



República Federativa do Brasil
Ministério da Economia
Instituto Nacional da Propriedade Industrial

(11) PI 1005300-0 B1



(22) Data do Depósito: 28/01/2010

(45) Data de Concessão: 29/06/2021

(54) Título: CODIFICADOR DE ÁUDIO, DECODIFICADOR DE ÁUDIO, INFORMAÇÕES DE ÁUDIO CODIFICADO E MÉTODOS PARA CODIFICAR E DECODIFICAR UM SINAL DE ÁUDIO COM BASE EM UMA INFORMAÇÃO DE ÁUDIO CODIFICADO E EM UMA INFORMAÇÃO DE ÁUDIO DE ENTRADA.

(51) Int.Cl.: G10L 19/16; G10L 19/022; G10L 19/00; G10L 19/24.

(52) CPC: G10L 19/167; G10L 19/022; G10L 19/0017; G10L 19/24.

(30) Prioridade Unionista: 28/01/2009 US 61/147,887.

(73) Titular(es): FRAUNHOFER - GESELLSCHAFT ZUR FÖRDERUNG DER ANGEWANDTEN TEN FORSCHUNG E.V..

(72) Inventor(es): RALF GEIGER; JÉRÉMIE LECOMTE; MARKUS MULTRUS; MAX NEUENDORF; CHRISTIAN SPITZNER.

(86) Pedido PCT: PCT EP2010050998 de 28/01/2010

(87) Publicação PCT: WO 2010/086373 de 05/08/2010

(85) Data do Início da Fase Nacional: 27/07/2011

(57) Resumo: CODIFICADOR DE ÁUDIO, DECODIFICADOR DE ÁUDIO, INFORMAÇÕES DE ÁUDIO CODIFICADO, MÉTODOS PARA CODIFICAR E DECODIFICAR UM SINAL DE ÁUDIO E PROGRAMA DE COMPUTADOR Um decodificador de áudio para prover uma informação de áudio decodificado com base em uma informação de áudio codificado que compreende um transformador de sinal com base em janela configurado para mapear uma representação de frequência de tempo, que é descrita pelas informações de áudio codificado, para uma representação de domínio de tempo. O transformador de sinal com base em janela é configurado para selecionar uma janela, de uma pluralidade de janelas compreendendo janelas de diferentes inclinações de transição e janelas de diferentes extensões de transformação, com base em uma informação de janela. O decodificador de áudio compreende um seletor de janela configurado para avaliar uma informação de extensão de palavra-chave variável de janela a fim de selecionar uma janela para o processamento de uma determinada parte da representação de frequência de tempo associada a uma determinada estrutura das informações de áudio.

**"CODIFICADOR DE ÁUDIO, DECODIFICADOR DE ÁUDIO,
INFORMAÇÕES DE ÁUDIO CODIFICADO E MÉTODOS PARA CODIFICAR E
DECODIFICAR UM SINAL DE ÁUDIO COM BASE EM UMA INFORMAÇÃO DE ÁUDIO
CODIFICADO E EM UMA INFORMAÇÃO DE ÁUDIO DE ENTRADA"**

5 HISTÓRICO DA INVENÇÃO

As configurações, de acordo com a invenção, referem-se a um codificador de áudio para prover uma informação de áudio codificado, com base em uma informação de áudio de entrada, e a um decodificador de áudio para prover uma informação de áudio decodificado, com base em uma informação de áudio codificado. As configurações adicionais, de acordo com a invenção, referem-se a uma informação de áudio codificado. Ainda, configurações adicionais, de acordo com a invenção, referem-se a um método para prover uma informação de áudio decodificado, com base em uma informação de áudio codificado, e a um método para prover uma informação de áudio codificado, com base em uma informação de áudio de entrada. As configurações adicionais se referem a programas de computador para realizar os métodos inventivos.

Uma configuração da invenção se refere a uma atualização proposta sobre uma sintaxe da corrente de bits da codificação de áudio e fala unificada (USAC).

A seguir, um pouco do histórico da invenção será explicado, a fim de facilitar o entendimento da invenção e suas vantagens. Durante a década passada, foi envidado grande esforço na criação da possibilidade de armazenar e distribuir digitalmente conteúdos de áudio. Uma conquista importante nesse sentido é a definição do padrão internacional ISO/IEC 14496-3. A Parte 3 desse padrão se refere a uma codificação e decodificação de conteúdos de

áudio e a subparte 4 da parte 3 se refere à codificação de áudio geral. O ISO/IEC 14496, parte 3, subparte 4, define um conceito para codificar e decodificar conteúdo de áudio geral. Além disso, foram propostas melhorias adicionais, a fim de melhorar a
5 qualidade e/ou reduzir a taxa de bits necessária.

Entretanto, de acordo com o conceito descrito no referido padrão, um sinal de áudio de domínio de tempo é convertido em uma representação "tempo-frequência". A transformação de domínio de tempo ao domínio de frequência de
10 tempo é tipicamente realizada utilizando blocos de transformação, que também são designados como "estruturas" de amostras de domínio de tempo. Descobriu-se que é vantajoso utilizar estruturas de sobreposição, que são trocadas, por exemplo, pela metade de uma estrutura, pois a sobreposição permite de maneira eficiente evitar
15 (ou pelo menos reduzir) artefatos. Além disso, descobriu-se que deve ser realizado um janelamento, a fim de evitar os artefatos originados a partir desse processamento de estruturas temporariamente limitadas. Também, o janelamento permite uma otimização de um processo de sobreposição-e-adição de subsequentes
20 estruturas trocadas, mas sobrepostas temporariamente.

Entretanto, descobriu-se que é problemático representar eficientemente margens, ou seja, transições agudas ou os denominados transientes dentro do conteúdo de áudio, utilizando janelas de extensão uniforme, pois a energia de uma transição será
25 expandida durante toda a duração de uma janela, que resulta em artefatos audíveis. Da mesma forma, foi proposto fazer desvio entre as janelas de extensões diferentes, de maneira que partes aproximadamente fixas de um conteúdo de áudio sejam codificadas

utilizando janelas longas, e de maneira que partes transicionais (por exemplo, partes que compreendem um transiente) do conteúdo de áudio sejam codificadas utilizando janelas menores.

Entretanto, em um sistema, que permite escolher entre diferentes janelas para transformar um conteúdo de áudio do domínio de tempo ao domínio de frequência de tempo, claro que é necessário sinalizar a um decodificador qual janela deverá ser utilizada para uma decodificação de um conteúdo de áudio codificado de uma determinada estrutura.

