de datos digitales), mientras se conserva un valor de información del archivo de datos digitales estructurado que se encuentra dentro de un valor de umbral del valor de información del archivo de datos digitales original. Un experto en la técnica apreciará además que el valor de información del archivo de datos digitales estructurado puede, en algunos casos, tener una cantidad mucho menor de datos o información en comparación con el archivo de información digital original, pero también puede tener una informatividad (es decir, valor de información) dentro de un valor de umbral del archivo de información digital original y, en casos particulares, incluso puede tener una mayor informatividad que el archivo de datos digitales original. Dicha estructuración de los datos de acuerdo con los principios en el presente documento y de acuerdo con diversos aspectos de la presente descripción se ilustran y se detallan con más detalle, por ejemplo, en las Figuras 9-11.

En aspectos particulares, la aplicación sucesiva de procesos de manipulación para obtener un valor de información resultante dentro del umbral seleccionado del valor de información original puede comprender además la aplicación sucesiva de dichos procesos de manipulación, siendo el umbral seleccionado, por ejemplo, un límite de la percepción humana (es decir, como se ha analizado anteriormente en relación con los criterios de información perceptiva o, en el caso de los datos de audio, por ejemplo, los límites espectrales de la audición humana), a fin de conservar la fidelidad del archivo de información digital estructurado primario con respecto al archivo de información digital original.

En el proceso de aplicar sucesivamente los procesos de manipulación para formar los archivos de información digital resultantes, los archivos de información digital resultantes se forman de tal manera que cada archivo de información digital resultante comprende un conjunto de objetos y agrupaciones de objetos, en donde cada objeto y agrupación de objetos tiene una definición respectiva. Más particularmente, puede aplicarse al menos un proceso de manipulación para eliminar al menos un elemento del archivo de información digital procesado, y/o representar una combinación de una pluralidad de elementos del archivo de información digital procesada con un elemento representativo, una primera indicación asociada con una interrelación entre el elemento representativo y al menos uno de la pluralidad de elementos en la combinación, y una segunda indicación asociada con una interrelación entre al menos dos de la pluralidad de elementos en la combinación. Por lo tanto, los archivos de información digital resultantes se forman de tal manera que los objetos y las agrupaciones de objetos en los mismos comprenden al menos uno de los elementos representativos y de tal manera que la definición asociada con cada objeto y agrupación de objetos comprende las primeras indicaciones de interrelación y, opcionalmente, las segundas indicaciones de interrelación, asociadas con al menos uno de los elementos representativos. En algunos casos, las definiciones pueden incluirse con el conjunto correspondiente de objetos y agrupaciones de objetos en cada archivo de información digital resultante (es decir, almacenadas juntos en el archivo). Sin embargo, en otros casos, las definiciones asociadas con el conjunto de objetos y agrupaciones de objetos se pueden mantener (es decir, almacenar) por separado de cada archivo de información digital resultante correspondiente. Por lo tanto, cada uno de los archivos de información digital resultante se puede formar de tal manera que cada archivo de información digital resultante posterior comprenda cualquier patrón reconocido de objetos y agrupaciones de objetos dentro del archivo de información digital resultante anterior.

En otro aspecto, los procesos de manipulación pueden aplicarse sucesivamente hasta que las aplicaciones sucesivas del proceso de manipulación no logren una reducción de umbral en la cantidad de información en el archivo de información digital resultante posterior. En tales casos, el último archivo de información digital resultante

puede tener una estructura de primotrón de información que comprende un conjunto de objetos primarios, agrupaciones de objetos primarios, y cualquier interrelación asociada con los mismos. Después de la aplicación de los procesos de manipulación, cada archivo de información digital resultante y el último archivo de información digital resultante se pueden analizar para determinar las correlaciones estadísticas entre los objetos posteriores en un

- 5 archivo de información digital posterior y las combinaciones anteriores de objetos anteriores en un archivo de información digital anterior. Además, para cualquier correlación estadística que cumpla al menos con un umbral de correspondencia, una indicación de correlación del objeto posterior correspondiente y la combinación anterior de los objetos se almacena para el análisis de un archivo de información digital original posterior.
- 10 En algunos casos, cada archivo de información digital resultante sucesivo puede analizarse para determinar al menos una correlación objetiva entre una combinación de objetos anteriores, y al menos una relación entre objetos entre los objetos anteriores en la combinación, en un archivo de información digital anterior. La al menos una correlación objetiva y la al menos una relación entre objetos se asocian entonces con un objeto posterior correspondiente en un archivo de información digital posterior. La al menos una correlación objetiva y la al menos
- 15 una relación entre objetos asociada con el objeto posterior correspondiente, sobre una pluralidad de objetos posteriores, pueden formar colectivamente de este modo un conjunto de correlaciones de objetos representativas de las interrelaciones asociadas con el conjunto de objetos primarios y agrupaciones de objetos primarios en la estructura de primotrón de información.
- 20 En otros casos, cada archivo de información digital resultante sucesivo puede analizarse para determinar al menos una correlación objetiva entre una combinación de objetos anteriores en un archivo de información digital anterior, y al menos una relación entre objetos puede asignarse o designarse de otro modo entre los objetos anteriores en la combinación. La al menos una correlación objetiva y la al menos una relación entre objetos pueden asociarse entonces con un objeto posterior correspondiente en un archivo de información digital posterior. La al menos una
- 25 correlación objetiva y la al menos una relación entre objetos asociada con el objeto posterior correspondiente, sobre una pluralidad de objetos posteriores, pueden formar colectivamente de este modo un conjunto de correlaciones de objetos representativas de las interrelaciones asociadas con el conjunto de objetos primarios y agrupaciones de objetos primarios en la estructura de primotrón de información.
- 30 En cualquier caso, el conjunto de objetos primarios y las agrupaciones de objetos primarios en la estructura de primotrón de información puede entonces interpretarse mediante la evaluación estadística de diversas combinaciones de las correlaciones de objetos en el conjunto de correlaciones de objetos y, para cualquier combinación de correlaciones de objetos que cumpla al menos un umbral estadístico, asociando las combinaciones con los objetos primarios y las agrupaciones de objetos primarios como una interpretación contextual del conjunto de
- 35 objetos primarios y las agrupaciones de objetos primarios en la estructura de primotrón de información. Un archivo de información digital representativo asociado con el archivo de información digital original se puede sintetizar entonces, por ejemplo, aplicando la interpretación contextual al conjunto de objetos primarios y agrupaciones de objetos primarios en la estructura de primotrón de información.
- 40 Además, a partir del análisis de cada archivo de información digital resultante sucesivo, al menos un objeto, seleccionado de los objetos anteriores, los objetos posteriores, y los objetos primarios y agrupaciones de objetos primarios, pueden asociarse con una de una pluralidad de características del archivo de información digital original. De esta manera, una disposición de la pluralidad de características puede determinarse con respecto a la duración del archivo de información digital original, y una distribución de al menos un objeto puede correlacionarse con
- 45 respecto a la misma. En tales casos, se puede formar un archivo de datos de resumen, incluyendo el archivo de datos de resumen una representación generalizada de la distribución de al menos un objeto asociado con cada una de la pluralidad de características a lo largo de la duración del archivo de información digital original.
- En casos en los que la aplicación sucesiva de procesos de manipulación hasta aplicaciones sucesivas del proceso
- 50 de manipulación no logra una reducción de umbral en la cantidad de información en el archivo de información digital resultante posterior, una cantidad y un orden de los procesos de manipulación aplicados sucesivamente al archivo de información digital resultante posterior para obtener el último archivo de información digital resultante que tiene la estructura de primotrón de información se pueden catalogar, en donde la cantidad de procesos de manipulación está asociada con los niveles correspondientes de una jerarquía de información.
- 55 Como tal, de acuerdo con algunos aspectos, un archivo de información digital representativo asociado con el archivo de información digital original puede sintetizarse, por ejemplo, de acuerdo con un proceso condicional que comprende las siguientes etapas:

1. Aplicar cualquier indicación de correlación aplicable al conjunto de objetos primarios y agrupaciones de objetos primarios en la estructura de primotrófon de información para determinar la combinación anterior correspondiente de objetos asociados con los mismos;
 - 5 2. Si no se determina ninguna indicación de correlación aplicable en la Etapa 1, aplicar la interpretación contextual al conjunto de objetos primarios y agrupaciones de objetos primarios en la estructura de primotrófon de información para determinar la combinación anterior correspondiente de objetos asociados con los mismos; y
 - 10 3. si la aplicación de la interpretación contextual en la Etapa 2 está por debajo de un valor de información de umbral asociado con el valor de información original, aplicar un proceso de manipulación inverso inicial al conjunto de objetos primarios y agrupaciones de objetos primarios en la estructura de primotrófon de información, en donde el proceso de manipulación inverso inicial es la primera manipulación en el orden inverso de los procesos de manipulación, para determinar la combinación anterior correspondiente de objetos asociados con el mismo.
- 15 En algunos casos, para la combinación anterior de objetos determinada a partir del conjunto de objetos primarios y agrupaciones de objetos primarios en la estructura de primotrófon de información, puede aplicarse al mismo el proceso condicional que incluye cualquier indicación de correlación aplicable, la interpretación contextual correspondiente, o el siguiente proceso de manipulación inversa en el orden inverso de los procesos de manipulación correspondientes a la combinación anterior de objetos, con el fin de determinar la siguiente combinación anterior de objetos
- 20 correspondientes con el mismo. Por lo tanto, puede desprenderse que el proceso condicional puede aplicarse sucesivamente a la siguiente combinación anterior de objetos hasta que la combinación anterior resultante de objetos se asocie con el nivel inicial de la jerarquía de información, y el archivo de información digital anterior resultante corresponde de este modo al archivo de información digital representativo sintetizado.
- 25 Como se ha descrito previamente, la formalización MSML puede referirse a un método de estructuración de información, en el que dicha información puede comprender, por ejemplo, información digital tal como datos de audio. Por consiguiente, otro aspecto más de la presente descripción puede implicar un método de evolución de primotrones en diferentes niveles de la formalización MSML (es decir, estructura jerárquica) para el análisis de sonido o audio, en donde la evolución de tales primotrones se puede lograr mediante la aplicación de interfaces
- 30 independientes a la formalización general MSML, en donde tales interfaces pueden comprender, por ejemplo, procesos de manipulación para analizar y modificar la estructura de datos de la formalización MSML para facilitar tal evolución de primotrones. Tal aspecto también puede incluir un método de síntesis de sonido que utiliza el conjunto o combinación de primotrones definidos y determinados en asociación con la formalización MSML, e instrucciones y reglas particulares que definen la señal compleja proporcionada por la formalización MSML, a través de la aplicación
- 35 de las interfaces de a las mismas, sucesiva y/o concurrentemente, así como cualquier evolución de dichos primotrones asociados con las mismas, para sintetizar o reconstruir una señal o archivo de sonido o audio que se corresponda sustancialmente con el archivo de audio digitalizado original.

En un ejemplo de tales interfaces (es decir, que representan un proceso de manipulación que se puede aplicar a un

40 archivo de audio digital u otro archivo de información) aplicado a la formalización general MSML, la reducción del ancho de banda y la de la tasa de bits de los datos de audio puede realizarse utilizando una interfaz de codificación sin fases para la formalización MSML de la composición de sonido (véase, por ejemplo, la Figura 3A). Tal interfaz de codificación sin fases se puede configurar, por ejemplo, para considerar una jerarquía de valores de información perceptiva para determinar qué datos del archivo se pueden conservar, transformar o descartar. Más

45 particularmente, una vez que se ha analizado la señal de audio digital y se ha determinado la formalización MSML correspondiente, se puede aplicar compresión sin pérdida, por ejemplo, mediante un codificador de rango para procesar 1) estructuras armónicas en forma de un raster 2D; 2) envolvente de energía total con alta resolución de tiempo y varias bandas de frecuencia; y 3) relación entre la energía del ruido y la energía total. El método de codificación sin fases se puede aplicar a la formalización MSML para realizar, por ejemplo, tareas dirigidas a la

50 reducción seleccionada de los datos en la formalización MSML, mientras se conservan las estructuras de sonido de alta informatividad (es decir, un alto valor de información) asociada con la composición del sonido, por ejemplo, a través de la comparación con métricas de sonido conocidas. Generalmente, un método de este tipo realiza una comparación de objetos de sonido de diferentes niveles en el espacio, para determinar la integridad de las estructuras y el grado de degradación de las mismas que puede resultar dentro de una representación reducida (es

55 decir, una cantidad de información o una tasa de bits inferior) de la señal de audio después de la filtración y/o el procesamiento, lo que facilita la selección de un parámetro adecuado para cada elemento de sonido para lograr la reducción de la cantidad de información/tasa de bits al tiempo que se conserva la informatividad. En algunos casos, la interfaz de codificación sin fases está configurada de manera apropiada de tal forma que la dependencia del tiempo de la señal no se conserve necesariamente, y las estructuras de sonido se conserven en lugar de la forma de

onda de la señal, dando como resultado de este modo una disminución del ancho de banda esencial y la tasa de bits. Para lograr una mayor velocidad de ancho de banda de audio y reducción de la tasa de bits, las estructuras de nivel relativamente más informativas de la formalización de composición de sonido MSML pueden estar configuradas para mantener parámetros constantes. Los parámetros de las estructuras de nivel relativamente menos informativas pueden, por lo tanto, conservarse con menos precisión o pueden generarse durante la decodificación, por ejemplo, con el uso de valores aleatorios o parámetros apropiados.

