

19



OFICINA ESPAÑOLA DE
PATENTES Y MARCAS

ESPAÑA



11 Número de publicación: **2 980 709**

51 Int. Cl.:

G10L 19/093 (2013.01)

G06F 30/327 (2010.01)

12

TRADUCCIÓN DE PATENTE EUROPEA

T3

96 Fecha de presentación y número de la solicitud europea: **07.01.2014** **E 21171483 (7)**

97 Fecha y número de publicación de la concesión europea: **24.04.2024** **EP 3893240**

54 Título: **Predicción basada en modelo en un banco de filtros críticamente muestreados**

30 Prioridad:

08.01.2013 US 201361750052 P

09.09.2013 US 201361875528 P

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

02.10.2024

73 Titular/es:

DOLBY INTERNATIONAL AB (100.0%)

**77 Sir John Rogerson's Quay, Block C, Grand
Canal Docklands
Dublin, D02 VK60, IE**

72 Inventor/es:

VILLEMOES, LARS

74 Agente/Representante:

LINAGE GONZÁLEZ, Rafael

ES 2 980 709 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

Predicción basada en modelo en un banco de filtros críticamente muestreados

5 Referencia cruzada a solicitudes relacionadas

Esta solicitud es una solicitud divisional europea de la solicitud de patente europea EP19208681.7 (referencia: D12110EP03), para la que el formulario 1001 de la OEP fue presentado el 12 de noviembre de 2019.

10 Campo técnico

El presente documento se refiere a sistemas de codificación de fuentes de audio. En particular, el presente documento se refiere a sistemas de codificación de fuentes de audio que usan predicción lineal en combinación con un banco de filtros.

15 Antecedentes

Hay dos herramientas importantes de procesamiento de señales aplicadas en sistemas para la codificación de fuentes de señales de audio, en concreto los bancos de filtros críticamente muestreados y la predicción lineal. Los bancos de filtros críticamente muestreados (por ejemplo, bancos de filtros basados en la transformada discreta de coseno modificada, MDCT) permiten un acceso directo a representaciones de tiempo-frecuencia en las que puede aprovecharse la irrelevancia perceptual y la redundancia de las señales. La predicción lineal permite modelar de manera eficaz fuentes de señales de audio, en particular de señales de voz. La combinación de las dos herramientas, es decir, el uso de la predicción en las subbandas de un banco de filtros, se ha usado principalmente en la codificación de audio de alta velocidad binaria. En lo que respecta a la codificación de baja velocidad binaria, un desafío que plantea la predicción en las subbandas es mantener bajo el coste (es decir, la velocidad binaria) para la descripción de los predictores. Otro desafío es controlar la conformación de ruido resultante de la señal de error de predicción obtenida por un predictor de subbanda.

En cuanto al reto que supone codificar la descripción del predictor de subbanda de una manera eficaz en cuanto a los bits, un posible enfoque es estimar el predictor a partir de partes ya descodificadas de la señal de audio y, por tanto, evitar completamente el coste de una descripción de predictor. Si el predictor puede determinarse a partir de partes ya descodificadas de la señal de audio, el predictor puede determinarse en el codificador y en el descodificador sin necesidad de transmitir una descripción de predictor desde el codificador al descodificador. Este esquema se denomina esquema de predicción adaptativa hacia atrás. Sin embargo, el esquema de predicción adaptativa hacia atrás normalmente se degrada considerablemente cuando la velocidad binaria de la señal de audio codificada disminuye. Un enfoque alternativo o adicional a la codificación eficiente de un predictor de subbanda es identificar una descripción de predictor más natural, por ejemplo una descripción que aproveche la estructura intrínseca de la señal de audio que va a codificarse. Por ejemplo, la codificación de voz de baja velocidad binaria aplica normalmente un esquema adaptativo hacia delante basado en una representación compacta de un predictor a corto plazo (que utiliza correlaciones a corto plazo) y de un predictor a largo plazo (que utiliza correlaciones a largo plazo debido al tono subyacente de la señal de voz).

En cuanto al reto que supone controlar la conformación de ruido de la señal de error de predicción, se observa que aunque la conformación de ruido de un predictor puede controlarse correctamente dentro de una subbanda, la señal de audio final de salida del codificador presenta normalmente artefactos de distorsión (excepto en señales de audio que presentan una forma de ruido espectral sustancialmente plana).

Un caso importante de un predictor de subbanda es la implementación de una predicción a largo plazo en un banco de filtros con ventanas solapadas. Un predictor a largo plazo utiliza normalmente las redundancias de las señales de audio periódicas y casi periódicas (tales como señales de voz que presentan un tono intrínseco), y puede describirse con un solo o con un número bajo de parámetros de predicción. El predictor a largo plazo puede definirse en el tiempo continuo mediante un retardo que refleja la periodicidad de la señal de audio. Cuando este retardo es grande en comparación con la longitud de la ventana del banco de filtros, el predictor a largo plazo puede implementarse en el dominio de tiempo discreto mediante un desplazamiento o un retardo fraccionario y puede volver a convertirse en un predictor causal en el dominio de subbanda. Tal predictor a largo plazo no presenta normalmente artefactos de distorsión, pero tiene como desventaja una gran complejidad computacional debido a la necesidad de operaciones adicionales de banco de filtros para la conversión desde el dominio de tiempo al dominio de subbanda. Por lo tanto, el enfoque para determinar el retardo en el dominio de tiempo y para convertir el retardo en un predictor de subbanda no puede aplicarse cuando el periodo de la señal de audio que va a codificarse es comparable o inferior al tamaño de ventana del banco de filtros.

El presente documento aborda los inconvenientes de predicción de subbanda mencionados anteriormente. En particular, el presente documento describe procedimientos y sistemas que permiten una descripción eficaz de velocidad binaria de predictores de subbanda y/o que permiten una reducción de los artefactos de distorsión provocados por los predictores de subbanda. En particular, el procedimiento y los sistemas descritos en el

presente documento permiten la implementación de codificadores de audio de baja velocidad binaria usando predicción de subbanda, lo que permite reducir el nivel de los artefactos de distorsión. El documento US 2006/0015239 A1 describe la codificación de fuentes de predicción de subbanda.

5 Compendio

El presente documento describe procedimientos y sistemas que mejoran la calidad de la codificación de fuentes de audio utilizando predicción en el dominio de subbanda De un banco de filtros críticamente muestreado. Los métodos y sistemas pueden hacer uso de una descripción compacta de los predictores de subbanda, en donde la descripción se basa en los modelos de señal. Además o como alternativa, los métodos y sistemas pueden hacer uso de una implementación eficiente de los predictores directamente en el dominio de subbanda. Además o como alternativa, los métodos y sistemas pueden hacer uso de términos cruzados de predictor de subbanda, tal como se describe en el presente documento, para permitir una reducción de los artefactos de distorsión.

Como se describe en el presente documento, la descripción compacta de predictores de subbanda puede comprender la frecuencia de una senoide, el periodo de una señal periódica, un espectro ligeramente inarmónico, como el que aparece en la vibración de una cuerda tensa, y/o una multitud de tonos de una señal polifónica. Se sabe que en el caso de un predictor a largo plazo, el modelo de señal periódica proporciona predictores causales de alta calidad para diversos parámetros de demora (o retardos) que incluyen valores que son menores y/o mayores que el tamaño de ventana del banco de filtros. Esto significa que puede usarse un modelo de señal periódica para implementar un predictor de subbanda a largo plazo de manera eficiente. La transición desde la predicción basada en modelos sinusoidales a la aproximación de un retardo arbitrario no presenta discontinuidades.

La implementación directa de predictores en el dominio de subbanda permite acceder de manera explícita a características perceptuales de las distorsiones de cuantificación producidas. Además, la implementación de predictores en el dominio de subbanda permite acceder a propiedades numéricas, tales como la ganancia de predicción y la dependencia de los predictores con respecto a los parámetros. Por ejemplo, un análisis basado en un modelo de señal puede revelar que la ganancia de predicción solo es significativa en un subconjunto de las subbandas consideradas, y la variación de los coeficientes de los predictores en función del parámetro elegido para la transmisión puede resultar útil en el diseño de los formatos de los parámetros, así como de algoritmos de codificación eficientes. Además, la complejidad computacional puede reducirse considerablemente en comparación con implementaciones de predictores que se basan en el uso de algoritmos que funcionan tanto en el dominio del tiempo como en el dominio de subbanda. En particular, los procedimientos y sistemas descritos en el presente documento pueden usarse para implementar la predicción de subbanda directamente en el dominio de subbanda sin necesidad de determinar y aplicar un predictor (por ejemplo, un retardo a largo plazo) en el dominio del tiempo.

El uso de términos de subbanda cruzada en los predictores de subbanda permite mejorar de manera considerable las propiedades de conformación de ruido en el dominio de la frecuencia en comparación con los predictores dentro de banda (que solo se basan en la predicción dentro de banda). De esta forma pueden reducirse los artefactos de distorsión, permitiendo así el uso de la predicción de subbanda en sistemas de codificación de audio de velocidad binaria relativamente baja.

Según un aspecto, se describe un procedimiento para estimar una primera muestra de una primera subbanda de una señal de audio. La primera subbanda de la señal de audio puede haberse determinado usando un banco de filtros de análisis que comprende una pluralidad de filtros de análisis que proporcionan una pluralidad de señales de subbanda en una pluralidad de subbandas, respectivamente, de la señal de audio. La señal de audio del dominio del tiempo puede enviarse a un banco de filtros de análisis, obteniéndose así una pluralidad de señales de subbanda en una pluralidad de subbandas. Cada una de la pluralidad de subbandas cubre normalmente una gama de frecuencias diferente de la señal de audio, proporcionándose así acceso a diferentes componentes de frecuencia de la señal de audio. La pluralidad de subbandas puede tener una separación entre subbandas idéntica o uniforme. La primera subbanda corresponde a una de la pluralidad de subbandas proporcionadas por el banco de filtros de análisis.

El banco de filtros de análisis puede tener varias propiedades. Un banco de filtros de síntesis que comprende una pluralidad de filtros de síntesis puede tener propiedades idénticas o similares. Las propiedades descritas para el banco de filtros de análisis y los filtros de análisis también pueden aplicarse a las propiedades del banco de filtros de síntesis y de los filtros de síntesis. Normalmente, la combinación de un banco de filtros de análisis y de un banco de filtros de síntesis permite una reconstrucción perfecta de la señal de audio. Los filtros de análisis del banco de filtros de análisis pueden ser invariantes al desplazamiento entre sí. Además o como alternativa, los filtros de análisis del banco de filtros de análisis pueden comprender una función de ventana común. En particular, los filtros de análisis del banco de filtros de análisis pueden comprender versiones moduladas de diferente manera de la función de ventana común. En una realización, la función de ventana común se modula usando la función coseno, obteniéndose así un banco de filtros de análisis modulado por coseno. En particular, el banco de filtros de análisis puede comprender (o puede corresponder a) uno o más de las siguientes

transformadas: MDCT, QMF y/o ELT. La función de ventana común puede tener una duración K finita. La duración de la función de ventana común puede ser tal que muestras subsiguientes de una señal de subbanda se determinan usando segmentos solapados de la señal de audio de dominio de tiempo. De este modo, el banco de filtros de análisis puede comprender una transformada solapada. Los filtros de análisis del banco de filtros de análisis pueden formar una base ortogonal y/o una base ortonormal. Una propiedad adicional es que el banco de filtros de análisis puede corresponder a un banco de filtros críticamente muestreado. En particular, el número de muestras de la pluralidad de señales de subbanda puede corresponder al número de muestras de la señal de audio del dominio del tiempo.

El método puede comprender determinar un parámetro de modelo de un modelo de señal. Debe observarse que el modelo de señal puede describirse usando una pluralidad de parámetros de modelo. De este modo, el procedimiento puede comprender determinar la pluralidad de parámetros de modelo del modelo de señal. El parámetro o parámetros de modelo pueden extraerse de un flujo de bits recibido que comprende o que indica el parámetro de modelo y una señal de error de predicción. Como alternativa, el parámetro o parámetros de modelo pueden determinarse ajustando el modelo de señal a la señal de audio (por ejemplo, en cada trama), por ejemplo usando un enfoque de error cuadrático medio.

El modelo de señal puede comprender una o más componentes de modelo sinusoidales. En tal caso, el parámetro de modelo puede indicar la una o más frecuencias de la una o más componentes de modelo sinusoidales. A modo de ejemplo, el parámetro de modelo puede indicar una frecuencia fundamental Ω de un modelo de señal multisinusoidal, donde la señal multisinusoidal comprende componentes de modelo sinusoidales a frecuencias que corresponden a múltiplos $q\Omega$ de la frecuencia fundamental Ω . De este modo, el modelo de señal multisinusoidal puede comprender una componente de señal periódica, donde la componente de señal periódica comprende una pluralidad de componentes sinusoidales y donde la pluralidad de componentes sinusoidales tienen una frecuencia que es un múltiplo de la frecuencia fundamental Ω . Como se mostrará en el presente documento, tal componente de señal periódica puede usarse para modelar un retardo en el dominio de tiempo (como se usa, por ejemplo, en predictores a largo plazo). El modelo de señal puede comprender uno o más parámetros de modelo que indican un desplazamiento y/o una desviación del modelo de señal con respecto a un modelo de señal periódico. El desplazamiento y/o desviación pueden indicar una desviación de las frecuencias de la pluralidad de componentes sinusoidales del modelo de señal periódico con respecto a múltiplos respectivos $q\Omega$ de la frecuencia fundamental Ω .

El modelo de señal puede comprender una pluralidad de componentes de señal periódica. Cada una de las componentes de señal periódica puede describirse usando uno o más parámetros de modelo. Los parámetros de modelo pueden indicar una pluralidad de frecuencias fundamentales $\Omega_0, \Omega_1, \dots, \Omega_{M-1}$ de la pluralidad de componentes de señal periódica. Además o como alternativa, el modelo de señal puede describirse mediante un parámetro de relajación predeterminado y/o ajustable (que puede ser uno de los parámetros de modelo). El parámetro de relajación puede configurarse para allanar o suavizar el espectro lineal de una componente de señal periódica. Ejemplos específicos de modelos de señal y de parámetros de modelo asociados se describen en la sección del presente documento que describe las formas de realización.

El parámetro o parámetros de modelo pueden determinarse de manera que se reduzca (por ejemplo, se minimice) el valor medio de una señal de error cuadrático de predicción. La señal de error de predicción puede determinarse en función de la diferencia entre la primera muestra y la estimación de la primera muestra. En particular, el valor medio de la señal de error cuadrático de predicción puede determinarse en función de una pluralidad de primeras muestras subsiguientes de la primera señal de subbanda y en función de una pluralidad correspondiente de primeras muestras estimadas. En particular, en el presente documento se propone modelar la señal de audio o, al menos, la primera señal de subbanda de la señal de audio usando un modelo de señal que se describe mediante uno o más parámetros de modelo. Los parámetros de modelo se usan para determinar el uno o más coeficientes de predicción de un predictor lineal que determina una primera señal de subbanda estimada. La diferencia entre la primera señal de subbanda y la primera señal de subbanda estimada proporciona una señal de subbanda de error de predicción. El uno o más parámetros de modelo pueden determinarse de manera que se reduzca (por ejemplo, se minimice) el valor medio de la señal de subbanda de error cuadrático de predicción.

El método puede comprender además determinar un coeficiente de predicción que se aplicará a una muestra anterior de una primera señal de subbanda descodificada obtenida a partir de la primera señal de subbanda. En particular, la muestra anterior puede determinarse añadiendo una versión cuantificada de la señal de error de predicción a una muestra correspondiente de la primera señal de subbanda. La primera señal de subbanda descodificada puede ser idéntica a la primera señal de subbanda (por ejemplo, en caso de un codificador sin pérdidas). Un intervalo de tiempo de la muestra previa es normalmente anterior a un intervalo de tiempo de la primera muestra. En particular, el procedimiento puede comprender determinar uno o más coeficientes de predicción de un filtro de predicción recursivo (respuesta finita al impulso) que está configurado para determinar la primera muestra de la primera señal de subbanda de una o más muestras anteriores.