Em sistemas convencionais, por exemplo, em um decodificador de áudio de acordo com o padrão internacional ISO/IEC 14496-3, parte 3, subparte 4, um elemento de dados denominado "window_sequence", que indica a sequência de janelas utilizada na estrutura atual, é escrito com dois bits em uma corrente de bits em um denominado elemento de corrente de bits "ics_info". Ao levar em conta a sequência de janelas da estrutura anterior, oito diferentes sequências de janelas são sinalizadas.

Tendo em vista a discussão acima, pode-se observar que uma carga de bits da corrente de bits codificada, que representa uma informação de áudio, é criada pela necessidade de sinalizar o tipo de janela utilizado.

Tendo em vista essa situação, há o desejo de criar um conceito que permita uma sinalização de taxa de bits mais eficiente de um tipo de janela utilizado para uma transformação entre uma representação no domínio de tempo de um conteúdo de áudio e uma representação no domínio de frequência de tempo do conteúdo de áudio.

SUMÁRIO DA INVENÇÃO

Esse problema é solucionado por um codificador de áudio, de acordo com a reivindicação 1, um decodificador de áudio, de acordo com a reivindicação 9, uma informação de áudio codificado, de acordo com a reivindicação 12, um método para
5 prover uma informação de áudio decodificado, de acordo com a reivindicação 14, um método para prover uma informação de áudio codificado, de acordo com a reivindicação 15, e um programa de computador, de acordo com a reivindicação 16.

Uma configuração, de acordo com a invenção, cria
10 um decodificador de áudio para prover uma informação de áudio decodificado com base em uma informação de áudio codificado. O decodificador de áudio compreende um transformador de sinal com base em janela configurado para mapear uma representação tempo-frequência, que é descrita pela informação de áudio codificado,
15 para uma representação no domínio de tempo do conteúdo de áudio. O transformador de sinal com base em janela é configurado para selecionar uma janela de uma pluralidade de janelas que compreendem janelas de diferentes inclinações de transição e janelas de diferentes extensões de transformação, com base em uma
20 informação de janela. O decodificador de áudio compreende um seletor de janela configurado para avaliar uma informação de extensão variável de palavra-código de janela a fim de selecionar uma janela para um processamento de uma determinada parte (por exemplo, estrutura) da representação "tempo-frequência" associada
25 a uma determinada estrutura das informações de áudio.

Essa configuração da invenção tem base na descoberta que uma taxa de bits necessária para armazenar ou transmitir uma informação que indique qual tipo de janela deverá

ser utilizada para transformar uma representação de domínio de frequência de tempo de um conteúdo de áudio em uma representação no domínio de tempo pode ser reduzida utilizando uma informação de extensão de palavra-código variável de janela. Descobriu-se que
5 uma informação de extensão de palavra-código variável de janela é bem adequada, pois as informações necessárias para selecionar a janela apropriada são bem adequadas para essa representação de extensão de palavra-código variável.

Por exemplo, ao utilizar uma informação de
10 extensão de palavra-código variável de janela, pode ser explorado que há uma dependência entre uma seleção de uma inclinação de transição e uma seleção de uma extensão de transformação, pois uma extensão de transformação curta não será tipicamente utilizada para uma janela que tem uma ou duas inclinações de transição
15 longas. Da mesma forma, uma transmissão de informações redundantes pode ser evitada ao utilizar uma informação de extensão de palavra-código variável de janela, melhorando, assim, a eficiência de taxa de bits das informações de áudio codificado.

Como um exemplo adicional, deve ser observado que
20 há uma correlação tipicamente entre formas de janela de estruturas adjacentes, que também pode ser explorada para reduzir seletivamente uma extensão de palavra-código da informação de janela para casos em que o tipo de janela de mais uma janela adjacente (adjacente à janela considerada atualmente) limita a
25 escolha dos tipos de janelas para a estrutura atual.

Para resumir o exposto acima, o uso de uma informação de extensão de palavra-código variável de janela permite ressalvar a taxa de bits sem aumentar de maneira

significativa uma complexidade do decodificador de áudio e sem alterar uma forma de onda de saída do decodificador de áudio (em comparação a uma informação de janela de extensão de palavra-código constante). Também, a sintaxe das informações de áudio codificado também pode ser simplificada em alguns casos, assim como será discutido em detalhes posteriormente.

Em uma configuração preferida, o decodificador de áudio compreende um separador de corrente de bits configurado para separar uma corrente de bits que representa a informação de áudio codificado e para extrair da corrente de bits uma informação de extensão de inclinação de janela de um bit e para extrair de maneira seletiva, dependendo de um valor de uma informação de extensão de inclinação de janela de um bit, da corrente de bits uma informação de extensão de transformação de um bit. Nesse caso, o seletor de janela é preferivelmente configurado para utilizar ou omitir seletivamente, dependendo de uma informação de extensão de inclinação de janela, uma informação de extensão de transformação, a fim de selecionar uma janela para um processamento de uma determinada parte da representação "tempo-frequência".

Ao utilizar esse conceito, uma separação entre uma informação de extensão de inclinação de janela e uma informação de extensão de transformação pode ser obtida, que contribui para uma simplificação do mapeamento em alguns casos. Também, uma divisão da informação de janela em um bit de extensão de inclinação de janela compulsória e um bit de extensão de transformação, cuja presença, que depende do estado do bit de extensão de inclinação de janela, permite uma redução muito eficiente da taxa de bits, que pode ser obtida enquanto se mantém

a sintaxe da corrente de bits suficientemente simples. Da mesma forma, a complexidade do separador de corrente de bits mantém-se suficientemente pequena.

Em uma configuração preferida, o seletor de
5 janela é configurado para selecionar um tipo de janela para processar uma parte atual da informação de frequência de tempo (por exemplo, uma estrutura de áudio atual) dependendo de um tipo de janela selecionado para processar uma parte anterior (por exemplo, uma estrutura de áudio anterior) da informação de
10 frequência de tempo, de maneira que uma extensão de inclinação de janela do lado esquerdo da janela para processar a parte atual da informação de frequência de tempo seja igual para uma extensão de inclinação de janela do lado direito da janela selecionada para processar a parte anterior da informação de frequência de tempo.
15 Ao explorar essa informação, uma taxa de bits necessária para selecionar um tipo de janela para processar a parte atual da informação de frequência de tempo é particularmente pequena, uma vez que a informação para selecionar um tipo de janela é codificada com complexidade particularmente baixa. Em particular,
20 não é necessário "desperdiçar" um bit para codificar uma extensão de inclinação de janela do lado esquerdo da janela associada à parte atual da informação de frequência de tempo. Da mesma forma, ao utilizar a informação sobre uma extensão de inclinação de janela do lado direito utilizada para o processamento de uma parte
25 anterior da informação de frequência de tempo, dois bits (por exemplo, o bit de extensão de inclinação de janela compulsória e o bit de extensão de transformação facultativa) podem ser utilizados para selecionar uma janela apropriada de uma pluralidade de mais

do que quatro janelas selecionáveis. Portanto, é evitada a redundância desnecessária, e a eficiência da taxa de bits da corrente de bits codificada é melhorada.