En otro ejemplo de tales interfaces/procesos de manipulación que pueden aplicarse a la formalización MSML general, una interfaz de análisis de estructura de datos de audio para una formalización de composición de sonido MSML puede configurarse para extraer elementos repetitivos difusos de tales composiciones de música (véase, por ejemplo, las Figuras 4A y 4B). Tal interfaz de análisis de estructura puede configurarse para procesar datos de audio asociados con el archivo de datos de audio de formalización MSML para identificar elementos de bucle repetitivos difusos dentro de un nivel relativamente bajo de la jerarquía de la formalización MSML. En tales casos, los elementos de bucle repetitivos difusos identificados como iguales pueden ser reemplazados por un elemento representativo y parámetros caóticos asociados para proporcionar de este modo la cantidad de información/datos o la reducción de la tasa de bits, sin disminuir apreciablemente el valor de información general o la informatividad. Tal interfaz puede comprender, por ejemplo, tres módulos. El primer módulo puede estar configurado para proporcionar la búsqueda, el posicionamiento y el marcado de longitud de todos los elementos repetitivos difusos a lo largo de todo el nivel inicial de la formalización MSML. El segundo módulo puede configurarse para extraer un componente de sonido general para cualquier conjunto de fragmentos repetitivos difusos y para definir el nivel de cohesividad de los datos formando una señal residual correspondiente. El tercer módulo puede configurarse para realizar una construcción predictiva del fragmento de audio resultante integrando los elementos repetitivos extraídos, los datos de marcado correspondientes y la información residual a partir de la señal residual. Puede usarse un método de filtrado para el modelado de fases para minimizar la información de la fase inicial requerida durante el proceso de construcción de sonido para el fragmento de audio resultante. Esta interfaz puede configurarse adicionalmente para usarse en una diversidad de aplicaciones, incluyendo, pero sin limitación, motores de búsqueda de música, visualización de música, y tecnologías de compresión profunda de audio.

Otra interfaz/proceso de manipulación ejemplar que puede aplicarse a la formalización MSML general implica la compresión de audio utilizando estructuras repetitivas de jerarquía en la formalización de composición de sonido MSML (véase, por ejemplo, la Figura 5). Esta avanzada interfaz de codificación de repetición sin fases identifica las estructuras formadas por el sonido de macrorepetición o los primotrones de audio en el nivel más alto de una jerarquía de sonido de formalización MSML para proporcionar de este modo una cantidad adicional de información/datos o una reducción de la tasa de bits, sin disminuir de forma apreciable el valor de la información general o la informatividad. Se utilizan dos tipos de datos para el proceso de codificación: datos a gran escala que definen la estructura general dentro de la formalización MSML y datos que comprenden fragmentos de señales originales descritos por un nivel relativamente inferior dentro de la jerarquía. Para comprimir los fragmentos de señal, se puede utilizar un codificador psicoacústico general, en algunos casos, posiblemente utilizando un enmascaramiento de señal externo. La reducción de la tasa de bits se produce debido a la eliminación de fragmentos repetitivos determinada en el nivel más alto de la jerarquía de formalización MSML, al tiempo que se almacena o se mantienen las ubicaciones de estos fragmentos. La equivalencia perceptiva de fragmentos repetitivos puede permitir una reducción en el número de muestras de referencia requeridas para las comparaciones.

Aún otra interfaz/proceso de manipulación ejemplar que puede aplicarse a la formalización MSML general implica el ancho de banda de audio y la reducción de la tasa de bits utilizando una interfaz de compresión sin pérdida para una formalización de composición de sonido de MSML (véase, por ejemplo, la Figura 6A) que puede configurarse para reducir el número de matrices de valores enteros asociadas con la formalización MSML mediante el uso de un codificador de entropía avanzado que realiza solo operaciones de enteros, sin divisiones y estimaciones de probabilidad secundarias. Dado que el codificador de rango y el modelo contextual de la interfaz no utilizan la operación matemática de división, permite la aplicación de esta interfaz a y/o la ejecución de la interfaz por parte de muchos microordenadores de bajo coste que no tienen una operación de división. Para mejorar la fiabilidad del modelo contextual de la interfaz, puede emplearse una estimación de probabilidad secundaria (véase, por ejemplo, la Figura 6B). La estimación de probabilidad secundaria es un modelo contextual complejo que utiliza la probabilidad predicha como un contexto con respecto a otro modelo. Es decir, por ejemplo, la estimación de probabilidad secundaria puede implementarse como uno de los parámetros caóticos que relacionan diferentes niveles de la jerarquía. En este sentido, la estimación de probabilidad secundaria bidimensional también se puede usar para mezclar dos modelos diferentes, en donde la estimación de probabilidad secundaria bidimensional es una modificación de la estimación de probabilidad secundaria, que utiliza dos probabilidades de entrada para determinar un contexto. En algunos casos, también se pueden usar contextos enteros adicionales.

Todavía otra interfaz/proceso de manipulación ejemplar que se puede aplicar a la formalización MSML general implica la compresión de señales de audio para mejorar la codificación aritmética y el modelado contextual, para aumentar de este modo las relaciones de compresión y las velocidades de procesamiento, y permitir que los datos espectrales se compriman con alta eficacia y mayor velocidad. Tal método de compresión puede ser completamente automatizado y no requiere inicialización previa para diferentes tipos de datos de audio. Es lo suficientemente flexible para ajustarse a diferentes tamaños de datos de audio espectral, lo que permite su uso con diferentes transformaciones espectrales. En lugar de un codificador aritmético estándar, se utiliza un codificador de rango más eficiente. El modelado de contexto se aplica al flujo de datos, los modelos algorítmicos construidos y la optimización algorítmica de una función de decodificador. Este aspecto también se basa, al menos parcialmente, en el uso de técnicas de codificador de rango adaptativo que implican aumentar la probabilidad del valor codificado. Para mejorar la fiabilidad del modelo contextual, se puede emplear una estimación de probabilidad secundaria.

Otra interfaz/proceso de manipulación ejemplar que se puede aplicar a la formalización MSML general implica una transmisión eficiente de datos de audio utilizando una interfaz de acelerador avanzado de datos de audio para una formalización de composición de sonido MSML, y el uso de un proceso de pérdida de calidad nula (NQL) para reducir el número de coeficientes espectrales de tiempo-frecuencia cuantificados (QMDCT), para proporcionar de este modo una entrega más rápida de los datos de audio sin una degradación significativa de la calidad en el sonido resultante (véanse, por ejemplo, las Figuras 7A y 7B). El proceso NQL realiza una clasificación de las estructuras de sonido dentro de la formalización MSML en función de su importancia, en una comparación basada, por ejemplo, en el límite o umbral de la percepción humana. Es decir, por ejemplo, la pérdida de calidad de información será mínima, si la hay, si la cantidad de información se reduce en datos que se aproximan o superan el límite o el umbral de la percepción humana. La avanzada interfaz del acelerador puede configurarse para reconstituir los coeficientes espectrales cuantificados de tiempo-frecuencia del formato inicial, sin la descuantificación de los datos ni la transcodificación inversa a través del formato PCM. La avanzada interfaz del acelerador puede, por ejemplo, dividir el archivo codificado inicialmente de los coeficientes espectrales del dominio del tiempo-frecuencia en tres grupos, donde los coeficientes de los grupos espectrales de baja frecuencia se mantienen sin cambios, mientras que algunos coeficientes de los grupos espectrales de frecuencia intermedia pueden ponerse a cero utilizando el proceso NQL, y los coeficientes del tercer grupo (frecuencias más altas) pueden eliminarse reemplazándolos con los subgrupos de coeficientes más similares, uniéndolos en grupos bajos e intermedios, y guardando una referencia a este subgrupo buscado, así como un factor integral, sin afectar a los elementos de sonido estructurales (armónicos y aciertos).

Otra interfaz/proceso de manipulación ejemplar que puede aplicarse a la formalización MSML general implica reducir el tamaño del archivo de datos de audio (cantidad de información o tasa de bits), utilizando una interfaz de codificación de transformadas afines para una formalización de composición de sonido MSML (véase, por ejemplo, la Figura 8A), conservando al mismo tiempo la integridad (valor de información) del archivo de datos de audio original. Tal interfaz puede implementar una transformada afín de primotrones de sonido/audio similares, lo que conduce a una disminución en la entropía de la señal (es decir, una estructuración más alta proporciona menos entropía y, posiblemente, una disminución en la cantidad de información o la tasa de bits). La interfaz de transformada afín asociada con la formalización MSML se puede usar como una adición o un complemento para cualquier procedimiento de codificación de audio, por ejemplo, como los descritos en el presente documento, para aumentar la relación de compresión del mismo, así como para proporcionar un proceso de compresión central, para señales de sonido/audio donde los microfragmentos similares afines en el dominio del tiempo-frecuencia son esenciales para el valor de la información de sonido. Esta interfaz de codificación de transformadas afines, al mismo tiempo que puede usarse junto con la formalización MSML, está configurada para utilizar microfragmentos espectrales similares de la señal de audio digital para aumentar la relación de compresión de esa señal. Durante el procesamiento, el rango espectral completo se divide en sub-bandas. Se realiza una búsqueda de fragmentos espectrales similares en la parte anterior de la señal de sonido de forma independiente y utilizando una etapa diferente en cada sub-banda. Se usa una transformada afín de cambio de amplitud, tiempo y espectral durante la búsqueda de fragmentos similares. Cada fragmento encontrado se resta del fragmento espectral original y el residuo se procesa adicionalmente como una diferencia de componentes espectrales, eliminando de este modo una redundancia significativa de los datos espectrales, y reduciendo de este modo la cantidad de información o la tasa de bits en el proceso.

En términos prácticos (como se ilustra, por ejemplo, en las Figuras 9-11), un archivo de audio digital está representado por una matriz de bits, que después puede analizarse de manera apropiada para detectar patrones de bits que representan diversos niveles de informatividad, y tales patrones de bits básicos pueden denominarse "primotrones" de acuerdo con diversos aspectos de la presente divulgación abordada en el presente documento.

Debido a que la matriz de bits puede extenderse sobre múltiples dimensiones, y dado que los patrones de bits pueden tener diferentes estructuras/combinaciones que pueden abarcar más de una dimensión de la matriz de bits (es decir, proporcionar diversos niveles de informatividad), los primotrones pueden caracterizarse como objetos o entidades de multinivel y multiestructurales (MSML) que proporcionan un paradigma novedoso para la formalización digital de medios tal como el sonido o el audio. Tal procedimiento de formalización y el archivo de audio digital formalizado producido a partir del mismo se describen en la presente solicitud, así como diversos métodos mediante los cuales los primotrones pueden identificarse o determinarse de otra manera, y diversos métodos mediante los cuales dichos primotrones pueden (re)combinarse o analizarse de otro modo para sintetizar una representación del archivo de audio digital original que se encuentra dentro de un umbral del mismo. Dado que un aspecto de la presente descripción implica reducir la cantidad de información o la tasa de bits del archivo multimedia digital original, mientras se mantiene una cantidad de información dentro de su umbral, un experto en la técnica apreciará que se pueden obtener beneficios y ventajas adicionales en términos de, por ejemplo, transmisión de datos, almacenamiento de datos, seguridad de datos y similares, como se detalla adicionalmente en el presente documento.

15

Dado que un primotrón, determinado por aspectos de un método de formalización como se describe en el presente documento, representa un patrón o combinación de uno o más bits o elementos de datos en los diversos niveles y/o dimensiones de la matriz de bits del archivo de audio digital, se deduce que el archivo de audio digital puede representarse posteriormente mediante una combinación de dichos primotrones identificados y/o determinados a incluir en el mismo. Es decir, el archivo de audio digital puede estar sujeto a formalización MSML, y luego se puede analizar la formalización MSML para determinar la combinación de dichos primotrones o elementos de datos presentes en el mismo. Tal análisis puede, por ejemplo, caracterizarse como un "sistema vivo" de patrones de sonido que interactúan en condiciones particulares dentro de una cantidad limitada de información. La evolución o "historia de vida" de dichos patrones puede caracterizarse adicionalmente por los principios de evolución de la información en un sistema caótico dinámico definido por un número finito de estructuras, concretamente, los primotrones tienen cada uno una "vida útil" que puede variar desde una fracción de segundo hasta varios minutos dentro del tiempo o la duración del tiempo-frecuencia del archivo de audio digital. Debido a que el archivo de audio digital ahora está representado por tal combinación o sistema vivo de primotrones, en donde cada primotrón representa un subconjunto de bits de la matriz de bits para u otra estructura o entidad asociada con ese archivo de audio digital, el archivo de audio MSML que incluye la combinación identificada de primotrones puede presentar, por ejemplo, un tamaño del archivo de datos reducido (cantidad de información) en comparación con el archivo de audio digitalizado original (al mismo tiempo que conserva la calidad de la información del archivo de audio digital dentro de un umbral perceptivo del archivo de audio digital original). En algunos casos, tal reducción puede ser, por ejemplo, del orden de aproximadamente 20 veces a aproximadamente 60 veces menor que el tamaño del archivo de datos original, aunque la reducción del tamaño del archivo de datos que puede lograrse puede ser incluso mayor de 60 veces, por ejemplo, a medida que otros primotrones y métodos asociados se identifican o se determinan y refinan de otra manera, en donde dichos primotrones pueden facilitar la reducción en la cantidad de información o la tasa de bits para el archivo de audio digital en particular, mientras se mantiene la informatividad del archivo de audio digital dentro de un umbral perceptivo del archivo de audio digital original. En otros casos, puede ser posible que un subconjunto de la combinación de primotrones u otros elementos de datos en el archivo de datos se represente mediante un objeto más general u otra entidad, lo que reduce aún más el tamaño del archivo de datos en comparación con el archivo de audio digital original (y/o formalización MSML del mismo). Por consiguiente, tal paradigma forma la base de un formato de datos digitales multifuncional y altamente estructurado configurado para un almacenamiento y reproducción altamente compactos y eficaces de cualquier material representado en forma de una señal de audio de alta fidelidad de base o construcción armónica (incluyendo música, voz y otros medios basados en un sonido rico), dentro de un umbral de la señal de datos digital original.

En la identificación de las diversas combinaciones de elementos que forman los diversos primotrones u otros elementos de datos, el elemento representativo resultante también puede asociarse con parámetros particulares (es decir, denominados en el presente documento como parámetros caóticos, o de otro modo como un archivo de datos asociado con características particulares del elemento representativo) que representan, por ejemplo, el contexto u otras características del elemento representativo y/o los elementos de datos o primotrones a partir de los cuales se determinó el elemento representativo. Dichos parámetros pueden comprender, por ejemplo, relaciones entre elementos de datos en el mismo nivel de la jerarquía, relaciones entre elementos de datos a través de diferentes niveles de la jerarquía, relaciones entre elementos de datos de una jerarquía de nivel inferior y su relación con un elemento de datos en una jerarquía de nivel superior. Además, dado que dichos primotrones se definen dentro de una cantidad limitada de información finita, los patrones identificados asociados con dichos primotrones pueden asociarse con diversos aspectos correspondientes del archivo de audio digital. Por ejemplo, ciertos primotrones pueden representar armónicos, mientras que otros pueden representar voces, instrumentos particulares, frecuencias

o rangos de frecuencia específicos, acústica del lugar, o cualquier otro aspecto identificable del archivo de audio digital asociado con una actuación musical u otra basada en armónicos.