El uno o más coeficientes de predicción pueden determinarse en base al modelo de señal, en base al parámetro de modelo y en base al banco de filtros de análisis. En particular, un coeficiente de predicción puede determinarse en función de una evaluación analítica del modelo de señal y del banco de filtros de análisis. La evaluación analítica del modelo de señal y del banco de filtros de análisis puede dar lugar a la determinación de una tabla de consulta y/o de una función analítica. De este modo, el coeficiente de predicción puede determinarse usando la tabla de consulta y/o la función analítica, donde la tabla de consulta y/o la función analítica pueden predeterminarse en función del modelo de señal y en función del banco de filtros de análisis. La tabla de consulta y/o la función analítica pueden proporcionar el coeficiente o coeficientes de predicción en función de un parámetro obtenido a partir del parámetro o parámetros de modelo. El parámetro obtenido del parámetro de modelo puede ser, por ejemplo, el parámetro de modelo o puede obtenerse del parámetro de modelo usando una función predeterminada. De este modo, el uno o más coeficientes de predicción pueden determinarse de manera computacionalmente eficaz usando una tabla de consulta predeterminada y/o una función analítica que proporcionan el uno o más coeficientes de predicción en función de, solamente, el uno o más parámetros obtenidos, solamente, del uno o más parámetros de modelo. Por tanto, la determinación de un coeficiente de predicción puede reducirse a la simple consulta de una entrada de una tabla de consulta.

Como se ha indicado anteriormente, el banco de filtros de análisis puede comprender o puede presentar una estructura modulada. Como resultado de tal estructura modulada, se observa que el valor absoluto del uno o más coeficientes de predicción depende de un número de índice de la primera subbanda. Esto significa que la tabla de consulta y/o la función analítica pueden ser invariantes al desplazamiento (aparte de un valor de signo) en relación con el número de índice de la pluralidad de subbandas. En tales casos, el parámetro obtenido a partir del parámetro de modelo, es decir, el parámetro que se introduce en la tabla de consulta y/o en la función analítica con el fin de determinar el coeficiente de predicción, puede obtenerse expresando el parámetro de modelo de manera relativa con respecto a una subbanda de la pluralidad de subbandas.

Como se ha indicado anteriormente, el parámetro de modelo puede indicar una frecuencia fundamental Ω de un modelo de señal multisinusoidal (por ejemplo, de un modelo de señal periódico). En tales casos, determinar el coeficiente de predicción puede comprender determinar un múltiplo de la frecuencia fundamental Ω que está dentro de la primera subbanda. Si un múltiplo de la frecuencia fundamental Ω está dentro de la primera subbanda, puede determinarse una desviación relativa del múltiplo de la frecuencia fundamental Ω con respecto a una frecuencia central de la primera subbanda. En particular, puede determinarse la desviación relativa del múltiplo de la frecuencia fundamental Ω más cercana a la frecuencia central de la primera subbanda. La tabla de consulta y/o la función analítica pueden predeterminarse de manera que la tabla de consulta y/o la función analítica proporcionen el coeficiente de predicción en función de posibles desviaciones relativas con respecto a una frecuencia central de una subbanda (por ejemplo, en función de una frecuencia normalizada f y/o en función de un parámetro de desplazamiento Θ , como se describe en el presente documento). De este modo, el coeficiente de predicción puede determinarse conforme a la tabla de consulta y/o conforme a la función analítica usando la desviación relativa determinada. Una tabla de consulta predeterminada puede comprender un número limitado de entradas para un número limitado de posibles desviaciones relativas. En tal caso, la desviación relativa determinada puede redondearse a la desviación relativa más cercana posible de entre el limitado número de posibles desviaciones relativas, antes de consultar el coeficiente de predicción en la tabla de consulta.

Por otro lado, si no hay ningún múltiplo de la frecuencia fundamental Ω en la primera subbanda o, más bien, dentro de una gama de frecuencias ampliada que rodea a la primera subbanda, el coeficiente de predicción puede fijarse a cero. En tales casos, la estimación de la primera muestra puede ser también cero.

Determinar el coeficiente de predicción puede comprender seleccionar una tabla de una pluralidad de tablas de consulta en función del parámetro de modelo. A modo de ejemplo, el parámetro de modelo puede indicar una frecuencia fundamental Ω de un modelo de señal periódico. La frecuencia fundamental Ω de un modelo de señal periódico corresponde a una periodicidad T del modelo de señal periódico. En el presente documento se expone que en caso de periodicidades T relativamente pequeñas, un modelo de señal periódico converge hacia un modelo de una única senoide. Además, en el presente documento se expone que en caso de periodicidades T relativamente grandes, las tablas de consulta varían lentamente con el valor absoluto de T y dependen principalmente de la desviación relativa (esto es, del parámetro de desplazamiento Θ). De este modo, diversas tablas de consulta pueden predeterminarse para una pluralidad de diferentes valores de la periodicidad T . El parámetro de modelo (esto es, la periodicidad T) puede usarse para seleccionar una tabla apropiada de la pluralidad de tablas de consulta, y el coeficiente de predicción puede determinarse en función de la tabla seleccionada de la pluralidad de tablas de consulta (usando la desviación relativa, por ejemplo usando el parámetro de desplazamiento Θ). De este modo, un parámetro de modelo (que representa, por ejemplo, la periodicidad T) que puede tener una precisión relativamente alta puede descodificarse en un par de parámetros (por ejemplo, la periodicidad T y la desviación relativa) con una precisión reducida. El primer parámetro (por ejemplo, la periodicidad T) del par de parámetros puede usarse para seleccionar una tabla de consulta particular, y el segundo parámetro (por ejemplo, la desviación relativa) puede usarse para identificar una entrada de la tabla de consulta seleccionada.

El método puede comprender además determinar una estimación de la primera muestra aplicando el coeficiente de predicción a la muestra anterior. Aplicar el coeficiente de predicción a la muestra anterior puede comprender multiplicar el coeficiente de predicción por el valor de la muestra anterior, obteniéndose así la estimación de la primera muestra. Normalmente, una pluralidad de primeras muestras de la primera señal de subbanda se determina aplicando el coeficiente de predicción a una secuencia de muestras anteriores. Determinar una estimación de la primera muestra puede comprender además aplicar una ganancia de escalado al coeficiente de predicción y/o a la primera muestra. La ganancia de escalado (o una indicación de la misma) puede usarse, por ejemplo, en la predicción a largo plazo (LTP). En otras palabras, la ganancia de escalado puede obtenerse a partir de un predictor diferente (por ejemplo, de un predictor a largo plazo). La ganancia de escalado puede ser diferente para subbandas diferentes. Además, la ganancia de escalado puede transmitirse como parte de la señal de audio codificada.

De este modo se proporciona una descripción eficaz de un predictor de subbanda (que comprende uno o más coeficientes de predicción) usando un modelo de señal que se describe mediante un parámetro de modelo. El parámetro de modelo se usa para determinar el uno o más coeficientes de predicción del predictor de subbanda. Esto significa que un codificador de audio no tiene que transmitir una indicación del uno o más coeficientes de predicción, sino una indicación del parámetro de modelo. Normalmente, el parámetro de modelo puede codificarse de manera más eficaz (es decir, con un menor número de bits) que el uno o más coeficientes de predicción. Por tanto, el uso de la predicción basada en modelo permite una codificación de subbanda de baja velocidad binaria.

El método puede comprender además determinar una máscara de predicción que indica una pluralidad de muestras anteriores en una pluralidad de subbandas de soporte de máscara de predicción. La pluralidad de subbandas de soporte de máscara de predicción puede comprender al menos una de la pluralidad de subbandas, que es diferente de la primera subbanda. De este modo, el predictor de subbanda puede configurarse para estimar una muestra de la primera señal de subbanda a partir de muestras de una o más otras señales de subbanda de la pluralidad de señales de subbanda, que son diferentes de la primera señal de subbanda. En el presente documento, esto se denomina predicción de subbanda cruzada. La máscara de predicción puede definir la disposición de la pluralidad de muestras anteriores (por ejemplo, una demora de tiempo con respecto al intervalo de tiempo de la primera muestra y/o una demora de índice de subbanda con respecto al número de índice de la primera subbanda) que se usan para estimar la primera muestra de la primera señal de subbanda.

El método puede proseguir con la determinación de una pluralidad de coeficientes de predicción que se aplicarán a la pluralidad de muestras anteriores. La pluralidad de coeficientes de predicción puede determinarse en función del modelo de señal, en función del parámetro de modelo y en función del banco de filtros de análisis (por ejemplo, usando los esquemas de predicción basados en modelo indicados anteriormente y en el presente documento). De este modo, la pluralidad de coeficientes de predicción puede determinarse usando uno o más parámetros de modelo. En otras palabras, un número limitado de parámetros de modelo puede ser suficiente para determinar la pluralidad de coeficientes de predicción. Esto significa que usando la predicción de subbanda basada en modelo, la predicción de subbanda cruzada puede implementarse de una manera eficiente en lo que respecta a la tasa binaria.

El método puede comprender determinar una estimación de la primera muestra aplicando la pluralidad de coeficientes de predicción a la pluralidad de muestras anteriores, respectivamente. Determinar una estimación de la primera muestra comprende normalmente determinar la suma de la pluralidad de muestras anteriores ponderadas mediante la pluralidad de coeficientes de predicción respectivos.

Como se ha señalado anteriormente, el parámetro de modelo puede indicar una periodicidad T . La pluralidad de tablas de consulta, que se usan para determinar el uno o más coeficientes de predicción, puede comprender tablas de consulta para diferentes valores de la periodicidad T . En particular, la pluralidad de tablas de consulta puede comprender tablas de consulta para diferentes valores de la periodicidad T dentro del intervalo $[T_{\min}, T_{\max}]$ con un valor de incremento ΔT predeterminado. Como se describirá en el presente documento, T_{\min} puede tener un valor de 0,25 y T_{\max} puede tener un valor de 2,5. T_{\min} puede seleccionarse de manera que $T < T_{\min}$, y la señal de audio puede modelarse usando un modelo de señal que comprende una única componente del modelo sinusoidal. T_{\max} puede seleccionarse de manera que $T > T_{\max}$, y las tablas de consulta para las periodicidades T_{\max} a $T_{\max} + 1$ corresponden sustancialmente a las tablas de consulta para las periodicidades $T_{\max} - 1$ a T_{\max} . Lo mismo se aplica normalmente a las periodicidades $T_{\max} + n$ a $T_{\max} + n + 1$, para $n \geq 0$, en general.

El método puede comprender determinar la tabla de consulta seleccionada como la tabla de consulta para la periodicidad T indicada por el parámetro de modelo. Tras haberse seleccionado la tabla de consulta que comprende o indica el uno o más coeficientes de predicción, puede usarse un parámetro de consulta para identificar la una o más entradas apropiadas de la tabla de consulta seleccionada, que indican el uno o más coeficientes de predicción, respectivamente. El parámetro de consulta puede corresponder a o puede obtenerse a partir del parámetro de desplazamiento Θ .

El método puede comprender, para un parámetro de modelo que indica una periodicidad $T > T_{\max}$, determinar una periodicidad residual T_r sustrayendo de T un valor entero de manera que la periodicidad residual T_r esté en el intervalo $[T_{\max} - 1, T_{\max}]$. La tabla de consulta para determinar el coeficiente de predicción puede determinarse entonces como la tabla de consulta para la periodicidad residual T_r .

5

El método puede comprender, para un parámetro de modelo que indica una periodicidad $T < T_{\min}$, seleccionar la tabla de consulta para determinar el uno o más coeficientes de predicción como la tabla de consulta para la periodicidad T_{\min} . Además, el parámetro de consulta (por ejemplo, el parámetro de desplazamiento Θ) para identificar la una o más entradas de la tabla de consulta seleccionada que proporcionan el uno o más

10

coeficientes de predicción, puede escalarse según la relación T_{\min}/T . El uno o más coeficientes de predicción pueden determinarse entonces usando la tabla de consulta seleccionada y el parámetro de consulta escalado. En particular, el uno o más coeficientes de predicción pueden determinarse en función de la una o más entradas de la tabla de consulta seleccionada correspondiente al parámetro de consulta escalado.

15

De este modo, el número de tablas de consulta puede limitarse a un intervalo predeterminado $[T_{\min}, T_{\max}]$, limitando así los requisitos de memoria de un codificador/descodificador de audio. Sin embargo, los coeficientes de predicción pueden determinarse para todos los posibles valores de la periodicidad T usando las tablas de consulta predeterminadas, permitiendo así una implementación eficaz desde el punto de vista computacional de un codificador/descodificador de audio.

20

Según un aspecto adicional, se describe un procedimiento para estimar una primera muestra de una primera señal de subbanda de una señal de audio. Como se ha indicado anteriormente, la primera señal de subbanda de la señal de audio puede determinarse usando un banco de filtros de análisis que comprende una pluralidad de filtros de análisis que proporcionan una pluralidad de señales de subbanda en una pluralidad de subbandas, respectivamente, de la señal de audio. Las características descritas anteriormente también pueden aplicarse al procedimiento descrito a continuación.

25

El método comprende determinar una máscara de predicción que indica una pluralidad de muestras anteriores en una pluralidad de subbandas de soporte de máscara de predicción. La pluralidad de subbandas de soporte de máscara de predicción comprende al menos una de la pluralidad de subbandas, que es diferente de la primera subbanda. En particular, la pluralidad de subbandas de soporte de máscara de predicción puede comprender la primera subbanda y/o la pluralidad de subbandas de soporte de máscara de predicción puede comprender una o más de la pluralidad de subbandas directamente adyacentes a la primera subbanda.

30

El método puede comprender además determinar una pluralidad de coeficientes de predicción que se aplicarán a la pluralidad de muestras anteriores. La pluralidad de muestras anteriores se obtiene normalmente a partir de la pluralidad de señales de subbanda de la señal de audio. En particular, la pluralidad de muestras anteriores corresponde normalmente a las muestras de una pluralidad de señales de subbanda descodificadas. La pluralidad de coeficientes de predicción puede corresponder a los coeficientes de predicción de un filtro de predicción recursivo (respuesta finita al impulso) que también tiene en cuenta una o más muestras de subbandas que son diferentes de la primera subbanda. Una estimación de la primera muestra puede determinarse aplicando la pluralidad de coeficientes de predicción a la pluralidad de muestras anteriores, respectivamente. De este modo, el procedimiento permite predecir subbandas usando una o más muestras de otras subbandas (por ejemplo, adyacentes). De esta manera pueden reducirse los artefactos de distorsión provocados por los

35

40

45

El método puede comprender además determinar un parámetro de modelo de un modelo de señal. La pluralidad de coeficientes de predicción puede determinarse en función del modelo de señal, en función del parámetro de modelo y en función del banco de filtros de análisis. De este modo, la pluralidad de coeficientes de predicción puede determinarse usando una predicción basada en modelo como la descrita en el presente documento. En particular, la pluralidad de coeficientes de predicción puede determinarse usando una tabla de consulta y/o una función analítica. La tabla de consulta y/o la función analítica pueden determinarse en función del modelo de señal y en función del banco de filtros de análisis. Además, la tabla de consulta y/o la función analítica pueden proporcionar la pluralidad de coeficientes de predicción (solamente) en función de un parámetro obtenido a partir del parámetro de modelo. Por tanto, el parámetro de modelo puede proporcionar directamente la pluralidad de coeficientes de predicción usando la tabla de consulta y/o la función analítica. De este modo, el parámetro de modelo puede usarse para describir de manera eficaz el coeficiente de un predictor de subbanda cruzada. Según un aspecto adicional, se describe un procedimiento para codificar una señal de audio. El procedimiento puede comprender determinar una pluralidad de señales de subbanda a partir de la señal de audio usando un banco de

50

55

60

65

La señal de audio codificada puede indicar (por ejemplo, puede comprender) la pluralidad de señales de subbanda de error de predicción cuantificadas. Además, la señal codificada puede indicar (por ejemplo, puede comprender) uno o más parámetros usados para estimar las muestras de la pluralidad de señales de subbanda estimadas, por ejemplo indicar uno o más parámetros de modelo usados para determinar uno o más coeficientes de predicción que se usan después para estimar las muestras de la pluralidad de señales de subbanda estimadas.