Em uma configuração preferida, o seletor de
5 janela é configurado para selecionar entre um primeiro tipo de janela e um segundo tipo de janela dependendo de um valor de uma informação de extensão de inclinação de janela de um bit, caso uma extensão de inclinação de janela do lado direito da janela para processar a parte anterior da informação de frequência de tempo
10 tiver um valor "longo" (que indica uma extensão de inclinação de janela comparavelmente maior quando comparado a um valor "pequeno" que indica uma extensão de inclinação de janela comparavelmente menor) e caso uma parte anterior da informação de frequência de tempo, uma parte atual da informação de frequência de tempo e uma
15 parte subsequente da informação de frequência de tempo forem todas codificadas em um modo central de domínio de frequência.

O seletor de janela também é preferivelmente configurado para selecionar um terceiro tipo de janela em resposta ao primeiro valor (por exemplo, um valor de "um") de uma
20 informação de extensão de inclinação de janela de um bit, se uma extensão de inclinação de janela do lado direito da janela para processar a parte anterior da informação de frequência de tempo tiver um valor "curto" (conforme discutido acima), e se uma parte anterior da informação de frequência de tempo, a parte atual da
25 informação de frequência de tempo e uma parte subsequente da informação de frequência de tempo forem todas codificadas em um modo central de domínio de frequência.

Ademais, o seletor de janela também é

preferivelmente configurado para selecionar entre um quarto tipo de janela e uma sequência de janelas (que pode ser considerada um quinto tipo de janela) dependendo de uma informação de extensão de transformação de um bit, se uma informação de extensão de inclinação de janela de um bit tiver um segundo valor (por exemplo, um valor de "zero") indicando uma inclinação de janela do lado direito curta, e se a extensão de inclinação de janela do lado direito da janela para processar a parte anterior da informação de frequência de tempo tiver um valor "curto" (conforme discutido acima), e se a parte anterior da informação de frequência de tempo, a parte atual da informação de frequência de tempo e a parte subsequente da informação de frequência de tempo são todas codificadas em um modo central de domínio de frequência.

Para esse caso, o primeiro tipo de janela compreende uma extensão de inclinação de janela do lado esquerdo (comparavelmente) longa, uma extensão de inclinação de janela do lado direito (comparavelmente) longa e uma extensão de transformação (comparavelmente) longa, o segundo tipo de janela compreende uma extensão de inclinação de janela do lado esquerdo (comparavelmente) longa, uma extensão de inclinação de janela do lado direito (comparavelmente) curta e uma extensão de transformação (comparavelmente) longa, o terceiro tipo de janela compreende uma extensão de inclinação de janela do lado esquerdo (comparavelmente) curta, uma extensão de inclinação de janela do lado direito (comparavelmente) longa e uma extensão de transformação (comparavelmente) longa, e o quarto tipo de janela compreende uma extensão de inclinação de janela do lado esquerdo (comparavelmente) curta, uma extensão de inclinação de janela do

lado direito (comparavelmente) curta e uma extensão de transformação (comparavelmente) longa. A "sequência de janelas" (ou quinto tipo de janela) define uma sequência ou superposição de uma pluralidade de sub-janelas associadas a uma única parte (por exemplo, estrutura) da informação de frequência de tempo, cada uma da pluralidade de sub-janelas tendo uma extensão de transformação (comparavelmente) curta, uma extensão de inclinação de janela do lado esquerdo (comparavelmente) curta e uma extensão de inclinação de janela do lado direito (comparavelmente) curta. Ao utilizar essa abordagem, um total de cinco tipos de janelas (incluindo o tipo "sequência de janelas") pode ser selecionado utilizando somente dois bits, onde uma informação de bit único (a saber, informação de extensão de inclinação de janela de um bit) é suficiente para sinalizar a sequência muito comum de uma pluralidade de janelas que têm extensão de inclinação de janelas comparavelmente longa tanto no lado esquerdo como no lado direito. Ao contrário, uma informação de janela de dois bits é somente necessária na preparação de uma sequência de janelas curtas ("sequência de janelas" ou "quinto tipo de janela") e durante uma série de estruturas de "sequência de janelas" temporariamente estendida (através de uma pluralidade de estruturas).

Para resumir, o conceito descrito acima para selecionar um tipo de janela de uma pluralidade de, por exemplo, cinco tipos diferentes de janelas permitem uma redução sólida da taxa de bits necessária. Enquanto, convencionalmente, três bits dedicados seriam necessários para selecionar um tipo de janela dos, por exemplo, cinco tipos de janelas, somente um ou dois bits são necessários, de acordo com a presente invenção, para realizar

essa seleção. Portanto, uma preservação significativa de bits pode ser alcançada, reduzindo, portanto, a taxa de bits necessária e/ou provendo a chance de melhorar a qualidade de áudio.

Em uma configuração preferida, o seletor de
5 janela é configurado para avaliar de maneira seletiva um bit de extensão de transformação de uma informação de extensão de palavra-código variável de janela somente se um tipo de janela para um processamento de uma parte anterior (por exemplo, estrutura) da informação de frequência de tempo compreender uma
10 extensão de inclinação de janela do lado direito correspondente a uma extensão de inclinação de janela do lado esquerdo de uma sequência de janelas curta e se uma informação de extensão de inclinação de janela de um bit associada à parte atual (por exemplo, estrutura atual) da informação de frequência de tempo
15 definir uma extensão de inclinação de janela do lado direito correspondente à extensão de inclinação de janela do lado direito da sequência de janelas curta.

Em uma configuração preferida, o seletor de janela é configurado adicionalmente para receber uma informação de
20 modo central anterior associada a uma parte anterior (por exemplo, estrutura) das informações de áudio e que descreve um modo central utilizado para codificar a parte anterior (por exemplo, estrutura) das informações de áudio. Nesse caso, o seletor de janela é configurado para selecionar uma janela para um processamento de
25 uma parte atual (por exemplo, estrutura) da representação "tempo-frequência" dependendo da informação de modo central anterior e também dependendo de uma informação de extensão de palavra-código variável de janela associada à parte atual da representação "tempo

frequência". Assim, o modo central de uma estrutura anterior pode ser explorado para selecionar uma janela apropriada para uma transição (por exemplo, na forma de uma operação de sobreposição-e-adição) entre a estrutura anterior e a estrutura atual.