Por lo tanto, se deduce que un aspecto de la formalización MSML (es decir, la representación del archivo de audio digitalizado por una combinación de primotrones) del archivo de audio digital es que los bits en la matriz de bits que no están necesariamente identificados/determinados como un componente de un primotrón en la combinación de primotrones, se puede designar como ruido (es decir, ruido blanco o, de otro modo, ruido periférico que no es necesariamente significativo para la naturaleza del archivo de audio digital que se está convirtiendo en una formalización MSML (es decir, tiene un efecto prejudicial limitado en la calidad de la información del archivo de audio digital). En tales casos, es posible que el ruido no se incluya o se elimine de otro modo de la representación MSML. Sin embargo, puede haber casos en los que algunos o todos los "ruidos" puedan ser deseables para permanecer incluidos en la representación MSML. Por ejemplo, en casos en que el archivo de audio digital representa una actuación musical en vivo, parte del ruido puede atribuirse a la acústica del lugar de la actuación o la reacción de la audiencia a la actuación. Como tal, puede ser deseable, en algunos casos, incluir dicho "ruido" con la representación MSML del archivo de audio digital para proporcionar vida, ambiente o de otro modo un contexto al contenido del archivo de audio. De este modo, el "ruido deseable" se puede separar en elementos de ruido, siendo cada elemento de ruido representado por un primotrón periférico respectivo. De tal manera, uno o más primotrones periféricos, o diversas combinaciones de los mismos, pueden incluirse selectivamente en la representación MSML. Es decir, dichos primotrones periféricos pueden incluirse en la representación de MSML si es necesario o deseable, o eliminarse de otro de la representación MSML.

Otro aspecto asociado con lo anterior es que, dado que la representación MSML del archivo de datos original está en la forma de una combinación de primotrones u otros elementos de datos, y dado que los primotrones periféricos pueden identificarse/determinarse e incluirse selectivamente en o excluirse de la representación MSML, entonces es posible que una o más subcombinaciones de la combinación de primotrones se identifiquen/determinen y se incluyan selectivamente o se excluyan de la representación MSML. Por ejemplo, con respecto a la interpretación musical en vivo mencionada anteriormente, la combinación de primotrones incluida en la representación MSML puede incluir solo la música de los instrumentos que se tocan. En tales casos, puede ser posible, por ejemplo, identificar la música de instrumentos individuales de tal manera que la interpretación se pueda separar en "pistas", cada una asociada con un instrumento respectivo. Por consiguiente, el procesamiento agregativo eficiente simultáneo y sincronizado puntualmente de la información relativa y relacionada con la señal de audio y su entrega en red rastreable, puede optimizarse dinámicamente sobre una infraestructura de IP terrestre y móvil, en donde tales archivos de audio representados por MSML pueden ser capaces de transportar de forma simultánea y extraer dinámicamente representaciones de señales planas y volumétricas en alta definición (incluyendo, por ejemplo, sonido estéreo 2.0 y sonido envolvente 5.1) y dentro de un solo archivo de datos a una tasa de bits baja, y también pueden ser capaces de extraer en tiempo real una voz cantada, con una calidad de sonido completamente transparente y reproducible de la fuente de sonido original. Dado que una característica particular de las representaciones MSML de los archivos de audio digital es una cantidad de información o tasa de bits significativamente inferior, que tiene un valor de información dentro de un umbral o incluso superior a la calidad de la información del archivo de audio digital original, dichas capacidades ejemplares descritas en el presente documento pueden reducir significativamente los gastos relacionados con el almacenamiento digital, la transmisión, la difusión y el ancho de banda de los sistemas de centralidad mediática y redes, al mismo tiempo que se abren nuevas fuentes de ingresos y se aumenta la monetización del contenido.

Por ejemplo, otro aspecto asociado con la capacidad de analizar el archivo de datos de audio original o de otro modo segregar el archivo de datos de audio de acuerdo con características particulares con respecto a una representación MSML es que, en algunos casos, puede ser posible crear o transferir primotrones secundarios particulares para la inclusión selectiva en la representación MSML. Por ejemplo, con respecto a la actuación musical en vivo mencionada anteriormente, las características de la acústica de un lugar diferente se pueden capturar o recrear digitalmente y luego convertir en una representación MSML como uno o más primotrones. En tales casos, puede ser posible que el componente del lugar se elimine de la representación MSML del archivo de audio digital y luego se reemplace con el primotrón o primotrones asociados con el lugar diferente. Por lo tanto, un experto en la técnica apreciará que los primotrones pueden disponerse en muchas combinaciones diferentes para proporcionar una representación MSML deseada, ya sea que los primotrones se originen o no a partir del archivo de audio digital original.

Además del concepto de disponer primotrones en muchas combinaciones diferentes para proporcionar una representación MSML deseada, un experto en la técnica también apreciará que se pueden crear primotrones u otros elementos de datos para representar otras entidades dentro de la representación MSML del archivo de audio digital.

Por ejemplo, un primotróon de este tipo u otro elemento de datos puede asociarse con indicaciones únicas que representan, por ejemplo, al propietario o poseedor de la representación MSML del archivo de audio digital. Más particularmente, tras la conversión del archivo de audio digital a formato MSML, se puede generar una combinación única de primotrones u otros elementos de datos para identificar al usuario particular que tiene derecho a la representación MSML del archivo de audio digital, ya sea el usuario, por ejemplo, un propietario de contenido, un licenciario de contenido, un comprador de contenido o, de otro modo, una entidad que adquiere un derecho de acceso al contenido. En algunos casos, las indicaciones únicas pueden configurarse para que sigan siendo una parte de la representación MSML (es decir, no "borrables" o removibles de otro modo) independientemente de la manipulación del archivo de representación MSML por parte del usuario o el poseedor o poseedores posteriores del mismo. Por consiguiente, como se apreciará por los expertos en la técnica, las indicaciones únicas proporcionar, en algunos casos, una "marca de agua" o, de lo contrario, una característica de seguridad para la representación MSML particular del archivo de audio digital original. Tales indicaciones únicas también pueden configurarse o disponerse, en algunos casos y configuradas para conservar una parte de la representación MSML o no, para representar otros aspectos del archivo de representación MSML, en donde tales indicaciones únicas pueden comprender, por ejemplo, metadatos (es decir, palabra clave, referencia, clasificación, datos de seguridad, etc.), como apreciará un experto en la técnica.

En algunos casos, las indicaciones únicas incluidas en la representación MSML del archivo multimedia digital pueden hacerse evidentes o controlarse de otra manera a través de un canal de datos dedicado configurado para funcionar junto con la señal armónica codificada para reflejar dinámicamente y mostrar información interestructural en el dominio del tiempo, permitiendo de este modo la reproducción sincronizada de audio con cualquier evento exterior o externo. Por ejemplo, tras determinar la estructura de primotróon de un archivo multimedia digital particular, los primotrones que comprenden el archivo, incluidas las indicaciones únicas, pueden disponerse en el dominio del tiempo. Con el conocimiento de la estructura o las características particulares de los primotrones particulares y la ubicación de los mismos en el dominio del tiempo en asociación con la duración del archivo multimedia digital, dicho conocimiento se puede usar para impulsar ciertos eventos o para tener ciertos casos de información asociados con los mismos. Dichos eventos externos pueden incluir, por ejemplo, vídeo o una secuencia de programación en un juego o película, letras de canciones, portadas, credenciales de artistas y compositores o partituras, varios efectos especiales visuales y pirotécnicos, incluyendo iluminación interior y exterior o audiolibros visuales, publicidad dirigida o cualquier otra información relacionada con los medios de sonido reproducidos. Es decir, por ejemplo, un primotróon o una combinación de primotrones puede designarse dentro del canal de datos dedicado como impulsor de un evento o efecto externo particular. Cuando el canal de datos dedicado tiene el archivo de audio formalizado en MSML combinado con el mismo o introducido de otro modo en el mismo, las apariciones de primotrones particulares o combinaciones de los mismos a lo largo de la duración del archivo de audio formalizado en MSML sirven para dirigir el evento o efecto externo correspondiente en el momento particular durante la ejecución del archivo de audio formalizado en MSML, aumentando de este modo la informatividad y la personalización de los medios recibidos, y mejorando sustancialmente el valor de entretenimiento y la monetización de los mismos.

Más particularmente, con respecto al aspecto del canal de datos dedicado, durante el proceso de codificación y conversión de archivos datos de audio en bruto (PCM, wav, aiff) y MP3, AAC, OGG, WMA de alta tasa de bits (es decir, aproximadamente 192 kbit/s, y superior) en la formalización MSML, se identifican y extraen un conjunto de estructuras que tienen propiedades y características únicas que representan los datos de audio (es decir, primotrones y descripciones de múltiples niveles que utilizan los mismos). La representación de los datos de audio de tal manera permite que se proporcione un canal de datos interactivo, completamente integrado e inteligente, en asociación con la formalización MSML. Por ejemplo, el canal de datos puede configurarse para incluir una diversidad de información relacionada y asociada con la formalización MSML de los datos de audio u otros datos multimedia. Más particularmente, el canal de datos puede configurarse para recibir datos indicativos de uno o más eventos asociados con un elemento de primotróon particular y/o de descripción de nivel múltiple en la formalización MSML. En otros casos, los datos pueden asociarse generalmente con la composición musical general. Por ejemplo, un evento de este tipo puede incluir un atributo de audio y/o sonido asociado con una nota, una o más letras con tiempo y duración particulares asociados con una palabra, la "partitura" asociada con la composición musical, la identificación de la licencia, los derechos de autor y la información legal junto con una diversidad de datos subordinados, tales como enlaces a otras opciones de audio, imagen y vídeo, materiales de texto relacionados, incluidas noticias, eventos o similares.

En un ejemplo, una composición de sonido puede procesarse en la formalización MSML asociada, que incluye los primotrones mencionados anteriormente y las descripciones de múltiples niveles. En la preparación de la formalización MSML, el archivo de datos correspondiente puede configurarse para que los datos en el mismo puedan comunicarse con datos externos (es decir, el archivo de datos puede configurarse para comunicarse con el

"canal de datos"). Dichos datos externos pueden configurarse para relacionar un primotrón particular, una descripción particular de múltiples niveles, o combinaciones de los mismos, con una acción, evento, efecto, etc. particular. Por ejemplo, los datos externos pueden configurarse para ejecutar el funcionamiento de fuentes de agua decorativas. Por consiguiente, diversas características de la composición de sonido pueden correlacionarse con 5 diversas características operativas de las fuentes de agua. En aspectos particulares, por ejemplo, el volumen a lo largo de la duración del tiempo de la composición de sonido puede correlacionarse con el volumen o la presión del agua dirigida a través de una o más de las fuentes; diversos instrumentos pueden estar correlacionados con diversos subgrupos de las fuentes; las letras pueden estar correlacionadas y sincronizadas con luces que iluminan las distintas fuentes. Por consiguiente, los diversos primotrones y las descripciones de múltiples niveles que se 10 determinan a partir de la formalización MSML para la composición de sonido particular pueden configurarse automáticamente para controlar o de otra manera afectar a la aplicación particular. Como tal, la aplicación particular puede cambiarse fácilmente simplemente asociando una formalización MSML de una composición de sonido diferente con la aplicación particular a través del canal de datos dedicado. Un ejemplo similar puede vincularse a la pirotecnia, tal como, por ejemplo, en una exhibición de fuegos artificiales. Como tal, cuando se configura de esta 15 manera, la composición de sonido se utiliza para "dirigir" la aplicación particular por la naturaleza misma de la formalización MSML de esa composición de sonido. Es decir, la configuración particular del canal de datos puede seguir siendo la misma (es decir, ciertos primotrones o combinaciones de los mismos controlan eventos externos correspondientes particulares), pero el cambio de la composición del sonido se produce una formalización MSML diferente asociada con esa composición de sonido particular y, debido a que los eventos de primotrón tendrán lugar 20 en diferentes momentos a lo largo de la duración del tiempo que otra composición de sonido, los eventos externos asociados con el canal de datos serán controlados en una secuencia diferente que con una composición de sonido diferente. Por consiguiente, la formalización MSML de una composición de sonido puede simplemente reemplazarse con una formalización MSML de otra composición de sonido para afectar a un cambio en la aplicación particular, por ejemplo, un videojuego basado en las características de una composición musical (es decir, un videojuego basado 25 en karaoke). Tal esquema contrasta fuertemente con el estado de la técnica en el que cada composición de sonido tendría que ser analizada y los eventos individuales asociados con la aplicación particular tendrían que estar correlacionados con aspectos particulares de esa composición de sonido, en un proceso que consume mucho tiempo. Para cambiar la composición del sonido en tales casos, habría que repetir el minucioso proceso de correlación manual.