Según otro aspecto, se describe un método para descodificar una señal de audio codificada. La señal de audio codificada indica normalmente una pluralidad de señales de subbanda de error de predicción cuantificadas y uno o más parámetros que se usarán para estimar muestras de una pluralidad de señales de subbanda estimadas. El método puede comprender descuantificar la pluralidad de señales de subbanda de error de predicción cuantificadas, obteniéndose así una pluralidad de señales de subbanda de error de predicción descuantificadas. Además, el método puede comprender estimar muestras de la pluralidad de señales de subbanda estimadas usando cualquiera de los procedimientos de predicción descritos en el presente documento. Muestras de una pluralidad de señales de subbanda descodificadas pueden determinarse en función de muestras correspondientes de la pluralidad de señales de subbanda estimadas y en función de muestras de la pluralidad de señales de subbanda de error de predicción descuantificadas. Una señal de audio descodificada puede determinarse a partir de la pluralidad de señales de subbanda descodificadas usando un banco de filtros de síntesis que comprende una pluralidad de filtros de síntesis.

Según un aspecto adicional, se describe un sistema configurado para estimar una o más primeras muestras de una primera señal de subbanda de una señal de audio. La primera señal de subbanda de la señal de audio puede determinarse usando un banco de filtros de análisis que comprende una pluralidad de filtros de análisis que proporcionan una pluralidad de señales de subbanda a partir de la señal de audio en una pluralidad de subbandas respectivas. El sistema puede comprender un calculador de predictor configurado para determinar un parámetro de modelo de un modelo de señal. Además, el calculador de predictor puede estar configurado para determinar uno o más coeficientes de predicción que se aplicarán a una o más muestras anteriores de una primera señal de subbanda descodificada obtenida a partir de la primera señal de subbanda. De este modo, el calculador de predictor puede estar configurado para determinar uno o más coeficientes de predicción de un filtro de predicción recursivo, específicamente de un filtro de predicción de subbanda recursivo. El uno o más coeficientes de predicción pueden determinarse en base al modelo de señal, en base al parámetro de modelo y en base al banco de filtros de análisis (por ejemplo, usando los procedimientos de predicción basados en modelo descritos en el presente documento). Los intervalos de tiempo de la una o más muestras previas son normalmente anteriores a los intervalos de tiempo de la una o más primeras muestras. El sistema puede comprender además un predictor de subbanda configurado para determinar una estimación de la una o más primeras muestras aplicando el uno o más coeficientes de predicción a la una o más muestras anteriores.

Según otro aspecto, se describe un sistema configurado para estimar una o más primeras muestras de una primera señal de subbanda de una señal de audio. La primera señal de subbanda corresponde a una primera subbanda de una pluralidad de subbandas. La primera señal de subbanda se determina normalmente usando un banco de filtros de análisis que comprende una pluralidad de filtros de análisis que proporcionan una pluralidad de señales de subbanda para la pluralidad de subbandas, respectivamente. El sistema comprende un calculador de predictor configurado para determinar una máscara de predicción que indica una pluralidad de muestras anteriores en una pluralidad de subbandas de soporte de máscara de predicción. La pluralidad de subbandas de soporte de máscara de predicción comprende al menos una de la pluralidad de subbandas, que es diferente de la primera subbanda. El calculador de predictor está configurado además para determinar una pluralidad de coeficientes de predicción (o un filtro de predicción recursivo) que se aplicarán a la pluralidad de muestras anteriores. Además, el sistema comprende un predictor de subbanda configurado para determinar una estimación de la una o más primeras muestras aplicando la pluralidad de coeficientes de predicción a la pluralidad de muestras anteriores, respectivamente.

Según otro aspecto, se describe un codificador de audio configurado para codificar una señal de audio. El codificador de audio comprende un banco de filtros de análisis configurado para determinar una pluralidad de señales de subbanda a partir de la señal de audio usando una pluralidad de filtros de análisis. Además, el codificador de audio comprende un calculador de predictor y un predictor de subbanda como los descritos en el presente documento, que están configurados para estimar muestras de la pluralidad de señales de subbanda, obteniéndose así una pluralidad de señales de subbanda estimadas. Además, el codificador puede comprender una unidad de diferencia configurada para determinar muestras de una pluralidad de señales de subbanda de error de predicción basadas en muestras correspondientes de la pluralidad de señales de subbanda y de la pluralidad de señales de subbanda estimadas. Puede usarse una unidad de cuantificación para cuantificar la pluralidad de señales de subbanda de error de predicción. Además, una unidad de generación de flujo de bits puede estar configurada para generar una señal de audio codificada que indica la pluralidad de señales de subbanda de error de predicción cuantificadas y uno o más parámetros (por ejemplo, uno o más parámetros de modelo) usados para estimar las muestras de la pluralidad de señales de subbanda estimadas.

Según un aspecto adicional, se describe un descodificador de audio configurado para descodificar una señal de

audio codificada. La señal de audio codificada indica (por ejemplo, comprende) la pluralidad de señales de subbanda de error de predicción cuantificadas y uno o más parámetros usados para estimar muestras de una pluralidad de señales de subbanda estimadas. El decodificador de audio puede comprender un cuantificador inverso configurado para descuantificar la pluralidad de señales de subbanda de error de predicción cuantificadas, obteniéndose así una pluralidad de señales de subbanda de error de predicción descuantificadas. Además, el decodificador comprende un calculador de predictor y un predictor de subbanda como los descritos en el presente documento, que están configurados para estimar muestras de la pluralidad de señales de subbanda estimadas. Una unidad de suma puede usarse para determinar muestras de una pluralidad de señales de subbanda descodificadas en función de muestras correspondientes de la pluralidad de señales de subbanda estimadas y en función de muestras de la pluralidad de señales de subbanda de error de predicción descuantificadas. Además, puede usarse un banco de filtros de síntesis para determinar una señal de audio descodificada a partir de la pluralidad de señales de subbanda descodificadas usando una pluralidad de filtros de síntesis.

Según un aspecto adicional, se describe un programa de software. El programa de software puede estar adaptado para ejecutarse en un procesador y para llevar a cabo las etapas de procedimiento descritas en el presente documento cuando se ejecuta en el procesador.

Según otro aspecto, se describe un medio de almacenamiento. El medio de almacenamiento puede comprender un programa de software adaptado para ejecutarse en un procesador y para llevar a cabo las etapas del método descritas en el presente documento cuando se ejecuta en el procesador.

Según un aspecto adicional, se describe un producto de programa de ordenador. El programa de ordenador puede comprender instrucciones ejecutables para llevar a cabo las etapas del método descritas en el presente documento cuando se ejecuta en un ordenador. La invención se expone mediante las reivindicaciones adjuntas.

Debe observarse que los procedimientos y sistemas, incluidas sus realizaciones preferidas descritas en la presente solicitud de patente, pueden usarse de manera independiente o en combinación con los otros procedimientos y sistemas descritos en este documento.

Breve descripción de las figuras

La presente invención se describe a continuación de manera ilustrativa, sin limitar el alcance de la invención, con referencia a los dibujos adjuntos, en los que:

la figura 1 ilustra el diagrama de bloques de un decodificador de audio de ejemplo que aplica predicción lineal en un dominio de banco de filtros (es decir, en un dominio de subbanda);

la figura 2 muestra máscaras de predicción de ejemplo en una cuadrícula de tiempo-frecuencia;

la figura 3 ilustra datos tabulados de ejemplo para un calculador de predictor basado en un modelo sinusoidal;

la figura 4 ilustra una conformación de ruido de ejemplo que se obtiene a partir de una predicción de subbanda dentro de banda;

la figura 5 ilustra una conformación de ruido de ejemplo que se obtiene a partir de una predicción de subbanda de banda cruzada; y

la figura 6a ilustra una cuadrícula de cuantificación bidimensional de ejemplo subyacente a los datos tabulados para un cálculo de predictor basado en un modelo periódico;

la figura 6b ilustra el uso de diferentes máscaras de predicción para diferentes intervalos de periodicidades de señal; y

las figuras 7a y 7b muestran diagramas de flujo de procedimientos de codificación y decodificación de ejemplo que usan predicción de subbanda basada en modelo.

Descripción detallada

Las realizaciones descritas a continuación simplemente ilustran los principios de la presente invención para una predicción basada en modelo en un banco de filtros críticamente muestreado. Debe entenderse que modificaciones y variaciones de las disposiciones y de los detalles descritos en el presente documento resultarán evidentes a los expertos en la técnica. Por lo tanto, solo estarán limitadas por el alcance de las reivindicaciones de patente adjuntas y no por los detalles específicos presentados con fines descriptivos y explicativos de las realizaciones del presente documento.

La figura 1 ilustra el diagrama de bloques de un descodificador 100 de audio de ejemplo que aplica predicción lineal en un dominio de banco de filtros (denominado también dominio de subbanda). El descodificador 100 de audio recibe un flujo de bits que comprende información relacionada con una señal de error de predicción (también denominada señal residual) y, posiblemente, información relacionada con la descripción de un predictor usada por un codificador correspondiente para determinar la señal de error de predicción a partir de una señal de audio de entrada original. La información relacionada con la señal de error de predicción puede referirse a subbandas de la señal de audio de entrada, y la información relacionada con la descripción del predictor puede referirse a uno o más predictores de subbanda.

Dada la información de flujo de bits recibida, un cuantificador 101 inverso puede proporcionar muestras 111 de las señales de subbanda de error de predicción. Estas muestras pueden añadirse a la salida 112 del predictor de subbanda 103 y la suma 113 puede transferirse a una memoria intermedia 104 de subbandas que mantiene un registro de muestras 113 descodificadas anteriormente de las subbandas de la señal de audio descodificada. La salida del predictor 103 de subbanda puede denominarse como señales 112 de subbanda estimadas. Las muestras 113 descodificadas de las subbandas de la señal de audio descodificada pueden transmitirse a un banco de filtros 102 de síntesis, que convierte las muestras de subbanda al dominio de tiempo, obteniéndose así muestras 114 de dominio de tiempo de la señal de audio descodificada.

En otras palabras, el descodificador 100 puede funcionar en el dominio de subbanda. En particular, el descodificador 100 puede determinar una pluralidad de señales 112 de subbanda estimadas usando el predictor 103 de subbanda. Además, el descodificador 100 puede determinar una pluralidad de señales 111 de subbanda residuales usando el cuantificador 101 inverso. Pueden añadirse parejas respectivas de la pluralidad de señales 112 de subbanda estimadas y de la pluralidad de señales 111 de subbanda residuales para obtener una pluralidad correspondiente de señales 113 de subbanda descodificadas. La pluralidad de señales 113 de subbanda descodificadas puede enviarse a un banco de filtros 102 de síntesis para obtener la señal 114 de audio descodificada del dominio del tiempo.

En una realización del predictor 103 de subbanda, una muestra dada de una señal 112 de subbanda estimada dada puede obtenerse mediante una combinación lineal de muestras de subbanda en la memoria intermedia 104, que corresponde a un tiempo diferente y a una frecuencia diferente (es decir, a una subbanda diferente) con respecto a la muestra dada de la señal 112 de subbanda estimada dada. En otras palabras, una muestra de una señal 112 de subbanda estimada en un primer instante de tiempo y en una primera subbanda puede determinarse en base a una o más muestras de las señales 113 de subbanda descodificadas relativas a un segundo instante de tiempo (diferente del primer instante de tiempo) y relativas a una segunda subbanda (diferente de la primera subbanda). La colección de coeficientes de predicción y su vinculación a una máscara de tiempo y frecuencia pueden definir el predictor 103, y esta información puede suministrarse por el calculador 105 de predictor del descodificador 100. El calculador 105 de predictor proporciona la información que define al predictor 103 mediante una conversión de datos de modelo de señal incluidos en el flujo de bits recibido. Puede transmitirse una ganancia adicional que modifica el escalado de la salida del predictor 103. En una forma de realización del calculador de predictor 105, los datos de modelo de señal se proporcionan en forma de un espectro de líneas parametrizadas de manera eficaz, donde cada línea del espectro de líneas parametrizadas, o un grupo de líneas subsiguientes del espectro de líneas parametrizadas, se usa para indicar valores tabulados de coeficientes de predictor. De este modo, los datos de modelo de señal proporcionados en el flujo de bits recibido pueden usarse para identificar entradas de una tabla de consulta predeterminada, donde las entradas de la tabla de consulta proporcionan uno o más valores de los coeficientes de predictor (también denominados coeficientes de predicción) que serán usados por el predictor 103. El método aplicado para la tabla de consulta puede depender de equilibrios entre los requisitos de complejidad y de memoria. Por ejemplo, puede usarse una consulta de tipo 'vecino más cercano' para conseguir la complejidad más baja, mientras que un procedimiento de consulta por interpolación puede proporcionar un rendimiento similar con un tamaño de tabla más pequeño.

Como se ha indicado anteriormente, el flujo de bits recibido puede comprender una o más ganancias transmitidas de manera explícita (o indicaciones de ganancias transmitidas de manera explícita). Las ganancias pueden aplicarse como parte de o después de la operación de predictor. La una o más ganancias transmitidas de manera explícita pueden ser diferentes para subbandas diferentes. Las indicaciones de ganancias adicionales transmitidas de manera explícita se proporcionan junto con uno o más parámetros de modelo que se usan para determinar los coeficientes de predicción del predictor 103. De este modo, las ganancias adicionales pueden usarse para escalar los coeficientes de predicción del predictor 103.

La figura 2 muestra soportes de máscara de predicción de ejemplo en una cuadrícula de tiempo-frecuencia. Los soportes de máscara de predicción pueden usarse en predictores 103 que funcionan en un banco de filtros con una resolución de tiempo-frecuencia uniforme, tal como un banco de filtros modulado por coseno (por ejemplo, un banco de filtros MDCT). La notación se ilustra en el diagrama 201, donde una muestra 211 de subbanda objetivo de color oscuro es la salida de una predicción basada en una muestra 212 de subbanda de color claro. En los diagramas 202 a 205, la colección de muestras de subbanda de color claro indica el soporte de máscara de predictor. La combinación de muestras 212 de subbanda fuente y las muestras 211 de subbanda objetivo se denominará máscara 201 de predicción. Puede usarse una cuadrícula de tiempo-frecuencia para disponer

muestras de subbanda cerca de la muestra de subbanda objetivo. El índice de intervalo de tiempo aumenta de izquierda a derecha, y el índice de frecuencia de subbanda aumenta de abajo arriba. La figura 2 muestra casos de ejemplo de máscaras de predicción y de soportes de máscara de predictor, y debe observarse que pueden usarse otras máscaras de predicción y otros soportes de máscara de predictor. Las máscaras de predicción de ejemplo son:

- La máscara 202 de predicción define la predicción dentro de banda de una muestra 221 de subbanda estimada en el instante de tiempo k a partir de dos muestras 222 de subbanda descodificadas anteriores en los instantes de tiempo $k-1$ y $k-2$.

- La máscara 203 de predicción define la predicción de banda cruzada de una muestra 231 de subbanda estimada en el instante de tiempo k y en la subbanda n en función de tres muestras 232 de subbanda descodificadas anteriores en el instante de tiempo $k-1$ y en las subbandas $n-1$, n , $n+1$.

- La máscara 204 de predicción define la predicción de banda cruzada de tres muestras 241 de subbanda estimadas en el instante de tiempo k y en tres subbandas diferentes $n-1$, n , $n+1$ en base a tres muestras 242 de subbanda descodificadas anteriores en el instante de tiempo $k-1$ y en las subbandas $n-1$, n , $n+1$. La predicción de banda cruzada puede realizarse de manera que cada muestra 241 de subbanda estimada pueda determinarse en función de las tres muestras 242 de subbanda descodificadas anteriores en las subbandas $n-1$, n , $n+1$.

- La máscara 205 de predicción define la predicción de banda cruzada de una muestra 251 de subbanda estimada en el instante de tiempo k y en la subbanda n en función de doce muestras 252 de subbanda descodificadas anteriores en los instantes de tiempo $k-2$, $k-3$, $k-4$, $k-5$ y en las subbandas $n-1$, n , $n+1$.