5 Novamente, o uso de uma informação de extensão de palavra-código variável de janela é muito vantajoso, pois é novamente possível salvar um número significativo de bits. Uma preservação particularmente boa pode ser obtida se o número de tipos de janelas, que está disponível (ou válido) para uma estrutura de
10 áudio codificado, por exemplo, em um domínio de previsão linear, é pequeno. Portanto, geralmente é possível utilizar uma palavra-código curta, de uma palavra-código maior e uma palavra-código menor, em uma transição entre dois modos centrais diferentes (por exemplo, entre um modo central de domínio de previsão linear e um
15 modo central de domínio de frequência).

Em uma configuração preferida, o seletor de janela é configurado adicionalmente para receber uma informação de modo central subsequente associada a uma parte subsequente (ou estrutura) das informações de áudio e que descreve um modo central
20 utilizado para codificar a estrutura subsequente das informações de áudio. Nesse caso, o seletor de áudio é preferivelmente configurado para selecionar uma janela para um processamento de uma parte atual (por exemplo, estrutura) da representação "tempo frequência" dependendo da informação de modo central subsequente e
25 também dependendo de uma informação de extensão de palavra-código variável de janela associada à parte atual da representação "tempo frequência". Novamente, uma informação de extensão de palavra-código variável de janela pode ser explorada, em combinação com a

informação de modo central subsequente, a fim de determinar o tipo de janela com uma necessidade de contagem de bits baixa.

Em uma configuração preferida, o seletor de janela é configurado para selecionar janelas que têm uma
5 inclinação do lado direito encurtada, se a informação de modo central subsequente indicar que uma estrutura subsequente das informações de áudio é codificada utilizando um modo central de domínio de previsão linear. Dessa forma, uma adaptação das janelas para uma transição entre o modo central de domínio de frequência e
10 o modo central de domínio de tempo pode ser estabelecida sem precisar de esforço de sinalização extra.

Uma outra configuração, de acordo com a invenção, cria um codificador de áudio para prover uma informação de áudio codificado com base em uma informação de áudio de entrada. O
15 codificador de áudio compreende um transformador de sinal com base em janela configurado para prover uma sequência de parâmetros de sinal de áudio (por exemplo, uma representação de domínio de frequência de tempo da informação de áudio de entrada) com base em uma pluralidade de partes janeladas (por exemplo, estruturas de
20 sobreposição ou sem sobreposição) das informações de áudio de entrada. O transformador de sinal com base em janela é preferivelmente configurado para adaptar uma forma de janela para obter as partes janeladas das informações de áudio de entrada dependendo das características das informações de áudio de
25 entrada. O transformador de sinal com base em janela é configurado para alternar entre o uso de janelas que têm inclinação de transição (comparavelmente) mais longa e janelas que têm inclinação de transição (comparavelmente) mais curta, e também

alterna entre um uso de janelas que têm duas ou mais extensões de transformação diferentes. O transformador de sinal com base em janela também é configurado para determinar um tipo de janela utilizado para transformar uma parte atual (por exemplo, estrutura) das informações de áudio de entrada dependendo de um tipo de janela utilizado para transformar uma parte anterior (por exemplo, estrutura) das informações de áudio de entrada e um conteúdo de áudio da parte atual das informações de áudio de entrada. Também, o codificador de áudio é configurado para codificar uma informação de janela que descreve um tipo de janela utilizado para transformar uma parte atual das informações de áudio de entrada utilizando uma palavra-código de extensão variável. Esse codificador de áudio provê as vantagens já discutidas em relação ao decodificador de áudio inventivo. Em particular, é possível reduzir a taxa de bits das informações de áudio codificado ao evitar o uso de uma palavra-código comparavelmente longa em algumas ou em todas as situações em que isso é possível.

Uma outra configuração, de acordo com a invenção, cria uma informação de áudio codificado. As informações de áudio codificado compreendem uma representação "tempo frequência" codificada que descreve um conteúdo de áudio de uma pluralidade de partes janeladas de um sinal de áudio. As janelas de inclinações de transição diferentes (por exemplo, extensões de inclinação de transição) e extensões de transformação diferentes são associadas a diferentes partes janeladas do sinal de áudio. As informações de áudio codificado também compreendem uma informação de janela codificada que codifica tipos de janelas utilizados para obter as

representações de frequência de tempo codificadas de uma pluralidade de partes janeladas do sinal de áudio. A informação de janela codificada é uma informação de janela de extensão variável que codifica um ou mais tipos de janelas utilizando um primeiro
5 número de bits menor e que codifica um ou mais outros tipos de janelas utilizando um Segundo número de bits maior. Essas informações de áudio codificado trazem consigo as vantagens já discutidas acima em relação ao decodificador de áudio inventivo e o codificador de áudio inventivo.

10 Uma outra configuração, de acordo com a invenção, cria um método para prover uma informação de áudio decodificado com base em uma informação de áudio codificado. O método compreende avaliar uma informação de extensão de palavra-código variável de janela a fim de selecionar uma janela, de uma
15 pluralidade de janelas compreendendo janelas de diferentes inclinações de transição (por exemplo, diferentes extensões de inclinação de transição) e janelas de diferentes extensões de transformação, para um processamento de uma determinada parte da representação "tempo frequência" associada a uma determinada
20 estrutura das informações de áudio. O método também compreende mapear a determinada parte da representação "tempo frequência", que é descrita pelas informações de áudio codificado, para uma representação no domínio de tempo utilizando a janela selecionada.

Uma outra configuração, de acordo com a invenção,
25 cria um método para prover uma informação de áudio codificado com base em uma informação de áudio de entrada. O método compreende prover uma sequência de parâmetros de sinal de áudio (por exemplo, uma representação no domínio de frequência de tempo) com base em

uma pluralidade de partes janeladas das informações de áudio de entrada. Para prover a sequência de parâmetros de sinal de áudio, é realizada uma troca entre um uso de janelas que têm uma inclinação de transição mais longa e janelas que têm uma
5 inclinação de transição mais curta, e também entre um uso de janelas que têm duas ou mais extensões de transformação diferentes, para adaptar formas de janela para obter as partes janeladas das informações de áudio de entrada dependendo das características das informações de áudio de entrada. O método
10 também compreende codificar uma informação de janela, que descreve um tipo de janela utilizado para transformar uma parte atual das informações de áudio de entrada, utilizando uma palavra-código de extensão variável.