30 Un experto en la técnica también apreciará que el concepto de canal de datos también puede tener capacidades adicionales con respecto a la cooperación con una formalización MSML. Por ejemplo, una composición de sonido particular puede, en algunos casos, tener letras y/o vídeos correspondientes asociados. Como se describe, la composición de sonido puede representarse por una formalización MSML de acuerdo con los diversos aspectos de 35 la presente descripción como se detalla en el presente documento. En algunos casos, las letras y/o el vídeo se pueden correlacionar con el perfil de primotrón de la composición de sonido en el dominio del tiempo o dominio de tiempo-frecuencia. Es decir, las apariciones dentro de las letras y/o el vídeo pueden estar asociadas con el perfil de primotrón en el dominio del tiempo de la composición del sonido. Como tal, la interacción entre la formalización MSML de la composición de sonido y el canal de datos que tiene las letras y/o el vídeo asociado con la misma, 40 puede presentar una situación en la que los perfiles entre los mismos pueden estar correlacionados (es decir, de acuerdo con una aparición en el dominio del tiempo, de primotrones particulares o combinaciones de los mismos). En la forma de realización de la correlación, la formalización MSML de la composición del sonido esencialmente se sincroniza con las letras y/o el vídeo asociados con el canal de datos debido a la correspondencia de los perfiles de primotrón del dominio del tiempo. Por consiguiente, en algunos casos, la sincronización dinámica puede lograrse o 45 realizarse, ya que la correlación entre el canal de datos y la formalización MSML de la composición del sonido es esencialmente en tiempo real. Una implicación práctica es, por ejemplo, que la formalización MSML de la composición del sonido puede estar en el proceso de realizarse o ejecutarse, y puede ponerse en interacción con el canal de datos en cualquier momento durante esa duración, en donde, tras dicha interacción, el vídeo y/o las letras asociadas con el canal de datos pueden sincronizarse dinámicamente con la composición de sonido esencialmente 50 en tiempo real. Esto puede permitir, por ejemplo, el acoplamiento y el desacoplamiento a voluntad del canal de datos con la formalización MSML de la composición del sonido, sin tener que recurrir al origen del tiempo (es decir, $t = 0$) para realizar la correlación.

En algunos aspectos, los datos externos asociados con el canal de datos pueden dirigirse a un repositorio especial o 55 ubicación dentro del archivo de datos de la formalización MSML de la composición de sonido. De tal manera, los datos externos se pueden implementar antes y de forma correspondiente con el sonido sintetizado y reproducido a partir de la formalización MSML de la composición de sonido. Por consiguiente, un aspecto de reproducción apropiado puede configurarse para proporcionar acceso al repositorio (datos externos) de información de sonido recopilada, y puede configurarse para recibir el contenido deseado (es decir, diferentes composiciones de sonido y/o

vídeo) a través del canal de datos dedicado, y en sincronidad completa entre los mismos, mientras que simultáneamente se decodifican los datos de sonido para crear una experiencia de usuario de contenido interactivo enriquecida. Dichos aspectos permiten la creación de una música inteligente de múltiples dimensiones y aplicaciones relacionadas que incluyen, pero sin limitación, una variedad de servicios interactivos de música y vídeo, 5 y juegos donde los usuarios y los entornos interactuarán con el medio y entre sí (es decir, una nueva forma de experimentar la música proporcionando a los usuarios contenido interactivo de primera calidad para experiencias de reproducción personalizadas, por ejemplo, donde la música dirige o controla de otro modo la experiencia).

El canal de datos puede estar compuesto por datos recopilados durante el proceso de formalización MSML y 10 síntesis de la misma composición de sonido/datos de audio digital y cualquier información externa vinculada a la composición de sonido, y puede configurarse para incorporar la información deseada en la formalización MSML durante un proceso de codificación correspondiente o como una adición de procesamiento posterior de una formalización MSML ya codificada de acuerdo con, pero sin limitación, la siguiente estructura.

15 Es decir, en algunos aspectos, el canal de datos señalado puede incluir varios tipos de datos informativos relacionados con eventos externos y/o en correspondencia con una formalización MSML de una composición de sonido configurada para interactuar con los mismos, en donde un ejemplo es el siguiente:

- General:

20

- Básico - datos codificados tanto en el encabezado del archivo como en cada bloque (típicamente en un intervalo de 10 s) identificando el contenido de la reproducción, y quién ha codificado y decodificado el archivo original y toda la información avanzada correspondiente:

25

- ID de canción - identificador único de la canción o composición de audio, mediante el cual se puede recibir cualquier información adicional de la canción desde un repositorio informático en la nube
- LDE

30

- ID de codificador
- ID de codificador
- Periodo de validez de la licencia

35

- Nombre de la canción
- Nombre del álbum
- Nombre(s) del (de los) artista(s)
- Número total de patrones
- Número total de repeticiones

- Avanzado

40

- Interno

- Bloque
- Trama

45

- Externo

- Primario
- Secundario

50

- Datos generales que representan la siguiente información:

- Avanzado (datos codificados en el encabezado del archivo)

55

- Datos identificativos de la canción:

- Número de pista dentro del álbum, número de disco (para una colección de múltiples discos), álbum con espacios o sin espacios
- Género musical, subgénero musical, estilo

- Composición primaria o remix
- Idioma(s) principal(es) de la actuación
- 5 - Derechos de autor:
 - Nombre del titular de los derechos de autor
 - Fecha de lanzamiento
 - Tipo de licencia
 - 10 - Duración de la licencia (fecha de inicio y vencimiento)
 - Región de aplicabilidad de la licencia
 - Nombre del estudio de grabación
- Credenciales:
 - 15 - Autor de la música (compositor(es) de la música)
 - Autor de la letra de la canción
 - Lista y número total de diversos instrumentos utilizados para crear la canción y su ID respectiva dentro de la pista
 - 20 - Cada artista instrumental (es decir, los nombres de todos los artistas que tocaron los instrumentos mencionados en la Información general)
 - Ingeniero(s) de audio
 - Diseñador gráfico
- Características técnicas:
 - 25 - Formato de la canción (mono, estéreo, multicanal 5.1, etc.)
 - Características del archivo de salida: Discretización (44,1/48/96/192 KHz), resolución de profundidad de bits (16/24/32)
 - 30 - Desviación del nivel de volumen del sonido respecto al promedio (es decir, cuánto debe ajustarse el nivel de la canción para proporcionar un nivel constante dentro de la canción y/o proporcionar una reproducción continua de nivel equivalente entre varias canciones)
- Interno:
 - 35 - por Bloque:
 - Datos básicos
 - Ritmo
 - Tempo
 - 40 - Definiciones de señales estéreo
 - Máscara de tiempo-frecuencia y otra información mediante la cual el usuario puede filtrar instrumentos específicos de la mezcla
 - Definiciones de melodía:
 - 45 - Tipo de escala
 - Enlace(s) con marca de tiempo a información remota (audio, vídeo, texto) y su duración correspondiente
 - Características sonoras de sello de tiempo y su duración correspondiente:
 - 50 - Aciertos con brillo correspondiente
 - Armonicidad:
 - 55 - Identificación del instrumento y correspondientes definiciones de espacio volumétrico
 - La proporción de cada instrumento dentro de la mezcla y dentro de un canal particular, incluido el retardo de tiempo, etc.
 - Número total de los instrumentos en el bloque
 - Definición de voz/instrumento
 - Vocal
 - Consonante

- Armonias disonantes
- Clases de paso
- Eventos musicales:
 - 5
 - Nivel medio de sonoridad (AVL)
 - Aumento del nivel de sonoridad
 - Disminución del nivel de sonoridad
 - 10
 - Secuencia ascendente de escala
 - Secuencia descendente de escala
 - Ataque de tono
 - Tono señuelo
 - Duración del tono
 - Paso de tono
 - 15
 - Intensidad de tono (sonoridad)
 - Dedal de tono (o calidad)
 - Aspectos aperiódicos:
 - 20
 - transitorios de ataque
 - vibrato
 - modulación de envolvente
 - Patrones y Repeticiones
 - 25
 - Número de patrones por bloque
 - Identificación de patrones
 - Número de repeticiones por bloque
 - Identificación de repetición
 - 30
 - Letras en idioma principal
- por Trama:
 - 35
 - Características de la composición:
 - Letras de canciones:
 - Ubicación en el tiempo y duración de cada palabra de la letra a través de la trama
 - 40
 - Ubicación de los aciertos
 - Características de los aciertos:
 - brillo
 - duración
 - 45
 - Etapas de escala
 - Ubicación de escenas sonoras altamente estructuradas armónicamente
 - nivel de armonicidad
 - nivel de pureza de señal
 - 50
 - ubicación de las repeticiones de armónicos
 - ubicación para cada patrón específico
- Externo:
 - 55
 - Primario (información almacenada, por ejemplo, en el repositorio de música en la nube)
 - Información general
 - Diseño correspondiente
 - Guías adicionales

- Contenido de huella digital de audio
- Música de la ficha de la canción

5 - Secundario (información almacenada, por ejemplo, externamente a un repositorio de música de computación en la nube)

10 - cualquier texto junto con su sincronización temporal y la duración de la visibilidad del texto a través de la canción (fragmento)
 cualquier enlace junto con su sincronización temporal y la duración de su visibilidad a través de la canción (fragmento)

En términos prácticos, el canal de datos dedicado puede ayudar a aumentar la versatilidad de la formalización MSML del archivo de audio digital. Más particularmente, dentro del esquema de formalización MSML, la caracterización de los datos de audio (o vídeo, en algunos aspectos) permite que las representaciones objetivas de esos datos actúen como un representante bajo la formalización MSML. Más particularmente, la cantidad relativamente limitada de representaciones objetivas, así como las definiciones particulares de esas representaciones, permiten disponer diversas combinaciones de estas representaciones de una manera que refleja la configuración de los datos de audio originales (es decir, los datos de audio están representados por diversas combinaciones dentro del conjunto conocido de representaciones objetivas, lo que proporciona un universo definido para caracterizar cualquier composición de sonido). De esta manera, el procedimiento y la estructura de formalización MSML minimizan o eliminan de otro modo los criterios subjetivos o no estándar encontrados cuando se analiza convencionalmente el sonido desde una perspectiva de onda.

El universo definido de representaciones objetivas permite además que las composiciones de sonido se descompongan en elementos individuales (es decir, instrumentos, letras de canciones, etc.), donde dichos elementos pueden incluirse, eliminarse, reemplazarse o manipularse de otra manera, según se desee. En algunos aspectos, una composición de sonido se puede formar utilizando las representaciones objetivas formalizadas MSML como los "componentes básicos" o la base de la composición. Además, los datos externos, tales como los metadatos, la información de seguridad, el material de derechos de autor, etc., pueden añadirse a la formalización MSML, en donde dichos datos externos, mientras se encuentran en forma digital, pueden proporcionarse con representaciones objetivas que pueden distinguirse de las representaciones objetivas de la composición de sonido. De esta manera, los datos externos pueden sintetizarse, implementarse, analizarse o manipularse de otra manera, sin alterar o afectar de otro modo a las características de la propia composición de sonido (es decir, los datos externos pueden configurarse para que puedan distinguirse objetivamente de los datos que representan la composición de sonido formalizada en MSML).

En aspectos particulares, las representaciones objetivas realizadas por la formalización MSML pueden ser útiles, por ejemplo, para analizar de manera objetiva varias composiciones de sonido (es decir, como una herramienta "forense de derechos de autor"), o como un controlador u otro componente asociado de un sistema "dirigido" (es decir, una pantalla pirotécnica, una fuente dinámica, un videojuego, etc.). Es decir, las representaciones objetivas pueden vincularse o configurarse de otro modo para activar las funciones correspondientes que pueden definirse a través de la asociación con datos externos. El canal de datos dedicado mencionado anteriormente se puede usar para introducir una composición de sonido en las funciones definidas (en un formato formalizado en MSML), o viceversa, dentro de la estructura armónica de la representación formalizada en MSML de la composición de sonido.

En otros aspectos, la formalización MSML de la composición de sonido, que da como resultado representaciones objetivas de la composición de sonido, puede facilitar un esquema de eventos basado en elementos, que puede no estar necesariamente basado en, pero podría incluir, un aspecto dependiente del tiempo. Dado que el esquema se dirige por elementos/eventos, diversas características de audio (es decir, tempo, tiempos, compases, sincronización en tiempo real, eventos audibles, eventos emocionales, etc.) podrían representarse de una manera más objetiva y, por lo tanto, manipularse de manera similar con respecto a los componentes básicos de la composición sonora.

En algunos aspectos, las indicaciones únicas mencionadas anteriormente que pueden incluirse como parte de los datos externos pueden evitar la manipulación de la representación MSML por cualquier persona que no sea el usuario. En otros aspectos, las indicaciones únicas pueden servir para identificar al usuario (es decir, el usuario particular que tiene el derecho a la representación MSML particular) en los casos en que dicho archivo es copiado, transferido o distribuido de otra manera por el usuario. Tal característica puede, en algunos casos, limitar o eliminar la copia, el intercambio de archivos o de otro modo el uso no autorizado de la representación MSML particular del archivo de audio digital original. En otros casos, una característica de este tipo puede servir como un mecanismo

para "rastrear" al usuario, incluyendo, por ejemplo, hábitos o tendencias de compra, datos demográficos, o uso de los datos (y tendencias de uso), o como un mecanismo para dirigir facturas al usuario para la compra, uso autorizado, o acceso permitido a la representación MSML. Por consiguiente, las indicaciones únicas pueden, en algunos casos, funcionar dentro del espacio de formato MSML para proporcionar archivos de datos rastreables que pueden atribuirse al usuario particular que tiene ciertos derechos sobre el mismo, en donde dicha funcionalidad puede facilitar una mayor eficiencia e ingresos seguros para el propietario del contenido limitando o disuadiendo de otro modo la redistribución no autorizada del archivo de datos, ya sea en formato MSML o no. Por consiguiente, una representación MSML puede configurarse para extraer una huella digital de sonido única de una señal de audio analógica que, junto con el adorno de datos inaudible, puede proporcionar un método altamente eficaz y seguro de identificación y trazabilidad de medios digitales, sin tener en cuenta o dependiendo de las transformaciones de formato y métodos de entrega.