La figura 3 ilustra datos tabulados para un calculador 105 de predictor basado en modelo sinusoidal que funciona en un banco de filtros modulado por coseno. El soporte de máscara de predicción es el del diagrama 204. Para un parámetro de frecuencia dado, la subbanda con la frecuencia central de subbanda más cercana puede seleccionarse como subbanda objetivo central. La diferencia entre el parámetro de frecuencia y la frecuencia central de la subbanda objetivo central puede calcularse en unidades de la separación de frecuencia del banco de filtros (celdas). Esto genera un valor comprendido entre $-0,5$ y $0,5$ que puede redondearse a la entrada disponible más cercana de los datos tabulados, ilustrada por las abscisas de los nueve gráficos 301 de la figura 3. Esto produce una matriz 3×3 de coeficientes que puede aplicarse a los valores más recientes de la pluralidad de señales de subbanda descodificadas 113 de la memoria intermedia 104 de subbandas de la subbanda objetivo y sus dos subbandas adyacentes. El vector 3×1 resultante constituye la contribución del predictor 103 de subbanda a estas tres subbandas para el parámetro de frecuencia dado. El proceso puede repetirse de manera aditiva para todas las componentes sinusoidales del modelo de señal.

En otras palabras, la figura 3 ilustra un ejemplo de una descripción basada en modelo de un predictor de subbanda. Se supone que la señal de audio de entrada comprende una o más componentes sinusoidales a frecuencias fundamentales $\Omega_0, \Omega_1, \dots, \Omega_{M-1}$. Un predictor de subbanda que usa una máscara de predicción predeterminada (por ejemplo, la máscara 204 de predicción) puede determinarse para cada componente sinusoidal de la una o más componentes sinusoidales. Una frecuencia fundamental Q de la señal de audio de entrada puede estar dentro de una de las subbandas del banco de filtros. Esta subbanda puede denominarse subbanda central para esta frecuencia fundamental particular Q . La frecuencia fundamental Q puede expresarse como un valor comprendido entre $-0,5$ y $0,5$ con respecto a la frecuencia central de la subbanda central. Un codificador de audio puede transmitir al descodificador 100 información relacionada con la frecuencia fundamental Q . El calculador 105 de predictor del descodificador 100 puede usar la matriz 3×3 de la figura 3 para determinar una matriz tres por tres de coeficientes de predicción determinando el valor 302 del coeficiente para el valor 303 de frecuencia relativa de la frecuencia fundamental Q . Esto significa que el coeficiente para un predictor de subbanda 103 que usa una máscara 204 de predicción puede determinarse usando solamente la información recibida relacionada con la frecuencia fundamental particular Q . En otras palabras, modelando una señal de audio de entrada usando, por ejemplo, un modelo de una de más componentes sinusoidales, puede proporcionarse una descripción eficaz de la tasa de bits de un predictor de subbanda.

La figura 4 ilustra una conformación de ruido de ejemplo que se obtiene de una predicción de subbanda dentro de banda en un banco de filtros modulado por coseno. El modelo de señal usado para realizar la predicción de subbanda dentro de banda es un proceso estocástico autorregresivo de segundo orden con una resonancia máxima, como se describe mediante una ecuación diferencial de segundo orden basada en ruido blanco gaussiano aleatorio. La curva 401 muestra el espectro de magnitud medido para una realización del proceso. En este ejemplo se aplica la máscara 202 de predicción de la figura 2. Es decir, el calculador 105 de predictor suministra el predictor 103 de subbanda para una subbanda 221 objetivo dada basándose en muestras 222 de subbanda anteriores solamente de la misma subbanda. Sustituyendo el cuantificador 101 inverso por un generador de ruido blanco gaussiano se obtiene un espectro 402 de magnitud sintetizado. Como puede observarse, en la síntesis se producen importantes artefactos de distorsión, ya que el espectro 402 sintetizado comprende picos que no coinciden con el espectro 401 original.

La figura 5 ilustra una conformación de ruido de ejemplo que se obtiene de la predicción de subbanda de banda cruzada. La configuración es idéntica a la de la figura 4, excepto que se aplica la máscara 203 de predicción. Por tanto, el calculador 105 suministra el predictor 103 para una subbanda 231 objetivo dada basándose en muestras 232 de subbanda anteriores de la subbanda objetivo y en sus dos subbandas adyacentes. Como puede observarse en la figura 5, el espectro 502 de la señal sintetizada coincide sustancialmente con el espectro 501 de la señal original, es decir, los problemas de distorsión se suprimen considerablemente cuando se usa predicción de subbanda de banda cruzada. De este modo, las figuras 4 y 5 ilustran que cuando se usa predicción de subbanda de banda cruzada, es decir, cuando se predice una muestra de subbanda en función de muestras de subbanda anteriores de una o más subbandas adyacentes, los artefactos de distorsión producidos por la predicción de subbanda pueden reducirse. Como resultado, la predicción de subbanda también puede aplicarse en el contexto de codificadores de audio de baja velocidad binaria sin riesgo de generar artefactos de distorsión audibles. El uso de la predicción de subbanda de banda cruzada aumenta normalmente el número de coeficientes de predicción. Sin embargo, como se muestra en el contexto de la figura 3, el uso de modelos para la señal de audio de entrada (por ejemplo, el uso de un modelo sinusoidal o de un modelo periódico) permite una descripción eficaz del predictor de subbanda, permitiéndose así el uso de la predicción de subbanda de banda cruzada en codificadores de audio de baja velocidad binaria.

A continuación, con referencia a las figuras 1 a 6, se ofrecerá una descripción de los principios de la predicción basada en modelo en un banco de filtros críticamente muestreado, utilizando una terminología matemática apropiada.

Un posible modelo de señal subyacente a la predicción lineal es el de un proceso estocástico débilmente estacionario de media cero $x(t)$ cuyas estadísticas se determinan mediante su función de autocorrelación $r(\tau) = E\{x(t)x(t-\tau)\}$. Como un buen modelo para los bancos de filtros críticamente muestreados a considerar, se toma $\{w_\alpha : \alpha \in A\}$ como una colección de formas de onda de síntesis de valores reales $w_\alpha(t)$ que constituyen una base ortonormal. En otras palabras, el banco de filtros puede representarse mediante las formas de onda $\{w_\alpha : \alpha \in A\}$. Muestras de subbanda de una señal de dominio de tiempo $s(t)$ se obtienen mediante productos internos

$$\langle s, w_\alpha \rangle = \int_{-\infty}^{\infty} s(t) w_\alpha(t) dt, \quad (1)$$

y la señal se obtiene de la siguiente manera:

$$s(t) = \sum_{\alpha \in A} \langle s, w_\alpha \rangle w_\alpha(t), \quad (2)$$

Las muestras de subbanda $\langle x, w_\alpha \rangle$ del proceso $x(t)$ son variables aleatorias, cuya matriz de covarianza $R_{\alpha\beta}$ se determina mediante la función de autocorrelación $r(\tau)$ de la siguiente manera:

$$R_{\alpha\beta} = E\{\langle x, w_\alpha \rangle \langle x, w_\beta \rangle\} = \langle W_{\alpha\beta}, r \rangle, \quad (3)$$

donde $W_{\alpha\beta}(\tau)$ es la correlación cruzada de dos formas de onda de síntesis

$$W_{\alpha\beta}(\tau) = \int_{-\infty}^{\infty} w_\alpha(t) w_\beta(t-\tau) dt. \quad (4)$$

Una predicción lineal de la muestra de subbanda $\langle x, w_\alpha \rangle$ de una colección o de muestras de subbanda descodificadas $\{\langle x, w_\beta \rangle : \beta \in B\}$ se define como:

$$\sum_{\beta \in B} c_\beta \langle x, w_\beta \rangle. \quad (5)$$

En la ecuación (5), el conjunto B define las muestras de subbanda fuente, es decir, el conjunto B define el soporte de máscara de predicción. El valor medio del error cuadrático de predicción viene dado por

$$E \left\{ \left(\sum_{\beta \in B} c_{\beta} \langle x, w_{\beta} \rangle - \langle x, w_{\alpha} \rangle \right)^2 \right\} = \sum_{\beta, \gamma \in B} c_{\gamma} R_{\gamma\beta} c_{\beta} - 2 \sum_{\beta \in B} R_{\alpha\beta} c_{\beta} + R_{\alpha\alpha}, \quad (6)$$

y la solución del error cuadrático medio (MSE) mínimo se obtiene resolviendo las ecuaciones normales de los coeficientes de predicción c_{β} ,

$$\sum_{\beta \in B} R_{\gamma\beta} c_{\beta} = R_{\gamma\alpha}, \quad \gamma \in B. \quad (7)$$

Cuando los coeficientes de predicción satisfacen la ecuación (7), el lado derecho de la ecuación (6) se reduce a $R_{\alpha\alpha} - \sum_{\beta \in B} R_{\alpha\beta} c_{\beta}$. Las ecuaciones normales (7) pueden resolverse de manera eficaz usando, por ejemplo, el algoritmo de Levinson-Durbin.

En el presente documento se propone transmitir una representación paramétrica de un modelo de señal a partir del cual los coeficientes de predicción $\{c_{\beta} : \beta \in B\}$ pueden obtenerse en el calculador 105 de predictor. Por ejemplo, el modelo de señal puede proporcionar una representación paramétrica de la función de autocorrelación $r(\tau)$ del modelo de señal. El descodificador 100 puede obtener la función de autocorrelación $r(\tau)$ usando la representación paramétrica recibida y puede combinar la función de autocorrelación $r(\tau)$ con la correlación cruzada de forma de onda de síntesis $W_{\alpha\beta}(\tau)$ con el fin de obtener las entradas de la matriz de covarianza requeridas para las ecuaciones normales (7). Estas ecuaciones pueden resolverse entonces para obtener los coeficientes de predicción.

En otras palabras, una señal de audio de entrada que va a codificarse puede modelarse mediante un proceso $x(t)$ que puede describirse usando un número limitado de parámetros de modelo. En particular, el proceso de modelado $x(t)$ puede ser tal que su función de autocorrelación $r(\tau) = E\{x(t)x(t-\tau)\}$ puede describirse usando un número limitado de parámetros. El número limitado de parámetros para describir la función de autocorrelación $r(\tau)$ puede transmitirse al descodificador 100. El calculador 105 de predictor del descodificador 100 puede determinar la función de autocorrelación $r(\tau)$ a partir de los parámetros recibidos y puede usar la ecuación (3) para determinar la matriz de covarianza $R_{\alpha\beta}$ de las señales de subbanda a partir de las cuales puede determinarse la ecuación normal (7). El calculador 105 de predictor puede resolver entonces la ecuación normal (7), obteniéndose así los coeficientes de predicción c_{β} .

A continuación se describen modelos de señal de ejemplo que pueden usarse para aplicar el esquema de predicción basado en modelo antes descrito de manera eficaz. Los modelos de señal descritos a continuación son, por lo general, muy relevantes para codificar señales de audio, por ejemplo para codificar señales de voz.

Un ejemplo de un modelo de señal viene dado por el proceso sinusoidal

$$x(t) = a \cos(\xi t) + b \sin(\xi t), \quad (8)$$

donde las variables aleatorias a , b no están correlacionadas, tienen una media cero y una varianza uno. La función de autocorrelación de este proceso sinusoidal viene dada por

$$r(\tau) = \cos(\xi \tau). \quad (9)$$

Una generalización de tal proceso sinusoidal es un modelo multiseno que comprende un conjunto de frecuencias (angulares) S , es decir, que comprende una pluralidad de diferentes frecuencias (angulares) ξ ,

$$x(t) = \sum_{\xi \in S} a_{\xi} \cos(\xi t) + b_{\xi} \sin(\xi t). \quad (10)$$

Suponiendo que todas las variables aleatorias a_{ξ} , b_{ξ} no están correlacionadas en parejas, tienen una media cero y una varianza uno, el proceso multiseno tiene la función de autocorrelación

$$r(\tau) = \sum_{\xi \in S} \cos(\xi \tau). \quad (11)$$

La densidad espectral de potencia (PSD) del proceso multiseno (que corresponde a la transformada de Fourier

de la función de autocorrelación), es el espectro de líneas

$$P(\omega) = \frac{1}{2} \sum_{\xi \in S} (\delta(\omega - \xi) + \delta(\omega + \xi)). \quad (12)$$

- 5 Consideraciones numéricas pueden dar lugar a la sustitución del proceso multiseno puro con función de autocorrelación del proceso de ecuación por un proceso multiseno relajado que presenta la función de autocorrelación

$$r(\tau) = \exp(-\varepsilon |\tau|) \sum_{\xi \in S} \cos(\xi \tau)$$

10

donde $\varepsilon > 0$ es un parámetro de relajación relativamente pequeño. Este último modelo da lugar a una PSD estrictamente positiva sin funciones de impulsos.

Ejemplos de descripciones compactas del conjunto S de frecuencias de un modelo multiseno son los siguientes:

15

1. Una única frecuencia fundamental $\Omega : S = \{\Omega v : v = 1, 2, \dots\}$

2. M frecuencias fundamentales: $\Omega_0, \Omega_1, \dots, \Omega_{M-1} : S = \{\Omega_k v : v = 1, 2, \dots, k = 0, 1, \dots, M-1\}$

20

3. Una única frecuencia fundamental desplazada de banda lateral $S_{2,\theta} : S = \{\Omega(v + \theta) : v = 1, 2, \dots\}$

4. Un modelo ligeramente inarmónico: $\Omega, \alpha : S = \{\Omega v \cdot (1 + \alpha v^2)^{1/2} : v = 1, 2, \dots\}$, donde α describe la componente inarmónica del modelo.

25

De este modo, un modelo multiseno (posiblemente relajado) que presenta una PSD dada por la ecuación (12) puede describirse de manera eficaz usando una de las descripciones de ejemplo antes enumeradas. A modo de ejemplo, un conjunto completo S de frecuencias del espectro de líneas de la ecuación (12) puede describirse usando solamente una única frecuencia fundamental Q. Si la señal de audio de entrada que va a codificarse puede describirse correctamente usando un modelo multiseno que presenta una única frecuencia fundamental Q, el predictor basado en modelo puede describirse mediante un único parámetro (es decir, mediante la frecuencia fundamental Q), independientemente del número de coeficientes de predicción (es decir, independientemente de la máscara 202, 203, 204, 205 de predicción) usada por el predictor 103 de subbanda.

30

El caso 1 para describir el conjunto S de frecuencias proporciona un proceso $x(t)$ que modela señales de audio de entrada con un periodo $T = 2\pi/\Omega$. Tras la inclusión de la contribución de frecuencia cero (DC) con varianza 1/2 a la ecuación (11) y sujeta al reescalado del resultado mediante el factor $2/T$, la función de autocorrelación del proceso de modelo periódico $x(t)$ puede escribirse como

35

$$r(\tau) = \sum_{k \in \mathbb{Z}} \delta(\tau - kT). \quad (13)$$

40

Con la definición de un factor de relajación $\rho = \exp(-T\varepsilon)$, la función de autocorrelación de la versión relajada del modelo periódico viene dada por

$$r(\tau) = \sum_{k \in \mathbb{Z}} \rho^{|k|} \delta(\tau - kT). \quad (14)$$

45

La ecuación (14) corresponde también a la función de autocorrelación de un proceso definido mediante un único bucle de retardo alimentado con ruido blanco $z(t)$, es decir, del proceso de modelo

$$x(t) = \rho x(t - T) + \sqrt{1 - \rho^2} z(t). \quad (15)$$

50

Esto significa que el proceso periódico que presenta una única frecuencia fundamental Ω corresponde a un retardo en el dominio de tiempo, donde el retardo es $T = 2\pi/\Omega$.

55

Los modelos de señal globales antes mencionados tienen normalmente un espectro de potencia plano de gran escala debido a la suposición de varianza unitaria de los parámetros de amplitud sinusoidales a_ξ , b_ξ . Sin embargo, debe observarse que los modelos de señal solo se consideran normalmente de manera local para un subconjunto de subbandas de un banco de filtros críticamente muestreado, donde el banco de filtros es

instrumental en la conformación de todo el espectro. En otras palabras, para una señal que tiene una forma espectral con una baja variación en comparación con los anchos de subbanda, los modelos de espectro de potencia plano proporcionarán una buena aproximación de la señal y, por consiguiente, los predictores basados en modelo ofrecerán niveles adecuados de ganancia de predicción.

Más en general, el modelo PSD puede describirse en lo que respecta a parametrizaciones estándar de procesos autorregresivos (AR) o procesos autorregresivos de media móvil (ARMA). Esto mejorará el rendimiento de la predicción basada en modelo, posiblemente a expensas de un aumento en los parámetros de modelo descriptivos.

Otra variante se obtiene rechazando la suposición estacionaria para el modelo de señal estocástico. La función de autocorrelación se convierte entonces en una función de dos variables $r(t,s) = E\{x(t)x(s)\}$. Por ejemplo, los modelos sinusoidales no estacionarios pertinentes pueden incluir modulación en amplitud (AM) y en frecuencia (FM).