Além disso, as configurações, de acordo com a
15 invenção, criam programas de computador para implementar os referidos métodos.

BREVE DESCRIÇÃO DOS DESENHOS

As configurações da invenção serão
subsequentemente descritas, tendo como referência as figuras
20 anexas, nas quais:

Figura 1 apresenta um diagrama em blocos esquemático de um codificador de áudio, de acordo com uma configuração da invenção;

Figura 2 apresenta um diagrama em blocos
25 esquemático de um decodificador de áudio, de acordo com uma configuração da invenção;

Figura 3 apresenta uma representação esquemática de diferentes tipos de janelas, que podem ser utilizados de acordo

com o conceito inventivo;

Figura 4 apresenta uma representação gráfica das transições admissíveis entre janelas de diferentes tipos de janelas, que podem ser aplicadas no desenho das configurações, de
5 acordo com a invenção;

Figura 5 apresenta uma representação gráfica de uma sequência de diferentes tipos de janelas, que pode ser gerada por um codificador inventivo ou que pode ser processada por um decodificador de áudio inventivo;

10 Figura 6 apresenta uma tabela que representa uma sintaxe da corrente de bits proposta, de acordo com uma configuração da invenção;

Figura 6b apresenta uma representação gráfica de um mapeamento de um tipo de janela da estrutura atual para uma
15 informação de "window_length" e uma informação de "transform_length";

Figura 6c apresenta uma representação gráfica de um mapeamento para obter o tipo de janela da estrutura atual com base em uma informação de modo central anterior, uma informação de
20 "window_length" da estrutura anterior, uma informação de "window_length" da estrutura atual e uma informação de "transform_length" da estrutura atual;

Figura 7a apresenta uma tabela que representa uma sintaxe de uma informação de "window_length";

25 Figura 7b uma tabela que representa uma sintaxe de uma informação de "transform_length";

Figura 7c apresenta uma tabela que representa uma nova sintaxe da corrente de bits e transições;

Figura 8 apresenta uma tabela que dá uma visão geral de todas as combinações da informação de "window_length" e a informação de "transform_length";

Figura 9 apresenta uma tabela que representa uma preservação de bits, que pode ser obtida utilizando uma configuração da invenção;

Figura 10a apresenta uma representação de sintaxe de um denominado Bloco de dados brutos USAC;

Figura 10b apresenta uma representação de sintaxe de um denominado elemento de canal único;

Figura 10c apresenta uma representação de sintaxe de um denominado elemento de par de canais;

Figura 10d apresenta uma representação de sintaxe de uma denominada informação ICS;

Figura 10e apresenta uma representação de sintaxe de uma denominada corrente de canal de domínio de frequência;

Figura 11 apresenta um fluxograma de um método para prover uma informação de áudio codificado com base em uma informação de áudio de entrada; e

Figura 12 apresenta um fluxograma de um método para prover uma informação de áudio decodificado com base em uma informação de áudio codificado.

DESCRIÇÃO DETALHADA DAS CONFIGURAÇÕES

Visão geral do Codificador de áudio

A seguir, um codificador de áudio será descrito, no qual o conceito inventivo pode ser aplicado. Entretanto, deve ser observado que o codificador de áudio descrito em relação à Figura 1 deve ser considerado somente um exemplo de um codificador

de áudio no qual a invenção pode ser aplicada. Entretanto, mesmo que seja discutido um codificador de áudio comparavelmente simples em relação à Figura 1, deve ser observado que a invenção também pode ser aplicada em codificadores de áudio muito mais elaborados, por exemplo, codificadores de áudio que são capazes de comutar entre diferentes modos centrais de codificação (por exemplo, entre codificação de domínio de frequência e codificação de domínio de previsão linear). Todavia, devido à simplicidade, parece ser útil entender as ideias básicas de um codificador de áudio de domínio de frequência simples.

O codificador de áudio apresentado na Figura 1 é muito similar ao codificador de áudio descrito no padrão internacional ISO/IEC 14496-3:2005 (E), parte 3, subparte 4 e também nos documentos com referência a ele. Da mesma forma, deve ser feita referência ao referido padrão, os documentos citados nele e a extensa literatura relacionada à codificação de áudio MPEG.

O codificador de áudio 100 apresentado na Figura 1 é configurado para receber uma informação de áudio de entrada 110, por exemplo, um sinal de áudio de domínio de tempo. O codificador de áudio 100 compreende adicionalmente um pré-processador opcional 120 configurado para pré-processar opcionalmente as informações de áudio de entrada 110, por exemplo, ao *downsampling* as informações de áudio de entrada 110 ou ao controlar um ganho das informações de áudio de entrada 110. O codificador de áudio 100 também compreende, como uma componente chave, um transformador de sinal com base em janela 130, que é configurado para receber as informações de áudio de entrada 110,

ou uma versão delas pré-processada 122, e para transformar as informações de áudio de entrada 110 ou a versão delas pré-processada 122 no domínio de frequência (ou domínio de frequência de tempo), a fim de obter uma sequência de parâmetros de sinal de áudio, que podem ser valores espectrais em um domínio de frequência de tempo. Para este fim, o transformador de sinal com base em janela 130 compreende um janelador/transformador 136, que pode ser configurado para transformar blocos de amostras (por exemplo, "estruturas") das informações de áudio de entrada 110, 122 em conjuntos de valores espectrais 132. Por exemplo, o janelador/transformador 136 pode ser configurado para prover um conjunto de valores espectrais para cada bloco de amostras (ou seja, para cada "estrutura") das informações de áudio de entrada. Entretanto, os blocos de amostras (ou seja, "estruturas") das informações de áudio de entrada 110, 122 podem ser preferivelmente sobrepostos, de maneira que blocos de amostras adjacentes (estruturas) temporariamente das informações de áudio de entrada 110, 122 compartilhem uma pluralidade de amostras. Por exemplo, dois blocos de amostras subsequentes temporariamente (estruturas) podem sobrepor 50% das amostras. Da mesma forma, o janelador/transformador 136 pode ser configurado para realizar uma denominada transformação sobreposta, por exemplo, uma transformada discreta do cosseno modificada (MDCT). Entretanto, ao realizar a transformada discreta do cosseno modificada, o janelador/transformador 136 pode aplicar uma janela a cada bloco de amostras, ponderando, assim, as amostras centrais (temporariamente organizadas na proximidade de um centro temporal de um bloco de amostras) mais sólidas do que as amostras