Muchas modificaciones y otros aspectos de las descripciones expuestas en el presente documento le vendrán a la mente a un experto en la técnica a la que pertenecen estas descripciones que se benefician de las enseñanzas presentadas en las descripciones anteriores y los dibujos asociados. Por ejemplo, los aspectos descritos en el presente documento pueden, en algunos casos, denominarse como una indicación de un formato de datos digital multifuncional configurado para un almacenamiento altamente compacto y eficaz y una reproducción de cualquier material representado en forma de una señal de alta fidelidad de base o construcción armónica (incluyendo música, voz y otros medios basados en sonido ricos), con un procesamiento agregativo eficiente simultáneo y oportunamente sincronizado de información relativa y relacionada con la señal y su entrega en red rastreada, optimizada dinámicamente en una infraestructura de IP móvil y terrestre. Dichos archivos multimedia pueden ser capaces de transportar simultáneamente y extraer dinámicamente representaciones de señales tanto planas como volumétricas en alta definición (incluyendo, por ejemplo, estéreo 2.0 y sonido envolvente 5.1) y en un solo archivo de datos a una tasa de bits baja (es decir, un solo archivo de datos que tiene una cantidad de información reducida), y también puede ser capaz de extraer en tiempo real una voz cantada, con una calidad de sonido totalmente transparente y reproducible de la fuente de sonido original (es decir, con alta calidad de información). Dichas capacidades pueden reducir significativamente los gastos relacionados con el almacenamiento digital, la difusión, la transmisión y el ancho de banda de los sistemas de centralidad mediática y redes, al mismo tiempo que se abren nuevas fuentes de ingresos y aumenta la monetización del contenido. A este respecto, dichos medios pueden configurarse para extraer una huella digital de sonido única de la señal analógica que, junto con el adorno de datos inaudible, puede proporcionar un método altamente eficaz y seguro de identificación y trazabilidad de medios digitales, sin tener en cuenta o dependiendo de las transformaciones de formato y métodos de entrega. Dichos medios también pueden configurarse para incluir un canal de datos dedicado integrado y mezclado con la señal armónica para reflejar dinámicamente y mostrar información interestructural de la señal en el dominio del tiempo o dominio de tiempo-frecuencia, permitiendo de este modo la reproducción sincronizada de audio con cualquier evento exterior o externo (en donde un evento de este tipo puede incluir, por ejemplo, vídeo o una secuencia de programación en un juego o película, letras de canciones, portadas, credenciales del artista y el compositor o partituras, diversos efectos especiales visuales y pirotécnicos que incluyen iluminación interior y exterior o audiolibros visuales, publicidad dirigida o cualquier otra información relacionada con los medios de sonido reproducidos) mediante una interacción de datos bidireccional. El canal de datos aumenta sustancialmente la informatividad (es decir, el valor de la información) y la personalización de los medios recibidos, mejorando de este modo sustancialmente el valor de entretenimiento y la monetización de los mismos.

Además de los métodos descritos en el presente documento, un experto en la técnica también apreciará que los métodos descritos pueden incorporarse en un ejecutable mediante dispositivos y sistemas apropiados tales como, por ejemplo, diversos dispositivos capaces de cooperar a través de los procesadores correspondientes u otros dispositivos informáticos y de comunicación, en donde dichos dispositivos pueden asociarse con disposiciones particulares de memoria o almacenamiento cuando sea necesario, se desee y/o sea apropiado. Un experto en la técnica apreciará además que los métodos descritos pueden incorporarse en al menos un medio de almacenamiento legible por ordenador que tiene partes de código de programa legibles por ordenador almacenadas en el mismo que, en respuesta a la ejecución por un procesador, hacen que un aparato realice al menos los métodos descritos. Por lo tanto, se entenderá que las descripciones no deben limitarse a los aspectos específicos descritos y que se pretende que las modificaciones y otros aspectos se incluyan dentro del alcance de las reivindicaciones adjuntas. Aunque se emplean términos específicos en el presente, se usan solo en un sentido genérico y descriptivo y no con fines limitativos.

REIVINDICACIONES

1. Un método para estructurar información que comprende información digital, comprendiendo dicho método:
 - 5 analizar un archivo de información digital original para determinar cuántos elementos de datos se incluyen en el archivo de información digital, y un valor de información original asociado con cada elemento de datos de este tipo, y por lo tanto una cantidad de información original y un valor de información original asociados con el archivo de información digital original;
 - 10 aplicar un proceso de manipulación inicial al archivo de información digital original para formar un primer archivo de información digital resultante, y aplicar un proceso de manipulación posterior al primer archivo de información digital resultante para formar un segundo archivo de información digital resultante, estando cada proceso de manipulación configurado para, al menos uno de, eliminar al menos un elemento del archivo de información digital procesado, y representar una combinación de una pluralidad de elementos del archivo de información digital procesado con un elemento representativo y una primera indicación asociada con una interrelación entre el elemento representativo y al menos uno de la pluralidad de elementos en la combinación, para reducir la cantidad de información de, y para estructurar el archivo de información digital procesado, estando el al menos uno del elemento eliminado y el elemento representativo determinados para reducir el valor de información del archivo de información digital procesado en no más que un umbral seleccionado; y
 - 20 aplicar sucesivamente procesos de manipulación al archivo de información digital previamente resultante hasta que las aplicaciones de proceso de manipulación sucesivas no logren una reducción de umbral en la cantidad de información en el archivo de información digital resultante posterior sobre el archivo de información digital previamente resultante, teniendo de este modo el último archivo de información digital resultante un estructura primaria con una cantidad de información reducida con respecto a la cantidad de información original y un valor de información resultante dentro del umbral seleccionado del valor de información original.

2. Un método de acuerdo con la reivindicación 1, en el que la aplicación sucesiva de procesos de manipulación comprende además la aplicación sucesiva de procesos de manipulación al archivo de información
 - 30 digital previamente resultante para formar el último archivo de información digital estructurada resultante configurado de acuerdo con una jerarquía de información que tiene una pluralidad de niveles de información.

3. Un método de acuerdo con la reivindicación 2, en el que un nivel de información superior de la jerarquía de información incluye una cantidad de información más pequeña que un nivel de información inferior, y en el que el
 - 35 método comprende además uno de:
 - formar el nivel de información superior a partir de una estructura del nivel de información inferior y los parámetros caóticos asociados con el mismo;
 - 40 formar el nivel de información superior de modo que el nivel de información superior incluya un valor de información superior que el valor de información original; y
 - formar una pluralidad de niveles de información inferiores a partir de una estructura del nivel de información superior, y seleccionar uno de la pluralidad de niveles de información inferiores para la asociación con el nivel de información superior.

4. Un método de acuerdo con la reivindicación 1, que comprende además aplicar simultáneamente una pluralidad de procesos de manipulación a un archivo de información digital anterior para formar un archivo de información digital estructurado secundario, en particular que comprende analizar los archivos de información digital
 - 45 estructurados primarios y secundarios para determinar al menos una de una interrelación entre los niveles de información dentro de uno de los archivos de información digital estructurados primarios y secundarios, y una interrelación entre los archivos de información digital estructurados primarios y secundarios.

5. Un método de acuerdo con la reivindicación 1, que comprende además aplicar al menos uno de una pluralidad de procesos de manipulación a un archivo de información digital anterior, estando los procesos de manipulación seleccionados del grupo que consiste en un proceso de manipulación de codificación sin fases, un
 - 55 proceso de manipulación de análisis estructurado para estructuras repetitivas, un proceso de manipulación de codificación de repetición sin fases, un método de manipulación de compresión de entropía sin pérdidas, un proceso de manipulación de acelerador sin pérdidas para formar una capa de compresión adicional para un archivo de datos de audio previamente reducido, un proceso de manipulación de acelerador para facilitar la transmisión eficiente de datos de audio, y un proceso de manipulación de codificación de transformada afín.

6. Un método de acuerdo con la reivindicación 1, en el que aplicar sucesivamente procesos de manipulación para obtener un valor de información resultante dentro del umbral seleccionado del valor de información original comprende además aplicar sucesivamente procesos de manipulación para obtener un valor de información
 5 resultante dentro del umbral seleccionado, que comprende un límite de percepción humana, del valor de información original para preservar la fidelidad del archivo de información digital estructurado primario.

7. Un método de acuerdo con la reivindicación 1, en el que la formación de los archivos de información digital resultantes, comprende además formar los archivos de información digital resultantes de tal forma que cada archivo
 10 de información digital resultante comprende un conjunto de objetos y agrupaciones de objetos, teniendo cada objeto y agrupación de objetos una definición respectiva.

8. Un método de acuerdo con la reivindicación 1, en el que aplicar un proceso de manipulación para, al menos uno de, eliminar al menos un elemento del archivo de información digital procesada, y representar una combinación
 15 de una pluralidad de elementos del archivo de información digital procesada con un elemento representativo, comprende además la aplicación de un proceso de manipulación para, al menos uno de, eliminar al menos un elemento del archivo de información digital procesado, y representar una combinación de una pluralidad de elementos del archivo de información digital procesada con un elemento representativo, una primera indicación asociada con un interrelación entre el elemento representativo y al menos uno de la pluralidad de elementos en la
 20 combinación, y una segunda indicación asociada con una interrelación entre al menos dos de la pluralidad de elementos en la combinación, en particular en donde la formación de los archivos de información digital resultantes comprende además formar los archivos de información digital resultantes de tal forma que los objetos y las agrupaciones de objetos en los mismos comprenden al menos uno de los elementos representativos, y de tal forma que la definición asociada con cada objeto y agrupación de objetos comprenda la primera indicación de interrelación
 25 y, opcionalmente, la segunda indicación de interrelación, asociadas con el al menos uno de los elementos representativos, en particular comprendiendo además uno de:

incluyendo las definiciones con el conjunto correspondiente de objetos y agrupaciones de objetos en cada
 30 archivo de información digital resultante; y
 mantener las definiciones asociadas con el conjunto de objetos y agrupaciones de objetos por separado de cada archivo de información digital resultante correspondiente.

9. Un método de acuerdo con la reivindicación 1, en el que la formación de los archivos de información digital resultantes, comprende además formar los archivos de información digital resultantes de manera que cada archivo
 35 de información digital resultante posterior comprenda cualquier patrón reconocido de objetos y agrupaciones de objetos dentro del archivo de información digital resultante anterior.

10. Un método de acuerdo con la reivindicación 1, en el que aplicar sucesivamente procesos de manipulación hasta que las aplicaciones de proceso de manipulación sucesivas no logren una reducción de umbral en la cantidad
 40 de información en el archivo de información digital resultante posterior, comprende además aplicar sucesivamente procesos de manipulación hasta que las aplicaciones de proceso de manipulación sucesivas no logren una reducción de umbral en la cantidad de información en el archivo de información digital resultante posterior, de tal forma que el último archivo de información digital resultante tenga una estructura de primotrópico de información que comprenda un conjunto de objetos primarios, agrupaciones de objetos primarios, y cualquier interrelación asociada
 45 los mismos, en particular que comprenda analizar cada archivo de información digital resultante y el último archivo de información digital resultante para determinar las correlaciones estadísticas entre los objetos posteriores en un archivo de información digital posterior y las combinaciones de objetos anteriores en un archivo de información digital anterior y, para cualquier correlación estadística que cumpla al menos un umbral de correspondencia, almacenar una indicación de correlación del objeto posterior correspondiente y una combinación anterior de objetos
 50 para el análisis de un archivo de información digital original posterior.

11. Un método de acuerdo con la reivindicación 10, que comprende además uno de:
 55 analizar cada archivo de información digital resultante sucesivo para determinar al menos una correlación objetiva entre una combinación de objetos anteriores, y al menos una relación entre objetos entre los objetos anteriores en la combinación, en un archivo de información digital anterior y asociar la al menos una correlación objetiva y la al menos una relación entre objetos con un objeto posterior correspondiente en un archivo de información digital posterior, la al menos una correlación objetiva y la al menos una relación entre objetos asociada con el objeto posterior correspondiente, sobre una pluralidad de objetos posteriores,

formando colectivamente un conjunto de correlaciones de objetos representativas de las interrelaciones asociadas con el conjunto de objetos primarios y agrupaciones de objetos primarios en la estructura de primotrón de información; y

5 analizar cada archivo de información digital resultante sucesivo para determinar al menos una correlación objetiva entre una combinación de objetos anteriores en un archivo de información digital anterior, y asignar al menos una relación entre objetos entre los objetos anteriores en la combinación, y asociar la al menos una correlación objetiva y la al menos una relación entre objetos con un objeto posterior correspondiente en un archivo de información digital posterior, la al menos una correlación objetiva y la al menos una relación entre objetos asociada con el objeto posterior correspondiente, sobre una pluralidad de objetos posteriores,
10 formando colectivamente un conjunto de correlaciones de objetos representativas de las interrelaciones asociadas con el conjunto de objetos primarios y agrupaciones de objetos primarios en la estructura de primotrón de información.

12. Un método de acuerdo con la reivindicación 11, que comprende además interpretar el conjunto de objetos primarios y la agrupación de objetos primarios en la estructura de primotrón de información mediante la evaluación estadística de diversas combinaciones de correlaciones de objetos en el conjunto de correlaciones de objetos y, para cualquier combinación de correlaciones de objetos, cumplir al menos con un umbral estadístico, asociar las combinaciones con los objetos primarios y agrupaciones de objetos primarios como una interpretación contextual del conjunto de objetos primarios y agrupaciones de objetos primarios en la estructura de primotrón de información, que en particular comprende además sintetizar un archivo de información digital representativo asociado con el archivo de información digital original aplicando la interpretación contextual al conjunto de objetos primarios y agrupaciones de objetos primarios en la estructura de primotrón de información.

13. Un método de acuerdo con la reivindicación 11, que comprende, además:

25 asociar al menos un objeto, estando el al menos un objeto seleccionado de los objetos anteriores, los objetos posteriores, y los objetos primarios y las agrupaciones de objetos primarios, con una de una pluralidad de características del archivo de información digital original; y
30 determinar una disposición de la pluralidad de características con respecto a la duración del archivo de información digital original y correlacionar una distribución del al menos un objeto con respecto al mismo, en particular que comprende formar un archivo de datos de resumen que incluye una representación generalizada de la distribución de al menos un objeto asociado con cada una de la pluralidad de características a lo largo de la duración del archivo de información digital original.

35 14. Un método de acuerdo con la reivindicación 10, en el que aplicar sucesivamente procesos de manipulación hasta que las sucesivas aplicaciones de proceso de manipulación no logren una reducción de umbral en la cantidad de información en el archivo de información digital resultante posterior, comprende además la catalogación de una cantidad y un orden de los procesos de manipulación aplicados sucesivamente a los archivos de información digital resultantes posteriores para obtener el último archivo de información digital resultante que tiene la estructura de primotrón de información, estando la cantidad de procesos de manipulación asociados con los niveles correspondientes de una jerarquía de información.