Además, puede utilizarse un modelo de señal más determinista. Como se observará en algunos de los ejemplos siguientes, la predicción puede tener un error decreciente en algunos casos. En tales casos, puede evitarse el enfoque probabilístico. Cuando la predicción es perfecta para todas las señales en un espacio de modelo, no es necesario calcular un valor medio del rendimiento de predicción mediante una medida de probabilidad del espacio de modelo considerado.

A continuación se describen varios aspectos relacionados con los bancos de filtros modulados. En particular, se describen aspectos que influyen en la determinación de la matriz de covarianza, proporcionándose así medios eficientes para determinar los coeficientes de predicción de un predictor de subbanda.

Un banco de filtros modulado puede describirse presentando un conjunto de índices bidimensional de formas de onda de síntesis $\alpha = (n, k)$ donde $n = 0, 1, \dots$ es el índice de subbanda (banda de frecuencia) y donde $k \in \mathbb{Z}$ es el índice de muestra de subbanda (ranura de tiempo). Para facilitar la exposición, se supone que las formas de onda de síntesis se proporcionan en tiempo continuo y se normalizan con respecto a lapso de tiempo unitario,

$$w_{n,k}(t) = u_n(t - k), \quad (16)$$

donde

$$u_n(t) = v(t) \cos \left[\pi \left(n + \frac{1}{2} \right) \left(t + \frac{1}{2} \right) \right], \quad (17)$$

en caso de un banco de filtros modulado por coseno. Se supone que la función de ventana $v(t)$ tiene valor real y es uniforme. Teniendo en cuenta pequeñas variantes de la regla de modulación, esto abarca varios casos muy importantes, tales como MDCT (Transformada Discreta del Coseno Modificada), QMF (Filtro de Espejo en Cuadratura) y ELT (Transformadas Solapadas Extendidas) con L subbandas tras el muestreo en un salto de tiempo $1/L$. Se supone que la ventana tiene una duración o longitud finita con soporte incluido en el intervalo $[-K/2, K/2]$, donde K es el factor de solapamiento de la transformada solapada y donde K indica la longitud de la función de ventana.

Debido a la estructura invariante al desplazamiento, se observa que la función de correlación cruzada de la forma de onda de síntesis (definida en la ecuación (4)) puede escribirse como

$$W_{n,k,m,l}(\tau) = \int_{-\infty}^{\infty} w_{n,k}(t) w_{m,l}(t - \tau) dt = \int_{-\infty}^{\infty} u_n(t) u_m(t - l + k - \tau) dt. \quad (18)$$

Es decir, $W_{n,k,m,l}(\tau) = U_{n,m}(\tau - l + k)$, con la definición $U_{n,m}(\tau) = W_{n,0,m,0}(\tau)$. La estructura de modulación (17) permite una expansión adicional a

$$U_{n,m}(\tau) = \frac{1}{2} \kappa_{n-m}(\tau) \cos \frac{\pi}{2} \left[(n+m+1)\tau + (n-m) \right] + \frac{1}{2} \kappa_{n+m+1}(\tau) \cos \frac{\pi}{2} \left[(n-m)\tau + (n+m+1) \right]. \quad (19)$$

donde la función kernel κ_v representa un muestreo con la etapa de subbanda de banco de filtros en la variable de frecuencia de la distribución Wigner-Ville de la ventana del banco de filtros

$$\kappa_v(\tau) = \int_{-\infty}^{\infty} v\left(t + \frac{\tau}{2}\right) v\left(t - \frac{\tau}{2}\right) \cos(\pi v t) dt. \quad (20)$$

5 La función kernel es real y uniforme tanto en v como en τ , debido a las suposiciones antes mencionadas de la función de ventana $v(t)$. Su transformada de Fourier es el producto de respuestas de ventana desplazadas,

$$\hat{\kappa}_v(\omega) = \hat{v}\left(\omega + \frac{\pi}{2}v\right) \hat{v}\left(\omega - \frac{\pi}{2}v\right). \quad (21)$$

10 A partir de las ecuaciones (20) y (21) puede observarse que la función kernel $\kappa_v(\tau)$ disminuye para $|\tau| > K$ y tiene un descenso rápido en función de $|v|$ para opciones típicas de las ventanas de banco de filtros $v(t)$. En consecuencia, el segundo término de la ecuación (19) que implica $v = n+m+1$ puede despreciarse normalmente, excepto en las subbandas más bajas.

15 En lo que respecta a la función de autocorrelación $r(\tau)$ de un modelo de señal dado, las fórmulas antes mencionadas pueden insertarse en la definición de la matriz de covarianza de muestras de subbanda proporcionada por la ecuación (3). Se obtiene $R_{n,k,m,l} = R_{n,m}[k-l]$ con la definición

$$R_{n,m}[\lambda] = \int_{-\infty}^{\infty} U_{n,m}(\tau) r(\tau + \lambda) d\tau. \quad (22)$$

20 En función de la densidad espectral de potencia $P(\omega)$ del modelo de señal dado (que corresponde a la transformada de Fourier de la función de autocorrelación $r(\tau)$), se observa que

$$R_{n,m}[\lambda] = \frac{1}{2\pi} \int_{-\infty}^{\infty} \hat{U}_{n,m}(\omega) P(\omega) \exp(-i\omega\lambda) d\omega. \quad (23)$$

25 donde $\hat{U}_{n,m}(\omega)$ es la transformada de Fourier de $U_{n,m}(\tau)$, donde n, m identifican índices de subbanda, y donde λ representa una demora de ranura de tiempo ($\lambda = k-l$). La expresión de la ecuación (23) puede reescribirse como

$$\begin{aligned} R_{n,m}[\lambda] = & \frac{1}{4\pi} \int_{-\infty}^{\infty} \hat{\kappa}_{n-m}\left(\omega - \frac{\pi}{2}(n+m+1)\right) P(\omega) \cos\left(\omega\lambda - \frac{\pi}{2}(n-m)\right) d\omega \\ & + \frac{1}{4\pi} \int_{-\infty}^{\infty} \hat{\kappa}_{n-m+1}\left(\omega - \frac{\pi}{2}(n-m)\right) P(\omega) \cos\left(\omega\lambda - \frac{\pi}{2}(n+m+1)\right) d\omega. \end{aligned} \quad (24)$$

30 Una observación importante es que el primer término de la ecuación (24) tiene esencialmente una propiedad de invarianza con respecto a los desplazamientos de frecuencia. Si se desprecia el segundo término de la ecuación (24) y $P(\omega)$ desplaza en un número entero v de veces la separación de subbanda π a $P(\omega - \pi v)$, se observa un desplazamiento correspondiente en las covarianzas $R_{n,m}[\lambda] = \pm R_{n-v,m-v}[\lambda]$, donde el signo depende de los valores (enteros) de la demora de tiempo λ . Esto refleja la ventaja de usar un banco de filtros con una estructura de modulación, en comparación con el caso general del banco de filtros.

35 La ecuación (24) proporciona un medio eficaz de determinar los coeficientes matriciales de la matriz de covarianza de muestras de subbanda cuando se conoce la PSD del modelo de señal subyacente. A modo de ejemplo, en caso de un esquema de predicción basado en modelo sinusoidal que utiliza un modelo de señal $x(t)$ que comprende una única senoide a la frecuencia (angular) ξ , la PSD viene dada por $P(\omega) = \frac{1}{2}(\delta(\omega - \xi) + \delta(\omega + \xi))$. Insertando $P(\omega)$ en la ecuación (24) se obtienen cuatro términos de los que tres pueden despreciarse al suponer que $n + m + 1$ tiene un valor elevado. El término restante pasa a ser

$$\begin{aligned}
 R_{n,m}[\lambda] &\approx \frac{1}{8\pi} \hat{\kappa}_{n-m} \left(\xi - \frac{\pi}{2}(n+m+1) \right) \cos \left(\xi \lambda - \frac{\pi}{2}(n-m) \right) \\
 &= \frac{1}{8\pi} \hat{v} \left(\xi - \pi \left(n + \frac{1}{2} \right) \right) \hat{v} \left(\xi - \pi \left(m + \frac{1}{2} \right) \right) \cos \left(\xi \lambda - \frac{\pi}{2}(n-m) \right).
 \end{aligned} \tag{25}$$

La ecuación (25) proporciona un medio eficaz para determinar la matriz de covarianza de subbanda $R_{n,m}$. Una muestra de subbanda $\langle x, w_{p,0} \rangle$ puede predecirse de manera fiable mediante una colección de muestras de subbanda circundantes $\{\langle x, w_{n,k} \rangle : (n,k) \in B\}$ que se supone que están muy influenciadas por la frecuencia considerada. La frecuencia absoluta ξ puede expresarse en términos relativos con respecto a la frecuencia central $\pi(p+\frac{1}{2})$ de una subbanda, como $\xi = \pi(p+\frac{1}{2}+f)$, donde p es el índice de subbanda de la subbanda que comprende la frecuencia ξ , y donde f es un parámetro de frecuencia normalizada que toma valores comprendidos entre -0,5 y +0,5, y que indica la posición de la frecuencia ξ con respecto a la frecuencia central de la subbanda p . Tras haberse determinado la matriz de covarianza de subbanda $R_{n,m}$, los coeficientes de predictor $c_m[l]$ que se aplican a una muestra de subbanda en la subbanda m en el índice de muestra l para estimar una muestra de subbanda en la subbanda n en el índice de muestra k se obtienen resolviendo las ecuaciones normales (7), que en este caso pueden escribirse como

$$\sum_{(m,l) \in B} R_{n,m}[k-l] c_m[l] = R_{n,p}[k], \quad (n,k) \in B. \tag{26}$$

En la ecuación (26), el conjunto B describe el soporte de máscara de predicción como se ilustra, por ejemplo, en la figura 2. En otras palabras, el conjunto B identifica las subbandas m y los índices de muestra l que se usan para predecir una muestra objetivo.

A continuación se proporcionan a modo de ejemplo soluciones de las ecuaciones normales (26) para diferentes soportes de máscara de predicción (como los mostrados en la figura 2). El ejemplo de un predictor causal dentro de banda y de segundo orden se obtiene seleccionando el soporte de máscara de predicción $B = \{(p,-1), (p,-2)\}$. Este soporte de máscara de predicción corresponde a la máscara 202 de predicción de la figura 2. Las ecuaciones normales (26) para esta predicción en dos etapas, que usa la aproximación de la ecuación (25), pasan a ser

$$\hat{v} \left(\xi - \pi \left(p + \frac{1}{2} \right) \right)^2 \sum_{l=-1, -2} \cos(\xi(k-l)) c_p[l] = \hat{v} \left(\xi - \pi \left(p + \frac{1}{2} \right) \right)^2 \cos(-\xi k), \quad k = -1, -2. \tag{27}$$

Una solución de la ecuación (27) viene dada por $c_p[-1] = 2\cos(\xi)$, $c_p[-2] = -1$ y es única siempre que la frecuencia $\xi = \pi(p+\frac{1}{2}+f)$ no se elija de manera que $\hat{v}(f) = 0$. Se observa que el valor medio del error cuadrático de predicción según la ecuación (6) disminuye. Por consiguiente, la predicción sinusoidal es perfecta, hasta la aproximación de la ecuación (25). La propiedad de invariación con respecto a los desplazamientos de frecuencia se ilustra aquí mediante el hecho de que, usando la definición $\xi = \pi(p+\frac{1}{2}+f)$, el coeficiente de predicción $c_p[-1]$ puede reescribirse en lo que respecta a la frecuencia normalizada f , como $c_p[-1] = -2(-1)^p \sin(\pi f)$. Esto significa que los coeficientes de predicción solo dependen de la frecuencia normalizada f dentro de una subbanda particular. Sin embargo, los valores absolutos de los coeficientes de predicción son independientes del índice de subbanda p .

Como se ha descrito anteriormente en la figura 4, la predicción dentro de banda tiene ciertos inconvenientes con respecto a los artefactos de distorsión en la conformación de ruido. El siguiente ejemplo se refiere al comportamiento mejorado ilustrado en la figura 5. Una predicción de banda cruzada causal como la descrita en el presente documento se obtiene seleccionando el soporte de máscara de predicción $B = \{(p-1,-1), (p,-1), (p+1,-1)\}$, que solo requiere un intervalo de tiempo anterior en lugar de dos, y que lleva a cabo una conformación de ruido con menos contribuciones de frecuencia de distorsión que la máscara de predicción clásica 202 del primer ejemplo. El soporte de máscara de predicción $B = \{(p-1,-1), (p,-1), (p+1,-1)\}$ corresponde a la máscara de predicción 203 de la figura 2. Las ecuaciones normales (26) basadas en la aproximación de la ecuación (25) se reducen en este caso a dos ecuaciones para los tres coeficientes no conocidos $c_m[-1]$, $m = p-1, p, p+1$,

$$\left\{ \begin{aligned} &\hat{v}(\pi f) c_p[-1] = (-1)^{p+1} \hat{v}(\pi f) \sin(\pi f) \\ &\hat{v}(\pi(f+1)) c_{p-1}[-1] - \hat{v}(\pi(f-1)) c_{p+1}[-1] = (-1)^p \hat{v}(\pi f) \cos(\pi f) \end{aligned} \right\}. \tag{28}$$

Se observa que cualquier solución de las ecuaciones (28) produce un valor medio decreciente del error cuadrático de predicción según la ecuación (6). Una posible estrategia para seleccionar una solución entre el

infinito número de soluciones de las ecuaciones (28) es minimizar la suma de cuadrados de los coeficientes de predicción. Esto hace que los coeficientes se obtengan de la siguiente manera:

$$\left\{ \begin{array}{l} c_{p-1}[-1] = \frac{(-1)^p \hat{v}(\pi f) \hat{v}(\pi(f+1)) \cos(\pi f)}{\hat{v}(\pi(f-1))^2 + \hat{v}(\pi(f+1))^2} \\ c_p[-1] = (-1)^{p+1} \sin(\pi f) \\ c_{p+1}[-1] = \frac{(-1)^{p+1} \hat{v}(\pi f) \hat{v}(\pi(f-1)) \cos(\pi f)}{\hat{v}(\pi(f-1))^2 + \hat{v}(\pi(f+1))^2} \end{array} \right\}. \quad (29)$$

5

A partir de las fórmulas (29) resulta evidente que los coeficientes de predicción solo dependen de la frecuencia normalizada f con respecto al punto central de la subbanda objetivo p , y dependen además de la paridad de la subbanda objetivo p .

- 10 Usando el mismo soporte de máscara de predicción $B = \{(p-1, -1), (p, -1), (p+1, -1)\}$ para predecir las tres muestras de subbanda $\langle x, w_{m,0} \rangle$ para $m = p-1, p, p+1$, como se ilustra mediante la máscara 204 de predicción de la figura 2, se obtiene una matriz de predicción 3×3 . Tras la introducción de una estrategia más natural para evitar la ambigüedad en las ecuaciones normales, principalmente insertando el modelo sinusoidal relajado $r(\tau) = \exp(-\varepsilon|\tau|)\cos(\xi\tau)$ correspondiente a $P(\omega) = \varepsilon((\varepsilon^2 + (\omega - \xi)^2)^{-1} + (\varepsilon^2 + (\omega + \xi)^2)^{-1})$, los cálculos numéricos obtienen
- 15 los elementos de la matriz de predicción 3×3 de la figura 3. Los elementos de la matriz de predicción se muestran en función de la frecuencia normalizada $f \in [-1/2, 1/2]$ en caso de un solapamiento $K = 2$ con una función de ventana sinusoidal $v(t) = \cos(\pi t / 2)$ y en caso de una subbanda impar p .

- 20 De este modo, se ha demostrado que los modelos de señal $x(t)$ pueden usarse para describir características subyacentes de la señal de audio de entrada que va a codificarse. Parámetros que describen la función de autocorrelación $r(\tau)$ pueden transmitirse a un decodificador 100, permitiéndose así que el decodificador 100 calcule el predictor a partir de los parámetros transmitidos y conociendo el modelo de señal $x(t)$. Se ha demostrado que en bancos de filtros modulados pueden obtenerse medios eficaces para determinar la matriz de covarianza de subbanda del modelo de señal y para resolver las ecuaciones normales para determinar los
- 25 coeficientes de predictor. En particular, se ha demostrado que los coeficientes de predictor resultantes son invariables con respecto a desplazamientos de subbanda y solo dependen normalmente de una frecuencia normalizada relativa a una subbanda particular. Como resultado, pueden proporcionarse tablas de consulta predeterminadas (como se ilustra, por ejemplo, en la figura 3) que permiten la determinación de coeficientes de predictor conociendo una frecuencia normalizada f que es independiente (aparte de un valor de paridad) del
- 30 índice de subbanda p para el que se determinan los coeficientes de predictor.