periféricas (temporariamente organizadas na proximidade temporal da extremidade esquerda e direita de um bloco de amostras). O janelamento pode ajudar a evitar artefatos, que seriam originados da segmentação das informações de áudio de entrada 110, 122 nos 5 blocos. Portanto, a aplicação de janelas antes ou durante a transformação do domínio de tempo para o domínio de frequência de tempo permite uma transição suave entre blocos de amostras subsequentes das informações de áudio de entrada 110, 122. Para detalhes em relação ao janelamento, novamente é feita referência 10 ao padrão internacional ISO/IEC 14496, parte 3, subparte 4 e aos documentos referenciados nele. Em uma versão muito simples do codificador de áudio, um número de $2N$ amostras de uma estrutura de áudio (definida como um bloco de amostras) será transformado em um conjunto de N coeficientes espectrais independentes das 15 características de sinal. Entretanto, descobriu-se que esse conceito, no qual uma extensão de transformação uniforme de $2N$ amostras das informações de áudio 110, 122 que é utilizado independente das características das informações de áudio de entrada 110, 122 resulta em uma grave degradação das transições, 20 pois no caso de uma transição, a energia da transição é propagada sobre toda a estrutura ao decodificar as informações de áudio. Todavia, descobriu-se que uma melhora na codificação de margens pode ser obtida se uma extensão de transformação mais curta (por exemplo, $2N/8=N/4$ amostras por transformação) for escolhida. 25 Entretanto, também se descobriu que a escolhas de uma extensão de transformação mais curta aumenta tipicamente a taxa de bits necessária, mesmo se menos valores espectrais forem obtidos para uma extensão de transformação mais curta quando comparados a uma

extensão de transformação mais longa. Da mesma forma, descobriu-se ser recomendável trocar de uma extensão de transformação longa (por exemplo, $2N$ amostras por transformação) para uma extensão de transformação curta (por exemplo, $2N/8=N/4$ amostras por transformação) na proximidade de uma transição (também designada como margem) do conteúdo de áudio, e para trocar de volta para a extensão de transformação longa (por exemplo, $2N$ -amostras por transformação) após a transição. A troca da extensão de transformação se refere a uma alteração de uma janela aplicada para janelar as amostras das informações de áudio de entrada 110, 122 antes ou durante a transformação.

Em relação a essa questão, deve ser observado que, em muitos casos, o codificador de áudio é capaz de utilizar mais de duas janelas diferentes. Por exemplo, uma denominada "only_long_sequence" pode ser utilizada para codificar uma estrutura de áudio atual, se tanto a estrutura anterior (anterior à estrutura considerada atualmente) como a estrutura seguinte (seguinte à estrutura considerada atualmente) são codificadas utilizando uma extensão de transformação longa (por exemplo, $2N$ amostras). Ao contrário, uma denominada "long_start_sequence" pode ser utilizada em uma estrutura, que é transformada utilizando uma extensão de transformação longa, que é precedida por uma estrutura transformada utilizando uma extensão de transformação longa e que é seguida por uma estrutura transformada utilizando uma extensão de transformação curta. Em uma estrutura, que é transformada utilizando uma extensão de transformação curta, uma denominada sequência de janelas "eight_short_sequence", que compreende oito (sub)janelas curtas e de sobreposição, pode ser aplicada. Além

disso, uma denominada janela de "long_start_sequence" pode ser aplicada para transformar uma estrutura, que é precedida por uma estrutura anterior transformada utilizando uma extensão de transformação curta e que é seguida por uma estrutura transformada utilizando uma extensão de transformação longa. Para detalhes em relação às possíveis sequências de janelas, é feita referência ao ISO/IEC 14496-3:2005 (E) parte 3, subparte 4. Também, é feita referência às Figuras 3, 4, 5, 6, que serão explicadas em detalhes abaixo.

Entretanto, deve-se observar que, em algumas configurações, um ou mais tipos adicionais de janelas podem ser utilizados. Por exemplo, uma denominada janela de "stop_start_sequence" pode ser aplicada se a estrutura atual for precedida por uma estrutura, na qual uma extensão de transformação curta é utilizada, e se a estrutura atual for seguida por uma estrutura na qual uma extensão de transformação curta é utilizada.

Da mesma forma, o transformador de sinal com base em janela 130 compreende um determinador de sequência de janelas 138, que é configurado para prover um tipo de informação de janela ao janelador/transformador 136, de maneira que o janelador/transformador 136 possa utilizar um tipo adequado de janela ("sequência de janelas"). Por exemplo, o determinador de sequência de janelas 130 pode ser configurado para avaliar diretamente as informações de áudio de entrada 110 ou as informações de áudio de entrada pré-processadas 122. Entretanto, de maneira alternativa, o codificador de áudio 100 pode compreender um processador de modelo psicoacústico 150, que é configurado para receber as informações de áudio de entrada 110 ou

as informações de áudio de entrada pré-processadas 122, e para aplicar um modelo psicoacústico a fim de extrair informações, que são relevantes para a codificação das informações de áudio de entrada 110, 122, das informações de áudio de entrada 110, 122.

5 Por exemplo, o processador de modelo psicoacústico 150 pode ser configurado para identificar transições dentro das informações de áudio de entrada 110, 122 e para prover uma informação de extensão de janela 152, que pode sinalizar estruturas nas quais uma extensão de transformação curta é desejada por causa da presença
10 de uma transição nas informações de áudio de entrada 110, 122 correspondentes.

O processador de modelo psicoacústico 150 também pode ser configurado para determinar quais valores espectrais precisam ser codificados com alta resolução (ou seja,
15 quantificação fina) e quais valores espectrais podem ser codificados com resolução menor (ou seja, quantificação grosseira) sem obter uma degradação grave do conteúdo de áudio. Para este fim, o processador de modelo psicoacústico 150 pode ser configurado para avaliar efeitos de encobrimento psicoacústico,
20 identificando, portanto, valores espectrais (ou bandas de valores espectrais) que são de menor relevância psicoacústica e outros valores espectrais (ou bandas de valores espectrais) que são de maior relevância psicoacústica. Da mesma forma, o processador de modelo psicoacústico 150 provê uma informação de relevância
25 psicoacústica 154.

O codificador de áudio 100 compreende ainda um processador espectral opcional 160, que é configurado para receber a sequência de parâmetros de sinal de áudio 132 (por exemplo, a

representação de domínio de frequência de tempo das informações de áudio de entrada 110, 122) e prover, com base em si, uma sequência pós-processada parâmetros de sinal de áudio 162. Por exemplo, o processador posterior espectral 160 pode ser configurado para
5 realizar uma modelagem de ruído temporal, a previsão em longo prazo, uma substituição de ruído perceptual e/ou um processamento de canal de áudio.