15. Un método de acuerdo con la reivindicación 10, 12 y 14, que comprende además sintetizar un archivo de información digital representativo asociado con el archivo de información digital original de acuerdo con un proceso condicional que comprende:

50 aplicar cualquier indicación de correlación aplicable al conjunto de objetos primarios y agrupaciones de objetos primarios en la estructura de primotrón de información para determinar la combinación anterior correspondiente de objetos asociados con los mismos;

si no se determina ninguna indicación de correlación aplicable, aplicar la interpretación contextual al conjunto de objetos primarios y agrupaciones de objetos primarios en la estructura de primotrón de información para determinar la combinación anterior correspondiente de objetos asociados con los mismos;

y
55 si la aplicación de la interpretación contextual está por debajo de un valor de información de umbral asociado con el valor de información original, aplicar un proceso de manipulación inverso inicial al conjunto de objetos primarios y agrupaciones de objetos primarios en la estructura de primotrón de información, siendo el proceso de manipulación inverso inicial la primera manipulación en el orden inverso de los procesos de manipulación, para determinar la combinación anterior correspondiente de objetos asociados con el mismo, comprendiendo en particular además, para la combinación anterior de objetos determinada a

partir del conjunto de objetos primarios y agrupaciones de objetos primarios en la estructura de primotrón de información, aplicar al mismo, el proceso condicional que incluye cualquier indicación de correlación aplicable, la interpretación contextual correspondiente o el siguiente proceso de manipulación inversa en el orden inverso de los procesos de manipulación correspondientes a la combinación anterior de objetos, con el fin de determinar la siguiente combinación anterior de objetos correspondientes al mismo, que comprende en particular además aplicar sucesivamente el proceso condicional a la siguiente combinación anterior de objetos hasta que la combinación anterior resultante de objetos se asocie con el nivel inicial de la jerarquía de información, correspondiendo así el archivo de información digital anterior resultante al archivo de información digital representativo sintetizado.

10 16. Un método de acuerdo con la reivindicación 1, en el que el archivo de información digital original es un archivo de audio digital, y el método comprende además asociar una cantidad de información del archivo de audio digital con una tasa de bits proporcional a un tamaño y una duración de tiempo del archivo de audio digital, y asociar un valor de información del archivo de audio digital con una calidad del archivo de audio digital, siendo la calidad
15 determinada a partir de al menos uno de un valor de umbral de energía de primotrón, un criterio de información perceptiva (PIC), una estructura de sonido identificada a partir de un evento de fluctuación de caos aleatorio determinado de acuerdo con una función de distribución de probabilidad que usa ruido aleatorio como señal de entrada, un criterio de pérdida de calidad nula, y un área de superficie de una esfera n-dimensional de un radio de
20 unidad única en un dominio de frecuencia de tiempo determinado a partir de una función implícita, en particular, en donde la aplicación sucesiva de procesos de manipulación al archivo de audio digital previamente resultante para formar el último archivo de música digital estructurado resultante comprende además la aplicación sucesiva de procesos de manipulación al archivo de audio digital previamente resultante para formar el último archivo de audio digital estructurado resultante configurado de acuerdo con una jerarquía de información que tiene una pluralidad de niveles de información, en donde un nivel de información inferior comprende una representación del audio como un
25 archivo de información digital de alta resolución, y un nivel de información superior comprende una representación en ficha que incluye una puntuación de nota escrita con las letras correspondientes, teniendo así el nivel de información superior una cantidad de información más pequeña que el nivel de información inferior, y siendo de este modo el nivel de información inferior capaz de formarse a partir del nivel de información superior y los parámetros caóticos asociados con el rendimiento y la digitalización del audio de acuerdo con la representación en ficha, en particular en donde la aplicación sucesiva de procesos de manipulación al archivo de audio digital previamente resultante para formar el último archivo de música digital estructurado resultante comprende además aplicar sucesivamente procesos de manipulación al archivo de audio digital previamente resultante para formar el último archivo de audio digital estructurado resultante que comprende al menos un primotrón definido en uno de un dominio del tiempo y un dominio del tiempo-frecuencia, representando el último archivo de audio digital estructurado
30 resultante una formalización digital compacta de una señal armónica, en donde la cantidad de información asociada con la tasa de bits de audio se reduce en comparación con el archivo de audio digital correspondiente al archivo de información digital original, mientras que el valor de información asociado con la calidad de audio se mantiene dentro de un umbral del archivo de audio digital correspondiente al archivo de información digital original.

40 17. Un método de acuerdo con la reivindicación 16, en el que cada archivo de audio digital representa datos relacionados con armónicos, y el método comprende además la definición de aciertos como estructuras jerárquicas de nivel superior en el dominio del tiempo; la definición de armónicos como estructuras jerárquicas de nivel superior en el dominio de la frecuencia que tienen un origen de primotrón armónico posterior que coincide con una terminación de primotrón armónico anterior; la eliminación de la redundancia en las estructuras jerárquicas de nivel inferior de dominio del tiempo para obtener una disminución de la entropía en las estructuras jerárquicas de nivel superior; y la definición de sobretonos como estructuras jerárquicas de nivel superior de dominio de la frecuencia que tienen una frecuencia base única correspondiente a las estructuras jerárquicas de nivel inferior y una pluralidad de copias de las estructuras jerárquicas de nivel superior desplazadas una frecuencia fija de la frecuencia base única y entre sí.

50 18. Un sistema para estructurar información, que comprende al menos un dispositivo de análisis, teniendo cada uno un procesador; al menos un dispositivo de reducción de archivos digitales, teniendo cada uno un procesador; y opcionalmente un dispositivo de sintetización que tiene un procesador; estando los dispositivos configurados para cooperar para realizar al menos el método de acuerdo con una cualquiera de las reivindicaciones 1-17.

55 19. Un medio de almacenamiento legible por ordenador que tiene partes de código de programa legibles por ordenador almacenadas en el mismo que, en respuesta a la ejecución por parte de un procesador, hace que un aparato realice al menos el método de acuerdo con una cualquiera de las reivindicaciones 1-17.

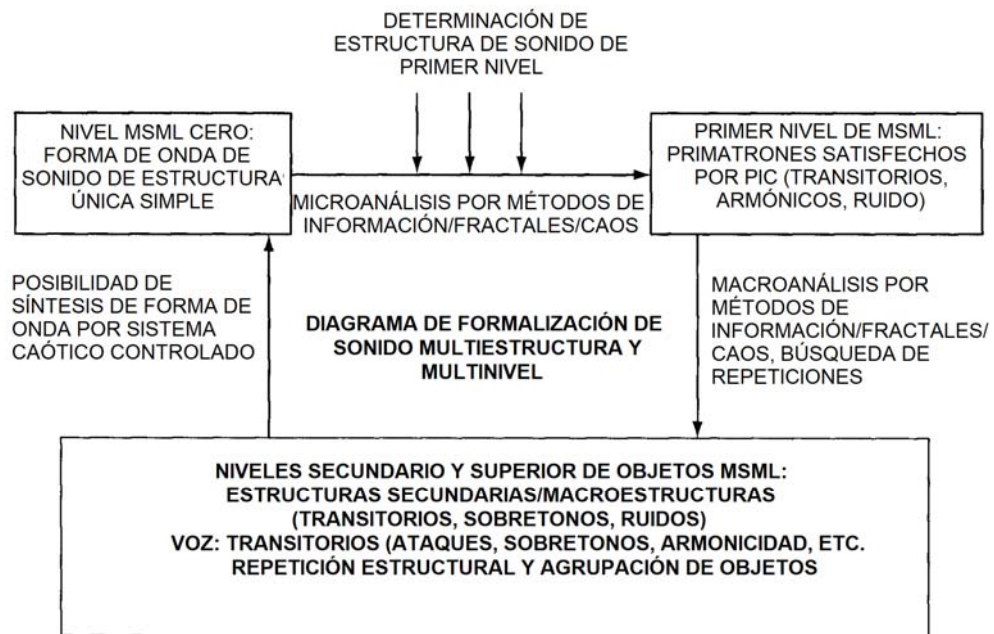
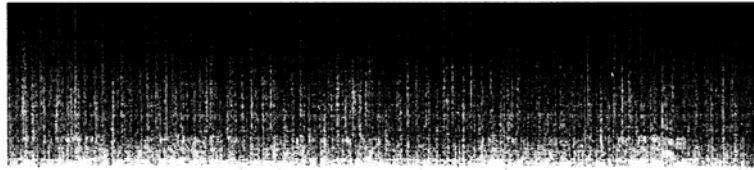


FIG. 1



SAMPLE OF STANDARD ORIGINAL SPECTRUM

FIG. 2A

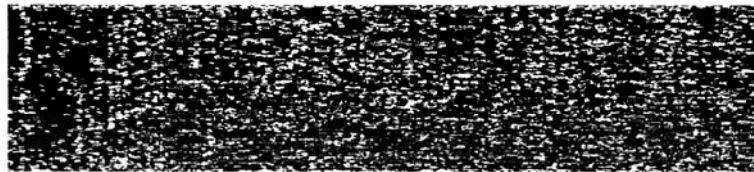


DIAGRAMA DE INFORMACIÓN DE PRIMOTRONES ARMÓNICOS PARA LA MISMA COMPOSICIÓN*

FIG. 2B



DIAGRAMA DE INFORMACIÓN DE PRIMOTRONES DE ACIERTO PARA LA MISMA COMPOSICIÓN*

FIG. 2C

*VISUALIZACIÓN DE PRIMOTRONES DE INFORMACIÓN DE SONIDO. LAS ÁREAS QUE NO ESTÁN EN NEGRO PARTICULARES SON PRIMATRONES QUE SATISFACEN LOS CRITERIOS PIC. LAS REGIONES EN NEGRO DON PERTENECEN AL CONJUNTO DE PRIMATRONES. OTRAS ÁREAS QUE NO ESTÁN EN NEGRO PUEDEN REQUERIR UN ANÁLISIS MÁS DETALLADO.