A continuación se describe en mayor detalle una predicción basada en un modelo periódico que usa, por ejemplo, una única frecuencia fundamental Ω . La función de autocorrelación $r(\tau)$ de tal modelo periódico viene dada por la ecuación (13). El espectro de líneas o PSD equivalente viene dado por

35

$$P(\omega) = \Omega \sum_{q \in \mathbb{Z}} \delta(\omega - q\Omega). \quad (30)$$

- 40 Cuando el periodo T del modelo periódico es suficientemente pequeño, por ejemplo $T \leq 1$, la frecuencia fundamental $\Omega = 2\pi/T$ es suficientemente grande como para permitir la aplicación de un modelo sinusoidal como el obtenido anteriormente usando la frecuencia parcial $\xi = q\Omega$ más cercana a la frecuencia central $\pi(p+1/2)$ de la subbanda p de la muestra de subbanda objetivo que va a predecirse. Esto significa que las señales periódicas que presentan un periodo pequeño T , es decir, un periodo que es pequeño con respecto al lapso de tiempo del banco de filtros, pueden modelarse y predecirse correctamente usando el modelo sinusoidal descrito anteriormente.

45

Quando el periodo T es suficientemente grande en comparación con la duración K de la ventana del banco de filtros $v(t)$, el predictor se reduce hasta una aproximación de un retardo en T . Como se mostrará, los coeficientes de este predictor pueden leerse directamente de la función de correlación cruzada de forma de onda dada por la ecuación (19).

50

La introducción del modelo según la ecuación (13) en la ecuación (22) da lugar a

$$R_{n,m}[\lambda] = \sum_{q \in \mathbb{Z}} U_{n,m}(qT - \lambda), \quad (31)$$

Una observación importante es que si $T \geq 2K$, entonces un término, a lo sumo, de la ecuación (31) es distinto de cero para cada λ ya que $U_{n,m}(\tau) = 0$ para $|\tau| > K$. Eligiendo un soporte de máscara de predicción $B = I \times J$ con un diámetro de intervalo de tiempo $D = |J| \leq T - K$ se observa que $(n, k), (m, l) \in B$ implica que $|k - l| \leq T - K$, y, por lo tanto, el único término de la ecuación (31) es aquél para $q = 0$. Se deduce que $R_{n,m}[k - l] = U_{n,m}(k - l)$, que es el producto interno de formas de onda ortogonales y que es nulo a no ser que $n = m$ y $k = l$. Con todo, las ecuaciones (7) normales pasan a ser

$$c_n[k] = R_{n,p}[k], \quad (n, k) \in B. \quad (32)$$

El soporte de máscara de predicción puede elegirse para que esté centrado alrededor de $k = k_0 \approx -T$, en cuyo caso el lado derecho de la ecuación (32) obtiene su única contribución de $q = -1$. Entonces, los coeficientes vienen dados por

$$c_n[k] = U_{n,p}[-k - T], \quad (n, k) \in B, \quad (33)$$

donde puede insertarse la expresión explícita de la ecuación (19). La geometría del soporte de máscara de predicción para este caso podría tener la apariencia del soporte de máscara de predicción de la máscara 205 de predicción de la figura 2. El valor medio del error cuadrático de predicción proporcionado por la ecuación (6) es igual a la norma cuadrática de la proyección de $u_p(t + T)$ en el espacio delimitado por el complemento de las formas de onda aproximadas $w_{m,l}(t)$, $(m, l) \notin B$.

En vista de lo anterior, en el presente documento se indica que la muestra de subbanda $\langle x, w_{p,0} \rangle$ (de la subbanda p y en el índice de tiempo 0) puede predecirse usando un soporte de máscara de predicción adecuado B centrado alrededor de $(p, -T)$ con un diámetro de tiempo aproximadamente igual a T . Las ecuaciones normales pueden resolverse para cada valor de T y p . En otras palabras, para cada periodicidad T de una señal de audio de entrada y para cada subbanda p , los coeficientes de predicción para un soporte de máscara de predicción dado B pueden determinarse usando las ecuaciones (33) normales.

Con un gran número de subbandas p y una gran variedad de periodos T , una tabulación directa de todos los coeficientes de predictor no resulta práctica. Pero de manera similar al modelo sinusoidal, la estructura de modulación del banco de filtros ofrece una reducción significativa del tamaño de tabla necesario, a través de la propiedad de invariación con respecto a los desplazamientos de frecuencia. Normalmente bastará con estudiar el modelo armónico desplazado con un parámetro de desplazamiento $-1/2 < \theta \leq 1/2$ centrado alrededor del centro de una subbanda p , es decir, centrado alrededor de $\pi(p+1/2)$, definido por el subconjunto $S(\theta)$ de frecuencias positivas entre la colección de frecuencias $\pi(p+1/2) + (q+\theta)\Omega$, $q \in \mathbb{Z}$,

$$P(\omega) = \Omega \sum_{\xi \in S(\theta)} (\delta(\omega - \xi) + \delta(\omega + \xi)). \quad (34)$$

De hecho, dado T y un índice de subbanda suficientemente grande p , el modelo periódico según la ecuación (30) puede recuperarse con una buena aproximación mediante el modelo desplazado según la ecuación (34) y una elección adecuada del parámetro de desplazamiento θ . La inserción de la ecuación (34) en la ecuación (24) con $n = p + v$ y $m = p + \mu$ (donde v y μ definen los índices de subbanda alrededor de la subbanda p del soporte de máscara de predicción) y las manipulaciones basadas en el análisis de Fourier dan lugar a la siguiente expresión para la matriz de covarianza,

$$R_{p+v, p+\mu}[\lambda] \approx \frac{(-1)^{p\lambda}}{2} \sum_{l \in \mathbb{Z}} \kappa_{v-\mu}(Tl - \lambda) \cos \left(2\pi l \theta + \frac{\pi}{2} ((v + \mu)(\lambda - Tl) + \lambda - v + \mu) \right). \quad (35)$$

Como puede observarse, la expresión (35) depende del índice de subbanda objetivo p solamente a través del factor $(-1)^{p\lambda}$. En el caso de un gran periodo T y de una pequeña demora de tiempo λ , solo el término para $l = 0$ contribuye en la expresión (35) y se observa de nuevo que la matriz de covarianza es la matriz identidad. El lado derecho de las ecuaciones normales (26) de un soporte de máscara de predicción adecuado B centrado alrededor de $(p, -T)$ hace que los coeficientes de predicción se obtengan directamente como

$$c_{p+v}[k] = \frac{(-1)^{p\lambda}}{2} \kappa_v(-T - k) \cos \left(-2\pi \theta + \frac{\pi}{2} (v(k + T) + k - v) \right), \quad (p + v, k) \in B. \quad (36)$$

Esto recupera la contribución del primer término de las ecuaciones (19) a (33) con la elección canónica de desplazamiento $\theta = -\pi(p+1/2)/\Omega$.

La ecuación (36) permite determinar los coeficientes de predicción $c_{p+v}[k]$ para una subbanda $(p + v)$ en un índice de tiempo k , donde la muestra que va a predecirse es una muestra de la subbanda p en el índice de tiempo 0. Como puede observarse en la ecuación (36), los coeficientes de predicción $c_{p+v}[k]$ dependen del índice de subbanda objetivo p solamente a través del factor $(-1)^{pk}$, que afecta al signo del coeficiente de predicción. Sin embargo, el valor absoluto del coeficiente de predicción es independiente del índice de subbanda objetivo p . Por otro lado, el coeficiente de predicción $c_{p+v}[k]$ depende de la periodicidad T y del parámetro de desplazamiento θ . Además, el coeficiente de predicción $c_{p+v}[k]$ depende de v y k , es decir, del soporte de máscara de predicción B , usados para predecir la muestra objetivo en la subbanda objetivo p .

En el presente documento, se propone proporcionar una tabla de consulta que permita consultar un conjunto de coeficientes de predicción $c_{p+v}[k]$ para un soporte de máscara de predicción predeterminado B . Para un soporte de máscara de predicción dado B , la tabla de consulta proporciona un conjunto de coeficientes de predicción $c_{p+v}[k]$ para un conjunto predeterminado de valores de la periodicidad T y valores del parámetro de desplazamiento θ . Para limitar el número de entradas de tabla de consulta, debe limitarse el número de valores predeterminados de la periodicidad T y el número de valores predeterminados del parámetro de desplazamiento θ . Como puede observarse en la expresión (36), un incremento de cuantificación adecuado para los valores predeterminados de la periodicidad T y del parámetro de desplazamiento θ debe depender de la periodicidad T . En particular, puede observarse que para periodicidades T relativamente grandes (relativas a la duración K de la función de ventana), pueden usarse etapas de cuantificación relativamente grandes para la periodicidad T y para el parámetro de desplazamiento θ . En el otro extremo, para periodicidades T relativamente pequeñas que tienden a cero, solo hay que tener en cuenta una contribución sinusoidal, de modo que la periodicidad T pierde importancia. Por otro lado, las fórmulas de la predicción sinusoidal según la ecuación (29) requieren que el desplazamiento de frecuencia absoluta normalizado $f = \Omega\theta/\pi = 1/2\theta/T$ varíe lentamente, de modo que el valor de incremento de cuantificación del parámetro de desplazamiento θ deba escalarse en función de la periodicidad T .

Con todo, en el presente documento se propone usar una cuantificación uniforme de la periodicidad T con un valor de incremento fijo. Sin embargo, el parámetro de desplazamiento θ también puede cuantificarse de manera uniforme con un valor de incremento que es proporcional a $\min(T, A)$, donde el valor de A depende de las especificaciones de la función de ventana del banco de filtros. Además, para $T < 2$, el intervalo de los parámetros de desplazamiento θ puede limitarse a $|\theta| \leq \min(CT, 1/2)$ para una C constante, que refleja un límite en los desplazamientos de frecuencia absolutos f .

La figura 6a ilustra un ejemplo de una cuadrícula de cuantificación resultante en el plano (T, θ) para $A = 2$. Solamente en el intervalo intermedio comprendido entre $0,25 \leq T \leq 1,5$ se tiene en cuenta la dependencia bidimensional total, mientras que las parametrizaciones esencialmente unidimensionales proporcionadas por las ecuaciones (29) y las ecuaciones (36) pueden usarse para el intervalo restante de interés. En particular, para periodicidades T que tienden a cero (por ejemplo, $T < 0,25$) la predicción basada en modelo periódico corresponde sustancialmente a una predicción basada en un modelo sinusoidal, y los coeficientes de predicción pueden determinarse usando las fórmulas (29). Por otro lado, para periodicidades T que superan sustancialmente la duración de ventana K (por ejemplo, $T > 1,5$) el conjunto de coeficientes de predicción $c_{p+v}[k]$ que usa predicción basada en modelo periódico puede determinarse usando la ecuación (36). Esta ecuación puede reinterpretarse mediante la sustitución $\theta = \varphi + 1/4Tv$. Se observa que

$$c_{p+v}[k] = \frac{(-1)^{pk}}{2} \kappa_v(-T - k) \cos\left(-2\pi\varphi + \frac{\pi}{2}((v+1)k - v)\right), \quad (p+v, k) \in B. \quad (37)$$

Al darle a φ la función dada al parámetro θ en la tabulación, se obtiene una estructura esencialmente separable en el plano (T, φ) equivalente. En lo que respecta a los cambios de signo en función de los índices de subbanda y de intervalo de tiempo, la dependencia de T viene dada por un primer factor que varía lentamente, y la dependencia de φ viene dada por un segundo factor periódico de la ecuación (37).

El parámetro de desviación modificado φ puede interpretarse como el desplazamiento de la serie armónica en unidades de la frecuencia fundamental medida desde el punto central de los puntos centrales de las celdas fuente y objetivo.

Resulta ventajoso mantener esta parametrización modificada (T, φ) para todos los valores de periodicidades T ya que las simetrías de la ecuación (37) que resultan evidentes con respecto a los cambios de signo simultáneos de φ y v se cumplirán, por lo general, y pueden aprovecharse para reducir los tamaños de tabla.

Como se ha indicado anteriormente, la figura 6a muestra una cuadrícula de cuantificación bidimensional subyacente a los datos tabulados para un cálculo de predictor basado en modelo periódico en un banco de filtros modulado por coseno. El modelo de señal es el de una señal con periodo $T/602$, medida en unidades del salto de

tiempo del banco de filtros. De manera equivalente, el modelo comprende las líneas de frecuencia de los múltiplos enteros, también conocidas como parciales, de la frecuencia fundamental correspondiente al periodo T . Para cada subbanda objetivo, el parámetro de desplazamiento θ 601 indica la distancia del parcial más cercano a la frecuencia central medida en unidades de la frecuencia fundamental Q . El parámetro de desplazamiento θ 601 tiene un valor entre $-0,5$ y $0,5$. Las cruces negras 603 de la figura 6a ilustran una densidad apropiada de puntos de cuantificación para la tabulación de predictores con una alta ganancia de predicción basada en el modelo periódico. Para periodos T grandes (por ejemplo, $T > 2$), la cuadrícula es uniforme. Normalmente se necesita una mayor densidad en el parámetro de desplazamiento θ cuando el periodo T disminuye. Sin embargo, en la región fuera de las líneas 604, la distancia θ es mayor que una celda de frecuencia del banco de filtros, de modo que puede despreciarse la mayor parte de los puntos de la cuadrícula en esta región. El polígono 605 delimita una región que es suficiente para una tabulación total. Además de las líneas escalonadas ligeramente fuera de las líneas 604, se introducen bordes en $T = 0,25$ y $T = 1,5$. Esto se permite debido a que pequeños periodos 602 pueden tratarse como sinusoides individuales y a que los predictores 602 de largos periodos pueden aproximarse mediante tablas esencialmente unidimensionales que dependen principalmente del parámetro de desviación θ (o del parámetro de desviación modificado ϕ). En la forma de realización ilustrada en la figura 6a, el soporte de máscara de predicción es normalmente similar a la máscara 205 de predicción de la figura 2 para grandes periodos T .

La figura 6b ilustra una predicción basada en un modelo periódico en el caso de periodos T relativamente grandes y en el caso de periodos T relativamente pequeños. En el diagrama superior puede observarse que para periodos T grandes, es decir, para frecuencias 613 fundamentales relativamente pequeñas Ω , la función 612 de ventana del banco de filtros captura un número relativamente grande de líneas o pulsos 616 de Dirac de la PSD de la señal periódica. Los pulsos 616 de Dirac están situados en las frecuencias $\omega = q\Omega$, con $q \in \mathbb{Z}$. Las frecuencias centrales de las subbandas del banco de filtros están ubicadas en las frecuencias $\omega = \pi(p + \frac{1}{2})$ con $p \in \mathbb{Z}$. Para una subbanda dada p , la asignación de frecuencia del pulso 616 con frecuencia $\omega = q\Omega$ más cercana a la frecuencia central de la subbanda dada $\omega = \pi(p + \frac{1}{2})$ puede describirse en términos relativos como $q\Omega - \omega = \pi(p + \frac{1}{2}) + \Theta\Omega$, donde el parámetro de desviación Θ oscila entre $-0,5$ y $+0,5$. De este modo, el término $\Theta\Omega$ refleja la distancia (en frecuencia) desde la frecuencia central $\omega = \pi(p + \frac{1}{2})$ hasta la componente de frecuencia más cercana 616 del modelo armónico. Esto se ilustra en el diagrama superior de la figura 6b, donde la frecuencia 617 central es $\omega = \pi(p + \frac{1}{2})$ y donde la distancia 618 $\Theta\Omega$ se ilustra para el caso de un periodo relativamente grande T . Puede observarse que el parámetro de desplazamiento Θ permite describir toda la serie armónica vista desde la perspectiva del centro de la subbanda p .