O codificador de áudio 100 também compreende um processador de codificação/quantificação/escalonamento opcional
10 170, que é configurado para escalar os parâmetros de sinal de áudio (por exemplo, valores de domínio de frequência de tempo ou "valores espectrais") 132, 162, para realizar uma quantificação e para codificar os valores quantificados e escalados. Para este fim, o processador de codificação/quantificação/escalonamento 170
15 pode ser configurado para utilizar as informações 154 providas pelo processador de modelo psicoacústico, por exemplo, a fim de decidir qual escalonamento e/ou qual quantificação deve ser aplicada a quais parâmetros de sinal de áudio (ou valores espectrais). Da mesma forma, o escalonamento e a quantificação
20 podem ser adaptados, de maneira que uma taxa de bits desejada dos parâmetros de sinal de áudio escalonados, quantificados e codificados (ou valores espectrais) seja obtida.

Além disso, o codificador de áudio 100 compreende um codificador de palavra-código de extensão variável 180, que é
25 configurado para receber o tipo de informação de janela 140 do determinador de sequência de janelas 138 e para prover, em sua base, uma palavra-código de extensão variável 182, que descreve o tipo de janela utilizado para a operação de

janelamento/transformação realizada pelo janelador/transformador 136. Os detalhes em relação ao codificador de palavra-código de extensão variável 180 serão subsequentemente descritos.

Ademais, o codificador de áudio 100 compreende
5 opcionalmente formatador de carga útil de bits 190, que é configurado para receber a informação espectral escalada, quantizada e codificada 172 (que descreve uma sequência de parâmetros de sinal de áudio ou valores espectrais 132) e a palavra-código de extensão variável 182 que descreve o tipo de
10 janela utilizado para a operação de janelamento/transformação. Da mesma forma, a corrente de formatador de carga útil de bits 190 provê uma corrente de bits 192, na qual as informações 172 e a palavra-código de extensão variável 182 são incorporadas. A corrente de bits 192 serve como uma informação de áudio
15 codificado, e pode ser armazenada em um meio e/ou transferida do codificador de áudio 100 a um decodificador de áudio.

Para resumir o descrito acima, o codificador de áudio 100 é configurado para prover as informações de áudio codificado 192 com base nas informações de áudio de entrada 110. O
20 codificador de áudio 100 compreende, como um componente importante, o transformador de sinal com base em janela 130, que é configurado para prover uma sequência de parâmetros de sinal de áudio 132 (por exemplo, uma sequência de valores espectrais) com base em uma pluralidade de partes janeladas das informações de
25 áudio de entrada 110. O transformador de sinal com base em janela 130 é configurado de maneira que um tipo de janela para obter as partes janeladas das informações de áudio de entrada é selecionado dependendo das características das informações de áudio. O

transformador de sinal com base em janela 130 é configurado para alternar entre um uso de janelas que têm uma inclinação de transição mais longa e janelas que têm uma inclinação de transição mais curta, e também para alternar entre um uso de janelas que têm
5 duas ou mais extensões de transformação diferentes. Por exemplo, o transformador de sinal com base em janela 130 é configurado para determinar um tipo de janela utilizado para transformar uma parte atual (por exemplo, estrutura) das informações de áudio de entrada dependendo de um tipo de janela utilizado para transformar uma
10 parte anterior (por exemplo, estrutura) das informações de áudio de entrada, e dependendo de um conteúdo de áudio da parte atual das informações de áudio de entrada. Entretanto, o codificador de áudio é configurado para codificar, por exemplo, utilizando o codificador de palavra-código de extensão variável 180, o tipo de
15 informação de janela 140 que descreve um tipo de janela utilizado para transformar uma parte atual (por exemplo, estrutura) das informações de áudio de entrada utilizando uma palavra-código de extensão variável.

Tipos de janelas de transformação

20 A seguir, uma descrição detalhada de diferentes janelas, que podem ser aplicadas pelo janelador/transformador 136, e que são selecionadas pelo determinador de sequência de janelas 138, será realizada. Entretanto, as janelas aqui discutidas devem ser consideradas somente como exemplo. Subsequentemente, os
25 conceitos inventivos para a codificação eficiente do tipo de janela serão discutidos.

Tendo como referência, agora, a Figura 3, que apresenta uma representação gráfica de diferentes tipos de janelas

de transformação, uma visão geral sobre novas janelas de amostra será dada. Entretanto, é feita referência adicional ao ISO/IEC 14496-3, parte 3, subparte 4, no qual os conceitos para aplicar janelas de transformação são descritos ainda em mais detalhes.

5 A Figura 3 apresenta uma representação gráfica de um primeiro tipo de janela 310, que compreende uma inclinação de janela do lado esquerdo (comparavelmente) longa 310a (1024 amostras) e uma inclinação de janela do lado direito longa 310b (1024 amostras). Um total de 2048 amostras e 1024 coeficientes espectrais é associado ao primeiro tipo de janela 310, de maneira
10 que o primeiro tipo de janela 310 compreende uma denominada "extensão de transformação longa".

Um segundo tipo de janela 312 é designado como "long_start_sequence" ou "long_start_window". O segundo tipo de
15 janela compreende uma inclinação de janela do lado esquerdo (comparavelmente) longa 312a (1024 amostras) e uma inclinação de janela do lado direito (comparavelmente) curta 312b (128 amostras). Um total de 2048 amostras e 1024 coeficientes espectrais é associado ao segundo tipo de janela, de maneira que o
20 segundo tipo de janela 312 compreenda uma extensão de transformação longa.

O terceiro tipo de janela 314 é designado como "long_start_sequence" ou "long_stop_window". O terceiro tipo de janela 314 compreende a inclinação de janela do lado esquerdo
25 curta 314a (128 amostras) e a inclinação de janela do lado direito longa 314b (1024 amostras). Um total de 2048 amostras e 1024 coeficientes espectrais é associado ao terceiro tipo de janela 314, de maneira que o terceiro tipo de janela compreenda a

extensão de transformação longa.

O quarto tipo de janela 316 é designado como uma "stop_start_sequence" ou "stop_start_window". O quarto tipo de janela 316 compreende uma inclinação de janela do lado esquerdo curta 316a (128 amostras) e uma inclinação de janela do lado direito curta 316b (128 amostras). Um total de 2048 amostras e 1024 coeficientes espectrais é associado ao quarto tipo de janela, de maneira que o quarto tipo de janela compreenda uma "extensão de transformação longa".