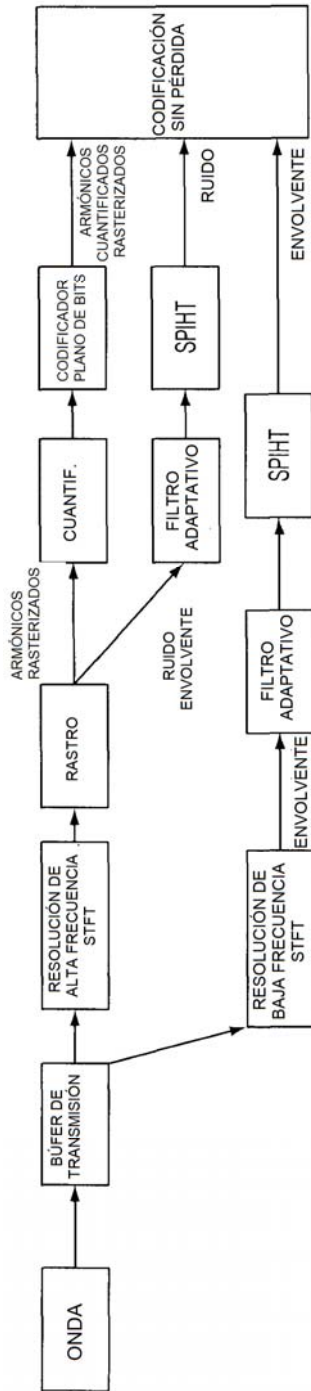


FIG. 3A

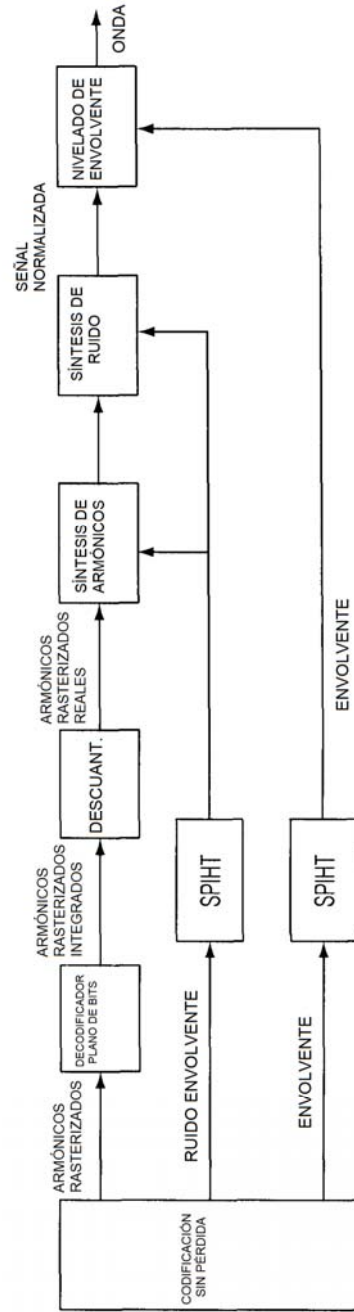


FIG. 3B

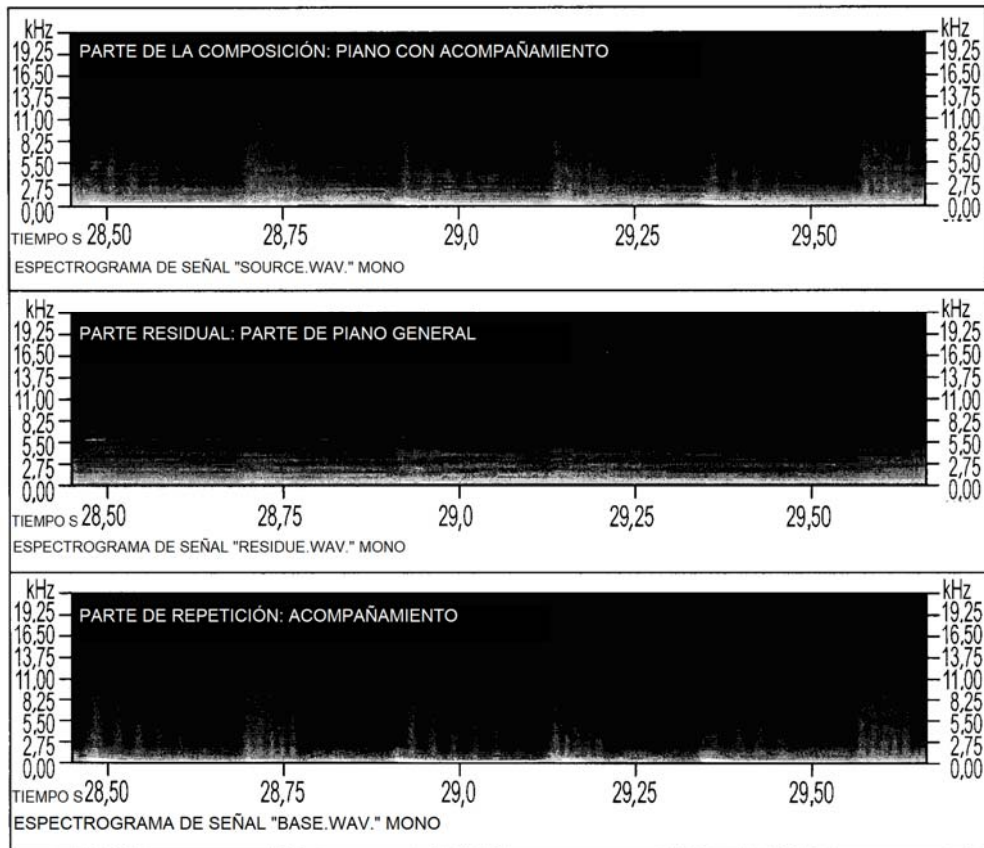


FIG. 4A

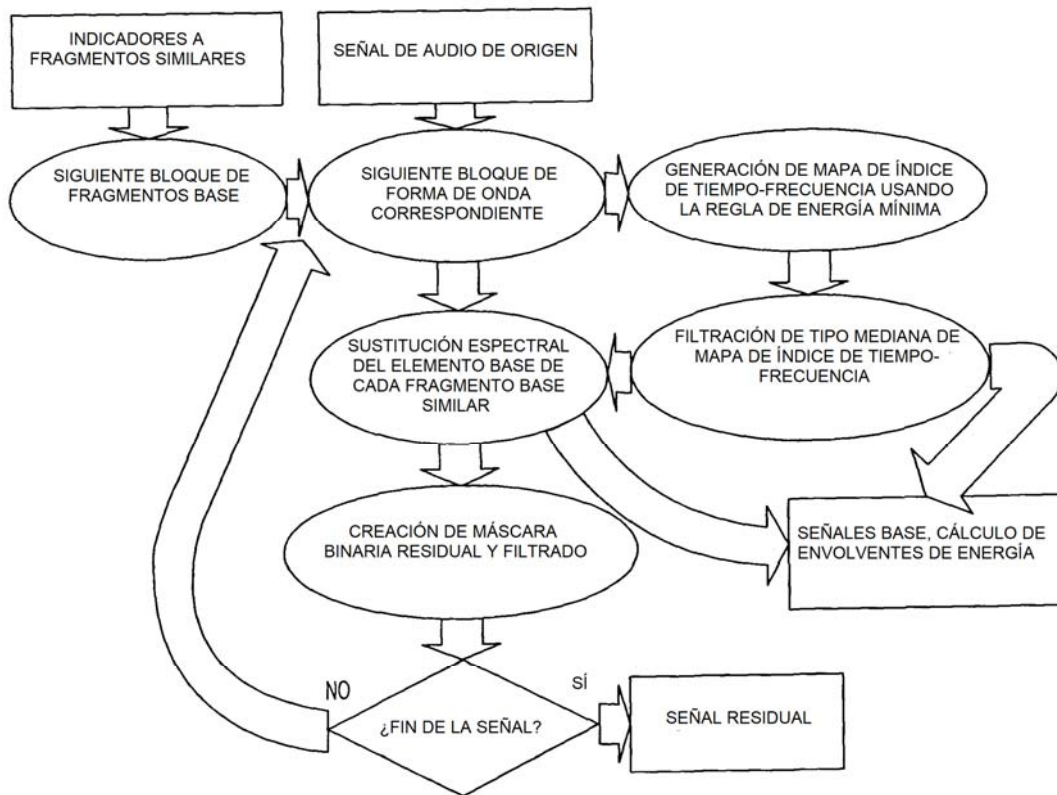


FIG. 4B

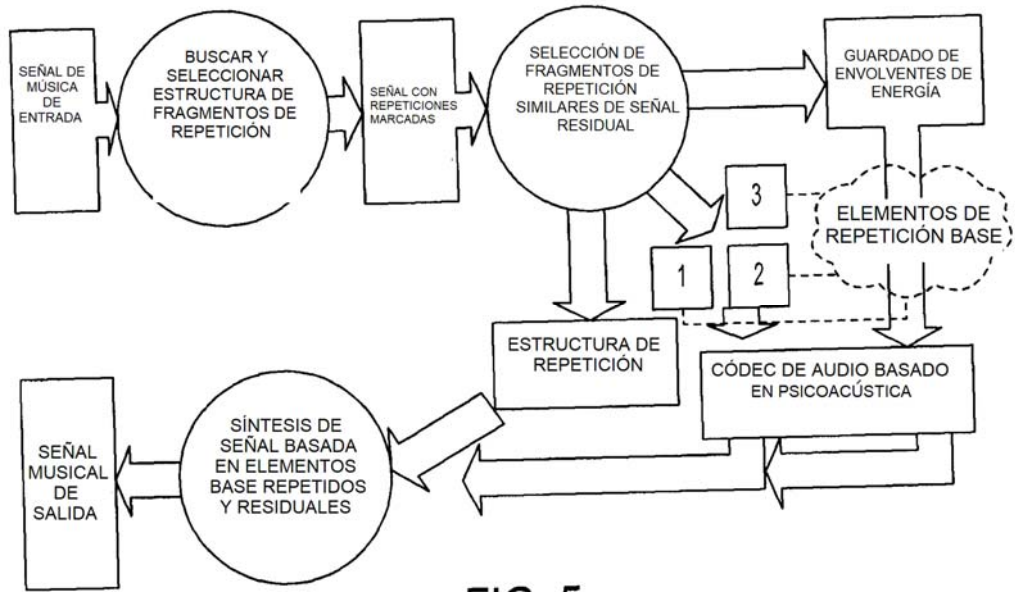


FIG. 5

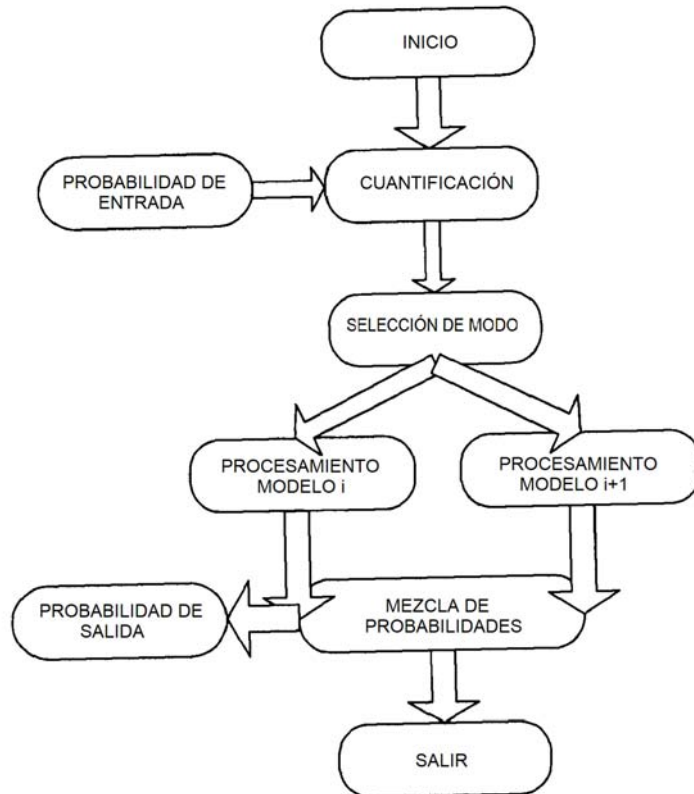


FIG. 6A

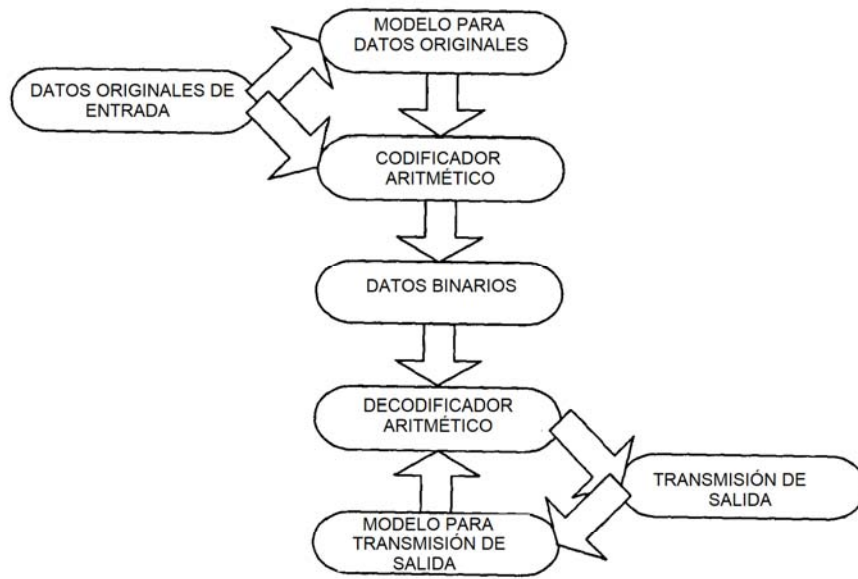


FIG. 6B

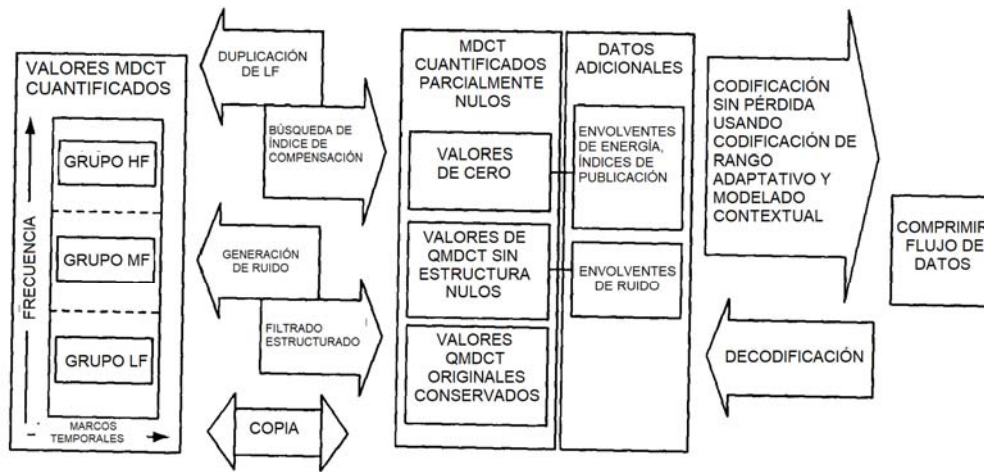