El diagrama inferior de la figura 6b ilustra el caso para periodos T relativamente pequeños, es decir para frecuencias 623 fundamentales Ω relativamente grandes, en concreto frecuencias fundamentales 623 que son mayores que el ancho de la ventana 612. Puede observarse que, en tales casos, una función 612 de ventana solo puede comprender un único pulso 626 de la señal periódica, de manera que la señal puede verse como una señal sinusoidal dentro de la ventana 612. Esto significa que en periodos T relativamente pequeños, el esquema de predicción basado en un modelo periódico converge hacia un esquema de predicción basado en un modelo sinusoidal.

La figura 6b ilustra también máscaras 611, 621 de predicción de ejemplo que pueden usarse en el esquema de predicción basado en un modelo periódico y en el esquema de predicción basado en un modelo sinusoidal, respectivamente. La máscara 611 de predicción usada en el esquema de predicción basado en un modelo periódico puede corresponder a la máscara 205 de predicción de la figura 2 y puede comprender el soporte 614 de máscara de predicción para estimar la muestra de subbanda 615 objetivo. La máscara 621 de predicción usada en el esquema de predicción basado en un modelo sinusoidal puede corresponder a la máscara 203 de predicción de la figura 2 y puede comprender el soporte 624 de máscara de predicción para estimar la muestra de subbanda objetivo 625.

La figura 7a ilustra un procedimiento 700 de codificación de ejemplo que utiliza predicción de subbanda basada en un modelo periódico (que comprende, por ejemplo, una única frecuencia fundamental Ω). Se considera una trama de una señal de audio de entrada. Para esta trama puede determinarse una periodicidad T o una frecuencia fundamental Ω (etapa 701). El codificador de audio puede comprender los elementos del descodificador 100 ilustrado en la figura 1; en particular, el codificador de audio puede comprender un calculador 105 de predictor y un predictor 103 de subbanda. La periodicidad T o la frecuencia fundamental Ω pueden determinarse de manera que se reduzca (por ejemplo, se minimice) el valor medio de las señales 111 de subbanda de error cuadrático de predicción según la ecuación (6). A modo de ejemplo, el codificador de audio puede aplicar un enfoque elemental que determina las señales 111 de subbanda de error de predicción usando diferentes frecuencias fundamentales Ω y que determina la frecuencia fundamental Ω para la que se reduce (por ejemplo, se minimiza) el valor medio de las señales 111 de subbanda de error cuadrático de predicción. El procedimiento prosigue con la cuantificación de las señales 111 de subbanda de error de predicción resultantes (etapa 702). Además, el procedimiento comprende la etapa 703 de generar un flujo de bits que comprende información que indica la frecuencia fundamental determinada Ω y las señales 111 de subbanda de error de

predicción cuantificadas.

Cuando se determina la frecuencia fundamental Ω en la etapa 701, el codificador de audio puede usar las ecuaciones (36) y/o (29) para determinar los coeficientes de predicción para una frecuencia fundamental particular Ω . El conjunto de posibles frecuencias fundamentales Ω puede limitarse por el número de bits que están disponibles para la transmisión de la información que indica la frecuencia fundamental determinada Ω .

Debe observarse que el sistema de codificación de audio puede usar un modelo predeterminado (por ejemplo, un modelo periódico que comprende una única frecuencia fundamental Ω o cualquier otro de los modelos proporcionados en el presente documento) y/o una máscara 202, 203, 204, 205 de predicción predeterminada. Por otro lado, el sistema de codificación de audio puede estar dotado de grados de libertad adicionales al permitir que el codificador de audio determine un modelo apropiado y/o una máscara de predicción apropiada para una señal de audio que va a codificarse. La información relacionada con el modelo seleccionado y/o con la máscara de predicción seleccionada se codifican después en el flujo de bits y se proporcionan al descodificador 100 correspondiente.

La figura 7b ilustra un procedimiento 710 de ejemplo para descodificar una señal de audio que se ha codificado usando predicción basada en modelo. Se supone que el descodificador 100 conoce el modelo de señal y la máscara de predicción usados por el codificador (ya sea a través del flujo de bits recibido o debido a configuraciones predeterminadas). Además, con fines ilustrativos, se supone que se ha usado un modelo de predicción periódico. El descodificador 100 extrae del flujo de bits recibido información referente a la frecuencia fundamental Ω (etapa 711). Usando la información referente a la frecuencia fundamental Ω , el descodificador 100 puede determinar la periodicidad T . La frecuencia fundamental Ω y/o la periodicidad T pueden usarse para determinar un conjunto de coeficientes de predicción para los diferentes predictores de subbanda (etapa 712). Los predictores de subbanda pueden usarse para determinar señales de subbanda estimadas (etapa 713) que se combinan (etapa 714) con las señales 111 de subbanda de error de predicción descuantificadas para obtener las señales 113 de subbanda descodificadas. Las señales 113 de subbanda descodificadas pueden filtrarse (etapa 715) usando un banco de filtros 102 de síntesis, obteniéndose así la señal 114 de audio descodificada del dominio del tiempo.

El calculador 105 de predictor puede usar las ecuaciones (36) y/o (29) para determinar los coeficientes de predicción de los predictores 103 de subbanda en función de la información recibida referente a la frecuencia fundamental Ω (etapa 712). Esto puede llevarse a cabo de manera eficaz usando una tabla de consulta, como se ilustra en las figuras 6a y 3. A modo de ejemplo, el calculador 105 de predictor puede determinar la periodicidad T y determinar si la periodicidad está por debajo de un umbral inferior predeterminado (por ejemplo, $T = 0,25$). Si este es el caso, se usa un esquema de predicción basado en un modelo sinusoidal. Esto significa que en función de la frecuencia fundamental recibida Ω se determinan subbandas p que comprenden un múltiplo $\omega = q\Omega$, con $q \in \mathbb{Z}$ de la frecuencia fundamental. Después se determina la frecuencia normalizada f usando la relación $\xi = \pi(p + \frac{1}{2}f)$, donde la frecuencia ξ corresponde al múltiplo $\omega = q\Omega$ que está en la subbanda p . El calculador 105 de predictor puede usar después la ecuación (29) o una tabla de consulta precalculada para determinar el conjunto de coeficientes de predicción (usando, por ejemplo, la máscara 203 de predicción de la figura 2 o la máscara 621 de predicción de la figura 6b).

Debe observarse que un conjunto diferente de coeficientes de predicción puede determinarse para cada subbanda. Sin embargo, en caso de un esquema de predicción basado en un modelo sinusoidal solo se determina normalmente un conjunto de coeficientes de predicción para las subbandas p , las cuales están muy influenciadas por un múltiplo $\omega = q\Omega$, con $q \in \mathbb{Z}$ de la frecuencia fundamental. Para las otras subbandas no se determina ningún coeficiente de predicción, lo que significa que las señales 112 de subbanda estimadas para estas otras subbandas son nulas. Para reducir la complejidad computacional del descodificador 100 (y del codificador que usa el mismo calculador 105 de predictor), el calculador 105 de predictor puede usar una tabla de consulta predeterminada que proporciona el conjunto de coeficientes de predicción, sujeto a valores para T y Θ . En particular, el calculador de predictor 105 puede usar una pluralidad de tablas de consulta para una pluralidad de diferentes valores de T . Cada una de la pluralidad de tablas de consulta proporciona un conjunto diferente de coeficientes de predicción para una pluralidad de diferentes valores del parámetro de desplazamiento Θ .

En una implementación práctica, una pluralidad de tablas de consulta puede proporcionarse para diferentes valores del parámetro de periodo T . A modo de ejemplo, las tablas de consulta pueden proporcionarse para valores de T en el intervalo comprendido entre 0,25 y 2,5 (como se ilustra en la figura 6a). Las tablas de consulta pueden proporcionarse para una granularidad predeterminada o valor de incremento de diferentes parámetros de periodo T . En una implementación de ejemplo, el valor de incremento para el parámetro de periodo normalizado T es $1/16$, y diferentes tablas de consulta para los coeficientes de predicción cuantificados se proporcionan para $T=8/32$ hasta $T=80/32$. Por tanto, puede proporcionarse un total de 37 tablas de consulta diferentes. Cada tabla puede proporcionar los coeficientes de predicción cuantificados en función del parámetro de desplazamiento Θ o en función del parámetro de desplazamiento modificado ϕ . Las tablas de consulta para $T=8/32$ hasta $T=80/32$ pueden usarse en un intervalo aumentado en la mitad del valor de incremento, es decir, $[9/32, 81/32]$. Para una

- periodicidad dada que difiere de las periodicidades disponibles, para las que se ha definido una tabla de consulta, puede usarse la tabla de consulta para la periodicidad disponible más cercana. Como se ha indicado anteriormente, en periodos T largos (por ejemplo, en periodos T que superan el periodo para el que se ha definido una tabla de consulta), puede usarse la ecuación (36). Como alternativa, en periodos T que superan los
- 5 periodos para los que se han definido las tablas de consulta, por ejemplo, en periodos $T > 81/32$, el periodo T puede dividirse en un retardo entero T_i y en un retardo residual T_r , de manera que $T = T_i + T_r$. La separación puede ser tal que el retardo residual T_r esté dentro del intervalo para el que puede aplicarse la ecuación (36) y para el que las tablas de consulta están disponibles, por ejemplo dentro del intervalo $[1,5; 2,5]$ o $[49/32, 81/32]$ para el ejemplo anterior. De este modo, los coeficientes de predicción pueden determinarse usando la tabla de
- 10 consulta del retardo residual T_r , y el predictor de subbanda 103 puede utilizar una memoria intermedia 104 de subbanda que se ha retardado con el retardo entero T_i . Por ejemplo, si el periodo es $T=3,7$, el retardo entero puede ser $T_i=2$, seguido de un retardo residual de $T_r=1,7$. El predictor puede aplicarse en función de los coeficientes para $T_r=1,7$ en una memoria intermedia de señales que está retardada, adicionalmente, con $T_i=2$.
- 15 El enfoque de separación se basa en la suposición razonable de que el extractor aproxima un retardo de T en el intervalo de $[1,5; 2,5]$ o $[49/32, 81/32]$. La ventaja del procedimiento de separación en comparación con el uso de la ecuación (36) es que los coeficientes de predicción pueden determinarse en función de operaciones de consulta de tabla eficaces desde el punto de vista computacional.
- 20 Como se ha indicado anteriormente, para periodos cortos ($T < 0,25$) puede usarse la ecuación (29) para determinar los coeficientes de predicción. Como alternativa, puede resultar beneficioso usar las tablas de consulta ya disponibles con el fin de reducir la complejidad computacional. Se observa que el parámetro de desplazamiento modificado φ está comprendido dentro del intervalo $|\varphi| \leq T$ con un valor de incremento de muestreo de $\Delta\varphi = T/32$ (para $T < 0,25$ y para $C=1$, $A=1/2$).
- 25 En el presente documento se propone reutilizar la tabla de consulta para el periodo más bajo, $T=0,25$, escalando el parámetro de desplazamiento modificado φ con T_i/T , donde T_i corresponde al periodo más bajo durante el cual una tabla de consulta está disponible (por ejemplo, $T_i=0,25$). A modo de ejemplo, con $T=0,1$ y $\varphi=0,07$, la tabla para $T=0,25$ puede consultarse con un parámetro de desplazamiento reescalado $\varphi = (0,25/0,1) \cdot 0,07 = 0,175$. De
- 30 este modo, los coeficientes de predicción para periodos cortos (por ejemplo, $T < 0,25$) pueden determinarse también de una manera eficaz desde el punto de vista computacional usando operaciones de consulta a tablas. Además, los requisitos de memoria del predictor pueden reducirse, ya que el número de tablas de consulta puede reducirse.
- 35 En el presente documento se ha descrito un esquema de predicción de subbanda basado en modelo. El esquema de predicción de subbanda basado en modelo permite una descripción eficiente de predictores de subbanda, es decir, una descripción que solo requiere un número de bits relativamente bajo. Como resultado de una descripción eficiente de predictores de subbanda, pueden usarse esquemas de predicción de subbanda cruzada, que reducen los artefactos de distorsión. En términos generales, esto permite proporcionar
- 40 codificadores de audio de baja tasa de bits usando predicción de subbanda.

REIVINDICACIONES

1. Un método en un decodificador para determinar una estimación de una muestra (221) de una señal de subbanda a partir de dos muestras (222) anteriores de la señal de subbanda, en donde la señal de subbanda
5 corresponde a una de una pluralidad de subbandas de una representación en el dominio de subbanda de una señal de audio determinada usando un banco de filtros de análisis que comprende una pluralidad de filtros de análisis, comprendiendo el método:

10 determinar datos de modelo de señal que comprenden un parámetro (613) de modelo;

15 determinar un primer coeficiente de predicción que se aplicará a una primera muestra (222) anterior de la señal de subbanda; en donde el primer coeficiente de predicción se determina como una función del parámetro (613) de modelo usando una tabla de consulta predeterminada y/o una función analítica predeterminada;

20 determinar un segundo coeficiente de predicción a aplicar a una segunda muestra (222) anterior de la señal de subbanda; en donde el segundo coeficiente de predicción se determina como una función del parámetro (613) de modelo usando la tabla de consulta y/o la función analítica; y

25 determinar la estimación de la muestra (221) aplicando el primer coeficiente de predicción a la primera muestra (222) anterior y aplicando el segundo coeficiente de predicción a la segunda muestra (222) anterior;

30 en el que un intervalo de tiempo de la primera muestra (222) anterior precede inmediatamente a un intervalo de tiempo de la muestra (221), en el que un intervalo de tiempo de la segunda muestra (222) anterior precede inmediatamente a un intervalo de tiempo de la primera muestra (222) anterior, y en el que valores absolutos de los coeficientes de predicción primero y segundo son independientes de un índice de subbanda p de la muestra.

35 2. El método de la reivindicación 1, que comprende además extraer el parámetro (613) de modelo de un flujo de bits recibido indicativo del parámetro (613) de modelo y una señal de error de predicción.

40 3. Un sistema (103, 105) de decodificación configurado para determinar una estimación de una muestra (221) de una señal de subbanda a partir de dos muestras (222) anteriores de la señal de subbanda, en donde la señal de subbanda corresponde a una de una pluralidad de subbandas de una representación en el dominio de subbanda de una señal de audio determinada usando un banco de filtros de análisis que comprende una pluralidad de filtros de análisis; en donde el sistema (103, 105) comprende:

45 un calculador (105) de predictor configurado para

50 determinar los datos de modelo de señal que comprende un parámetro (613) de modelo;

55 determinar un primer coeficiente de predicción a aplicar a una primera muestra (222) anterior de la señal de subbanda; en donde el primer coeficiente de predicción se determina como una función del parámetro (613) de modelo usando una tabla de consulta predeterminada y/o una función analítica predeterminada; y

60 determinar un segundo coeficiente de predicción a aplicar a una segunda muestra (222) anterior de la señal de subbanda; en donde el segundo coeficiente de predicción se determina como una función del parámetro (613) de modelo usando la tabla de consulta y/o la función analítica; y

65 un predictor (103) de subbanda configurado para determinar la estimación de la muestra (221) aplicando el primer coeficiente de predicción a la primera muestra (222) anterior y aplicando el segundo coeficiente de predicción a la segunda muestra (222) anterior;

70 en el que un intervalo de tiempo de la primera muestra (222) anterior precede inmediatamente a un intervalo de tiempo de la muestra (221), en el que un intervalo de tiempo de la segunda muestra (222) anterior precede inmediatamente a un intervalo de tiempo de la primera muestra (222) anterior, y en el que valores absolutos de los coeficientes de predicción primero y segundo son independientes de un índice de subbanda p de la muestra.

75 4. Un producto de programa de ordenador que comprende una secuencia de instrucciones que, al ser ejecutadas por un ordenador, provocan que el ordenador realice el método de la reivindicación 1 o la reivindicación 2.