FIG. 7A

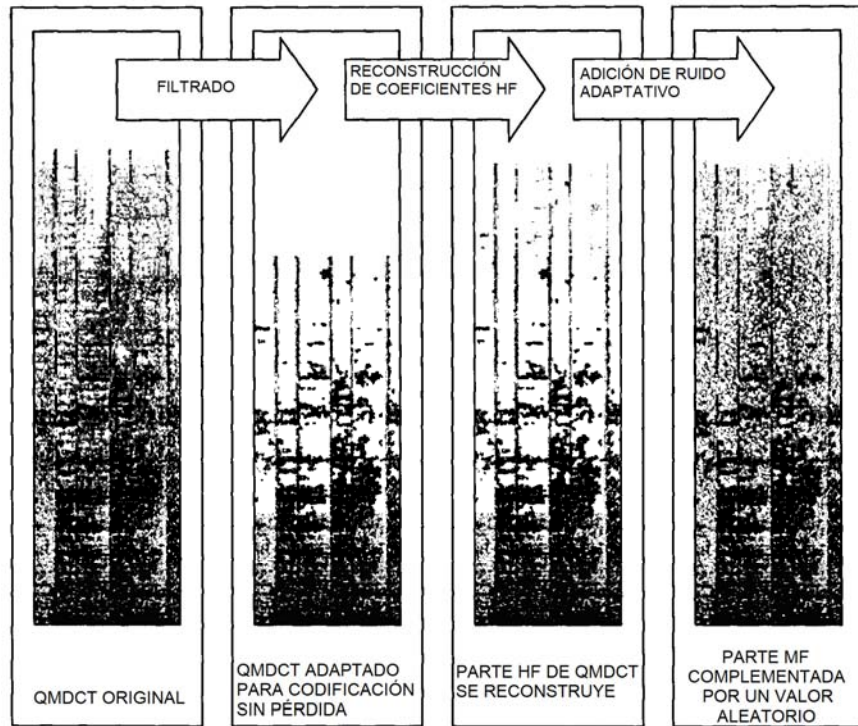


FIG. 7B

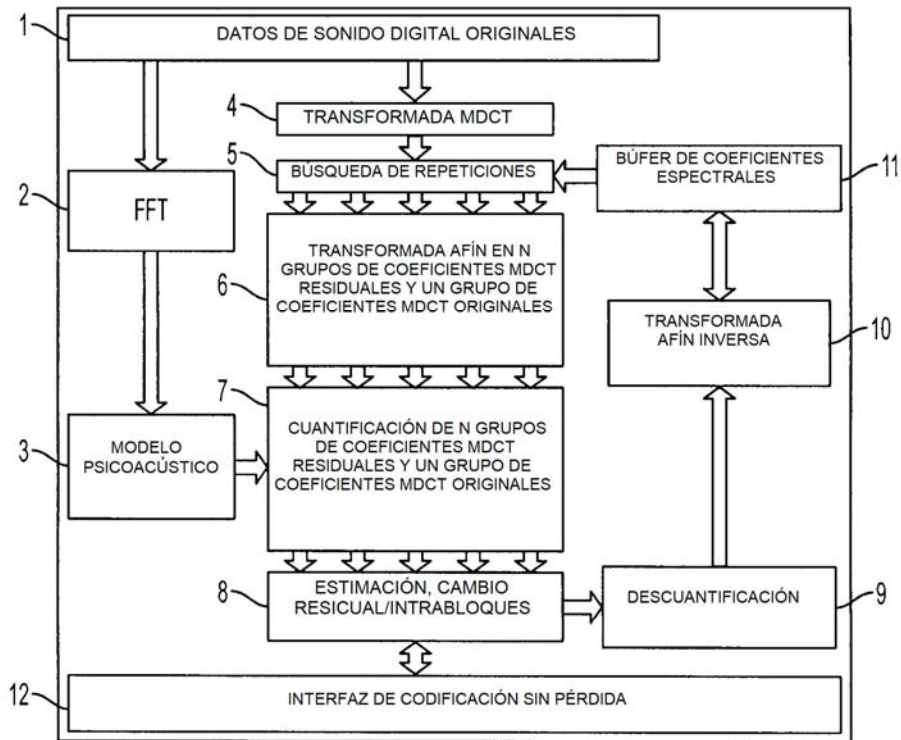


FIG. 8A

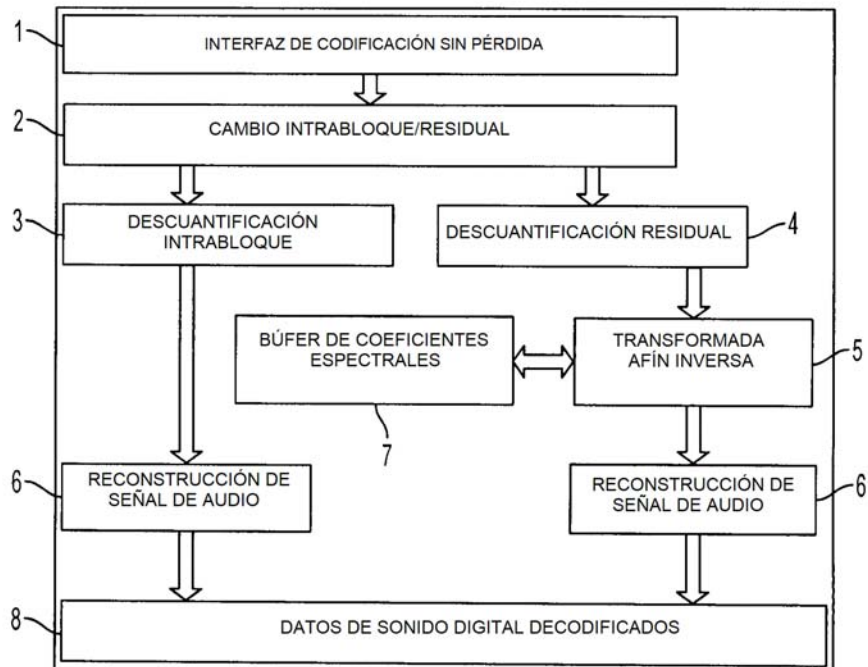


FIG. 8B

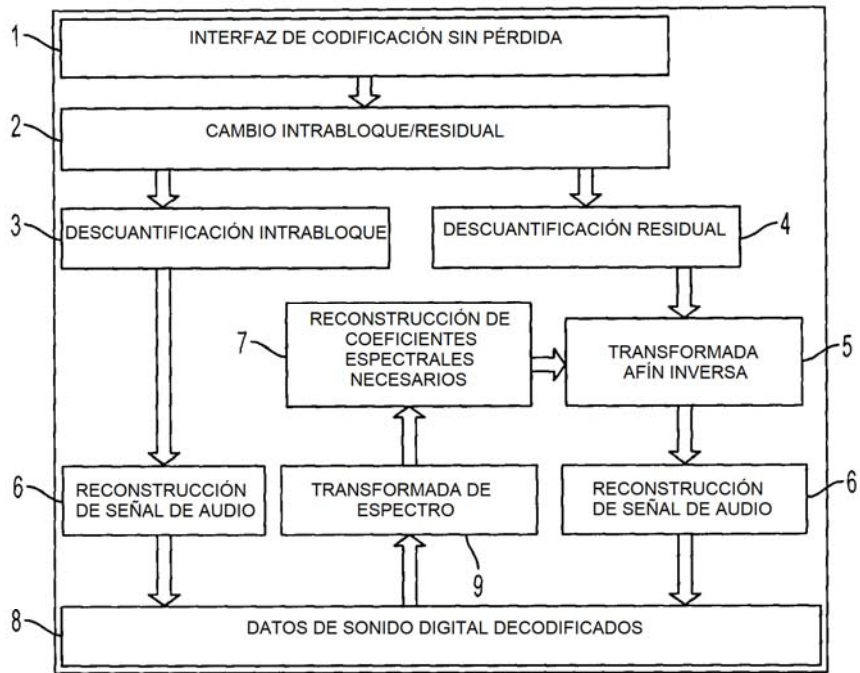


FIG. 8C

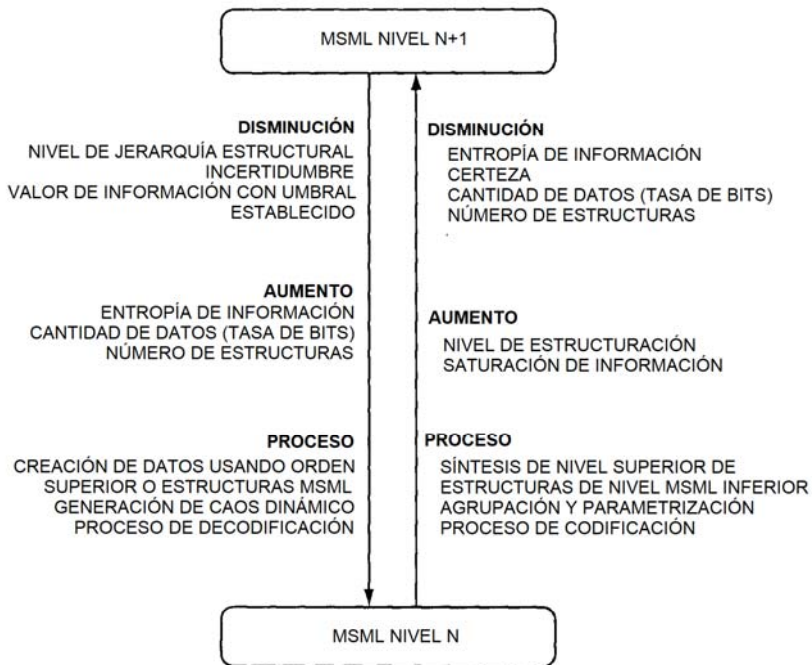


FIG. 9

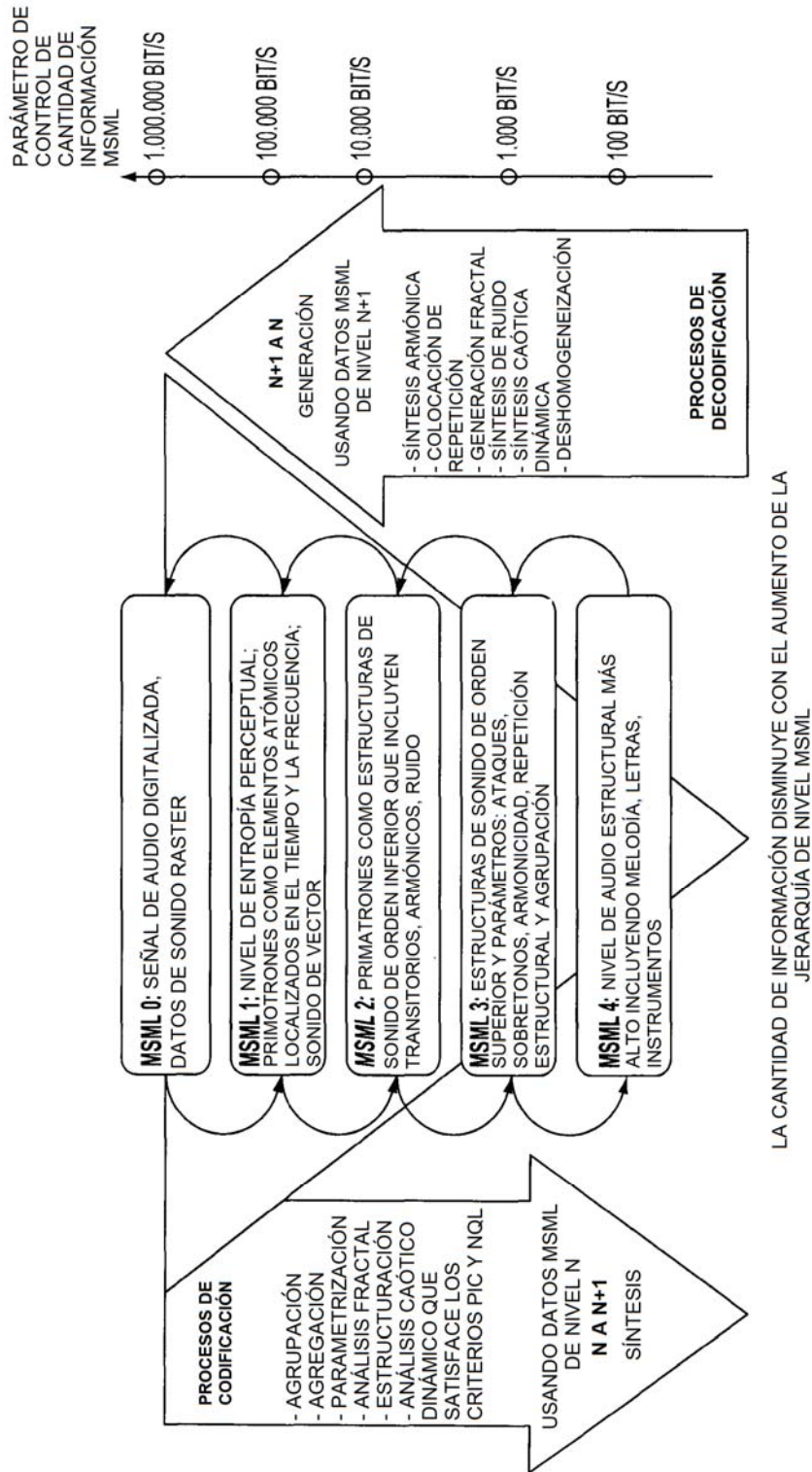


FIG. 10

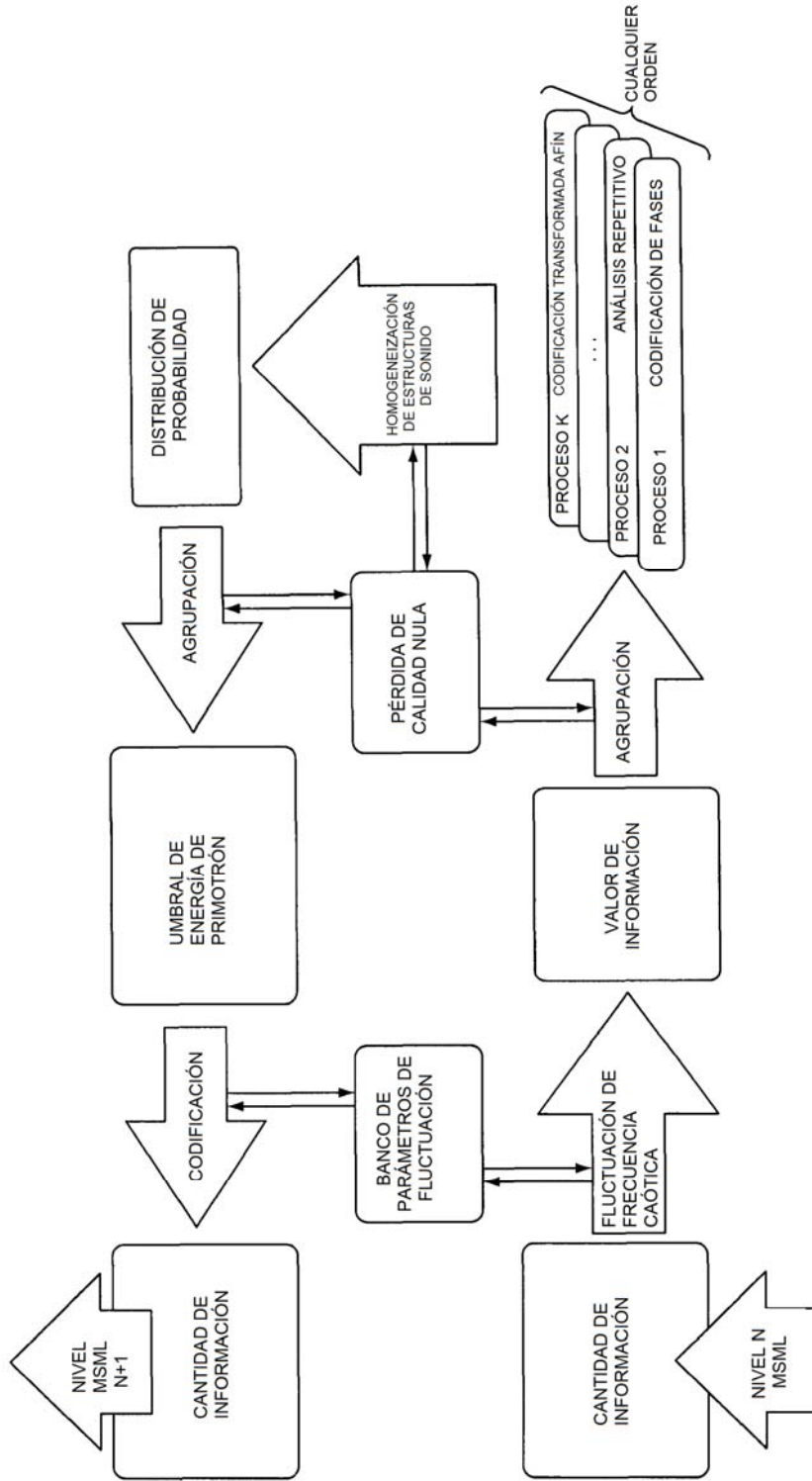


FIG. 11