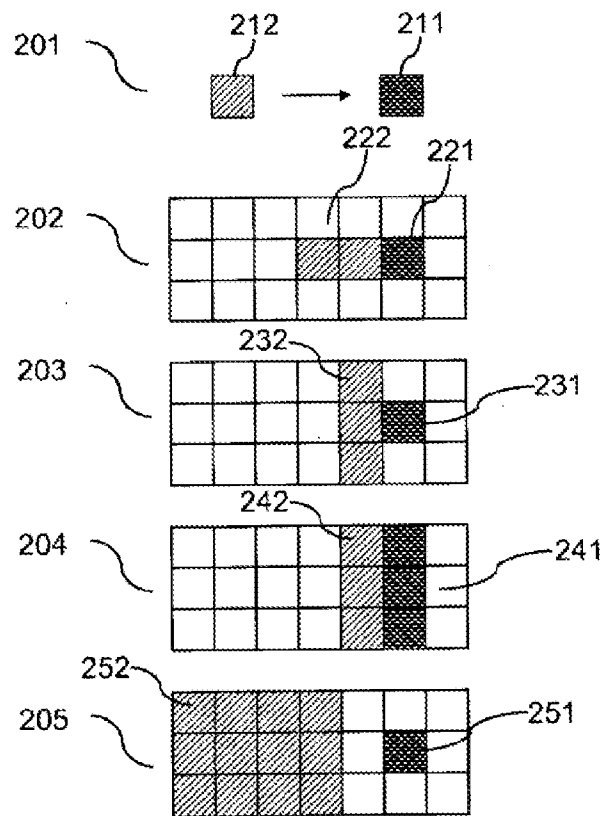
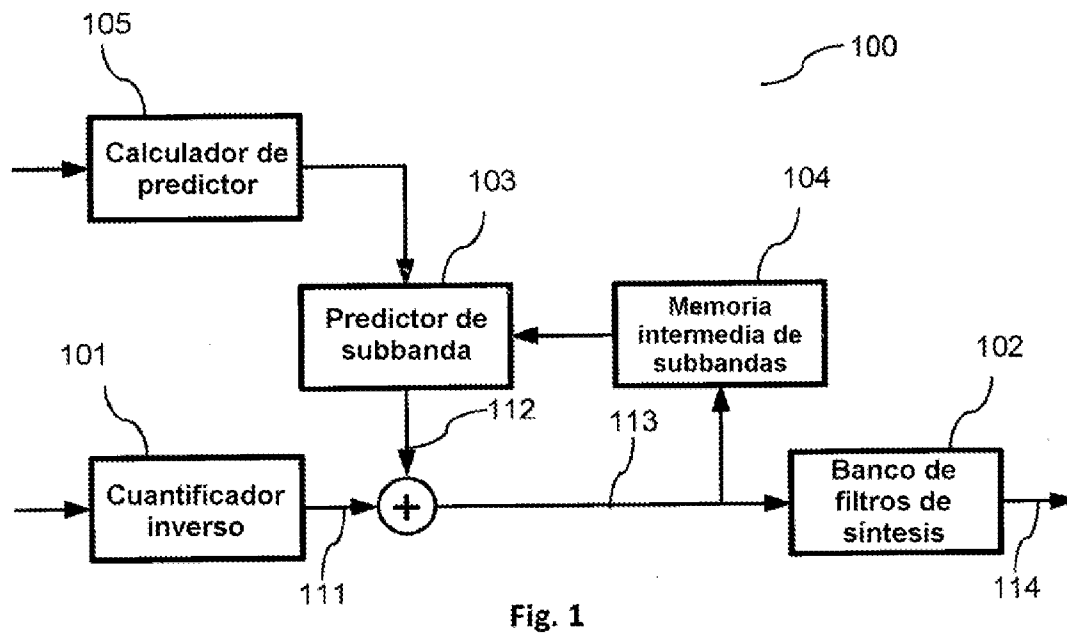


Fig. 2

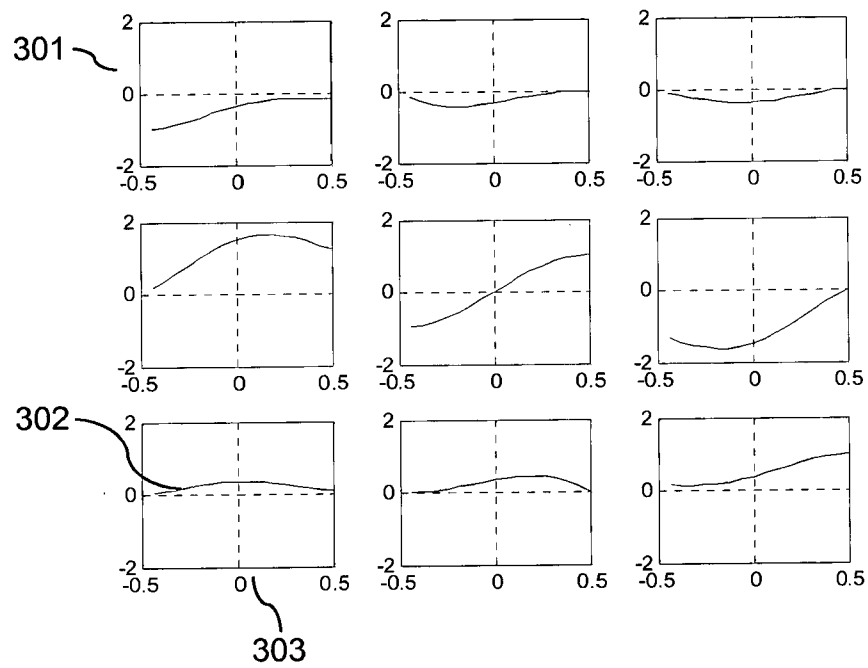


Fig. 3

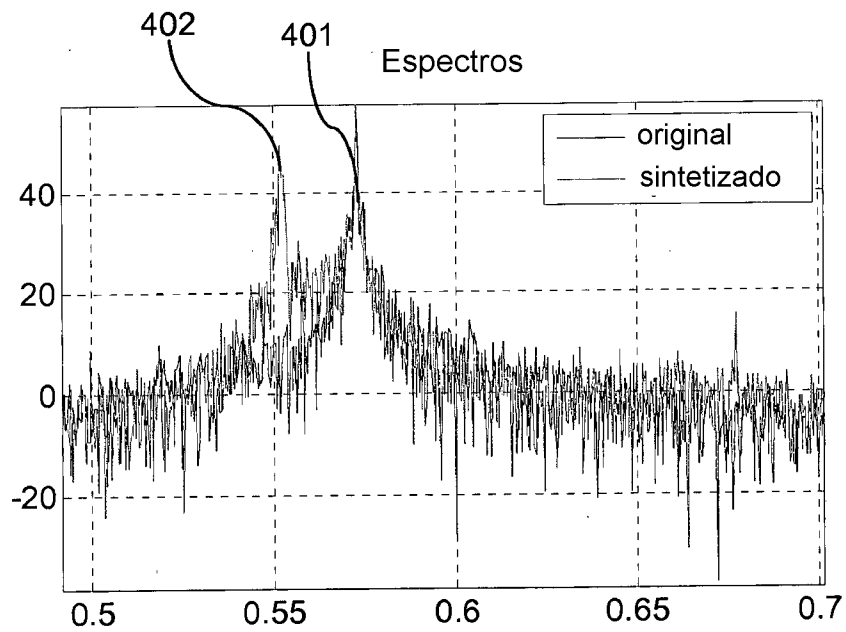


Fig. 4

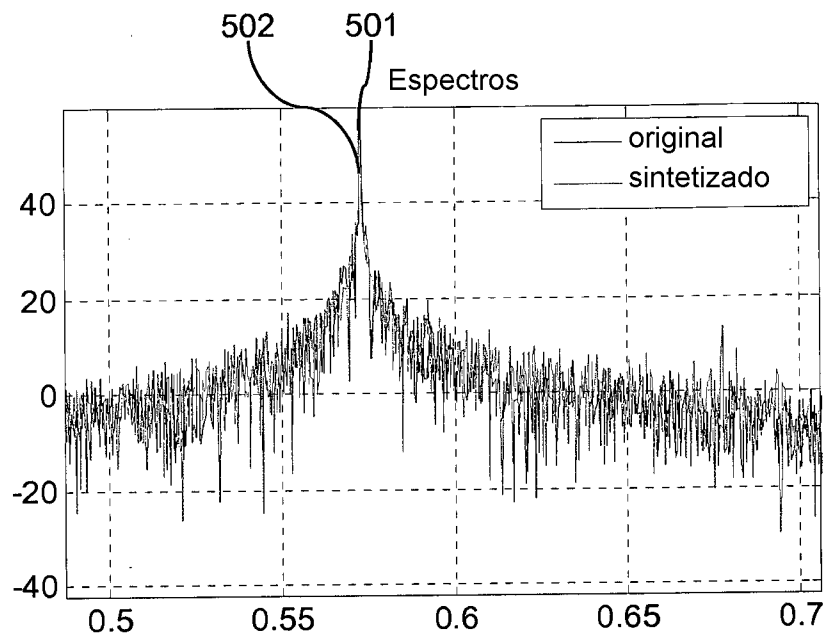


Fig. 5

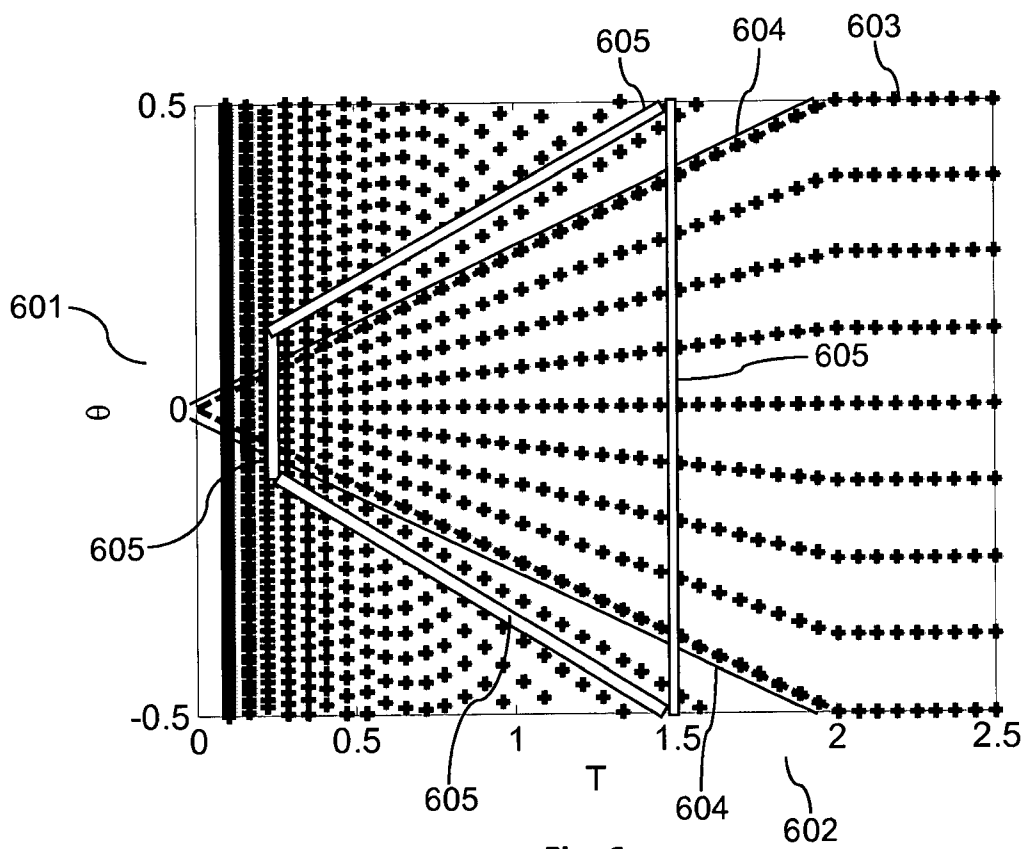


Fig. 6a

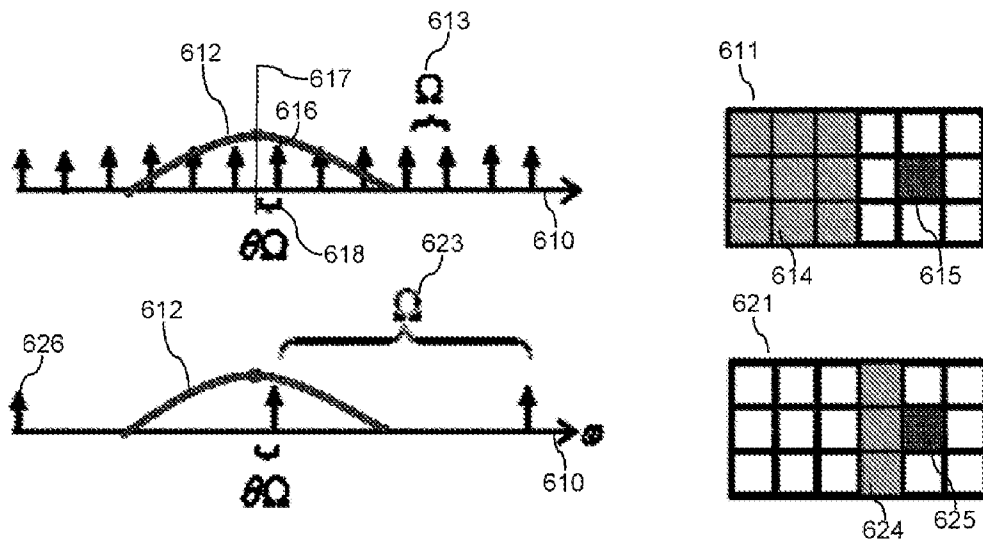


Fig. 6b

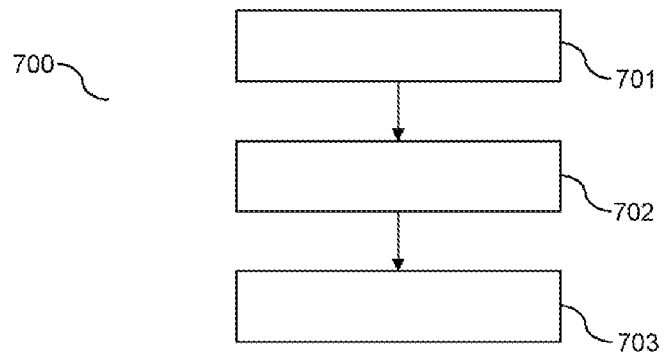


Fig. 7a

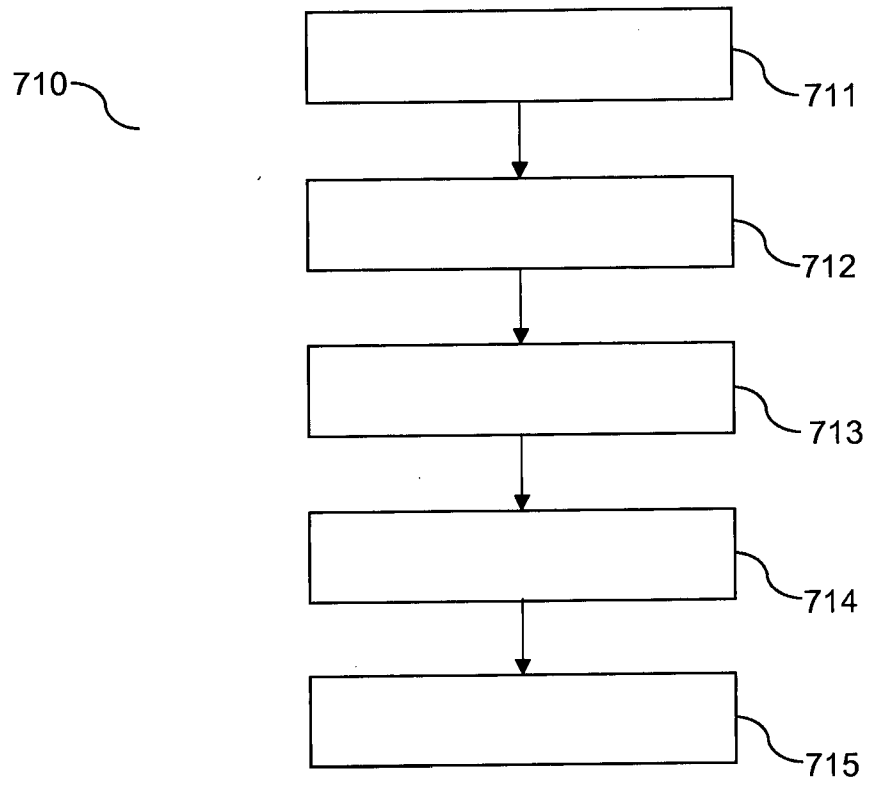


Fig. 7b