Um quinto tipo de janela 318 se difere significativamente do primeiro ao quarto tipo de janelas. O quinto tipo de janela compreende uma superposição de oito "janelas curtas" ou sub-janelas 319a a 319h, que são organizadas para se sobreporem temporariamente. Cada uma das janelas curtas 319a a 319h compreende uma extensão de 256 amostras. Da mesma forma, uma transformação MDCT "curta", que transforma 256 amostras em 128 valores espectrais, é associada a cada uma das janelas curtas 319a-319h. Da mesma forma, oito conjuntos de 128 valores espectrais, cada um associado ao quinto tipo de janela 318, enquanto um único conjunto de 1024 valores espectrais é associado a cada, do primeiro ao quarto, tipos de janelas 310, 312, 314, 316. Da mesma forma, pode-se dizer que o quinto tipo de janela compreende uma extensão de transformação "curta". Todavia, o quinto tipo de janela compreende uma inclinação de janela do lado esquerdo curta 318a e uma inclinação de janela do lado direito curta 318b.

Portanto, para uma estrutura à qual o primeiro tipo de janela 310, o segundo tipo de janela 312, o terceiro tipo

de janela 314 ou o quarto tipo de janela 316 é associado, 2048 amostras das informações de áudio de entrada são juntamente janeladas e transformada de MDCT, como um único grupo, no domínio de frequência de tempo. Ao contrário, para uma estrutura à qual o

5 quinto tipo de janela 318 é associado, oito subconjuntos (pelo menos parcialmente sobrepostos) de 256 amostras, cada um é individualmente (ou separadamente) transformado de MDCT, de maneira que oito conjuntos de coeficientes de MDCT (valores de frequência de tempo) sejam obtidos.

10 Tendo como referência novamente a Figura 3, deve ser observado que a Figura 3 apresenta uma pluralidade de janelas adicionais. Essas janelas adicionais, a saber, uma denominada "stop_1152_sequence" ou "stop_window_1152" 330 e uma denominada "stop_start_1152_sequence" ou "1152_stop_start_window" 332 pode

15 ser aplicada se a estrutura atual for precedida por uma estrutura anterior, que é codificada em um domínio de previsão linear. Nesses casos, uma extensão da transformação é adaptada, a fim de permitir um cancelamento de artefatos *aliasing* de domínio de tempo.

20 Também, janelas adicionais 362, 366, 368, 382 podem ser opcionalmente aplicadas se a estrutura atual for seguida por uma estrutura subsequente, que é codificada no domínio de previsão linear. Entretanto, tipos de janelas 330, 332, 362, 366, 368, 382 devem ser considerados como opcional e não são

25 necessários para a implementação do conceito inventivo.

Transições entre tipos de janelas de transformação

Tendo agora como referência a Figura 4, que

apresenta uma representação esquemática das transições permitidas entre sequências de janelas (ou tipos de janelas de transformação), alguns detalhes adicionais serão explicados. Observando que duas janelas de transformação subsequentes, cada uma tendo um dos tipos de janelas 310, 312, 314, 316, 318, são aplicadas a blocos parcialmente sobrepostos de amostras de áudio, pode-se entender que uma inclinação de janela do lado direito de uma primeira janela deverá corresponder-se a uma inclinação de janela do lado esquerdo de uma segunda janela subsequente, a fim de evitar artefatos causados pela sobreposição parcial. Da mesma forma, uma escolha de tipos de janelas para a segunda estrutura (de duas estruturas subsequentes) é limitada, se o tipo de janela para a primeira estrutura (das duas estruturas subsequentes) for dado. Como pode ser visto na Figura 4, se a primeira janela for uma janela de "only_long_sequence", a primeira janela somente pode ser seguida por uma janela de "only_long_sequence" ou uma janela de "long_start_sequence". Ao contrário, não é permitido utilizar uma janela de "eight_short_sequence", uma janela de "long_stop_sequence" ou uma janela de "stop_start_sequence" para a segunda estrutura após a primeira estrutura, se a janela de "only_long_sequence" for utilizada para transformar a primeira estrutura. Semelhantemente, se uma janela de "long_stop_sequence" for utilizada na primeira estrutura, a segunda estrutura pode utilizar uma janela de "only_long_sequence" ou uma janela de "long_start_sequence", mas a segunda estrutura não pode usar uma janela de "eight_short_sequence", uma janela de "long_stop_sequence" ou uma janela de "stop_start_sequence".

Ao contrário, se a primeira estrutura (das duas

estruturas subsequentes) utilizar uma janela de "long_start_sequence", uma janela de "eight_short_sequence" ou uma janela de "stop_start_sequence", a segunda estrutura (das duas estruturas subsequentes) não pode usar uma janela de "only_long_sequence" ou uma janela de "long_start_sequence", mas pode usar uma janela de "eight_short_sequence", uma janela de "long_stop_sequence" ou uma janela de "stop_start_sequence".

As transições admissíveis entre os tipos de janelas de "only_long_sequence", "long_start_sequence", "eight_short_sequence", "long_stop_sequence" e "stop_start_sequence" são apresentadas por uma "verificação" na Figura 4. Ao contrário, as transições entre os tipos de janelas, para os quais não há "verificação", não são admissíveis em algumas configurações.

Ademais, deve ser observado que tipos adicionais de janelas "LPD_sequence", "stop_1152_sequence" e "stop_start_1152_sequence" podem ser utilizáveis, se as transições entre um modo central de domínio de frequência e um modo central de domínio de previsão linear forem possíveis. Todavia, essa possibilidade deveria ser considerada opcional e será discutida posteriormente.

Exemplo de sequência de janelas

A seguir, uma sequência de janelas será descrita, que faz uso dos tipos de janelas 310, 312, 314, 316, 318. A Figura 5 apresenta uma representação gráfica dessa sequência de janelas. Como pode ser visto, uma abscissa 510 indica o tempo. As estruturas que sobrepõe aproximadamente 50% são marcadas na Figura 5 e designadas com "estrutural" a "estrutura7". A Figura 5

apresenta uma primeira estrutura 520, que pode, por exemplo, compreender 2048 amostras. Uma segunda estrutura 522 é temporariamente alternada em relação à primeira estrutura 520 por (aproximadamente) 1024 amostras, de maneira que a segunda
5 estrutura sobreponha a primeira estrutura 520 por (aproximadamente) 50 %. Um alinhamento temporal de uma terceira estrutura 524, uma quarta estrutura 526, uma quinta estrutura 528, uma sexta estrutura 530 e uma sétima estrutura 532 podem ser vistas na Figura 5. Uma janela de "only_long_sequence" 540 (de
10 tipo 310) é associada à primeira estrutura 520. Também, uma janela de "only_long_sequence" 542 (de tipo 310) é associada à segunda estrutura 522. Uma janela de "long_start_sequence" 544 (de tipo 312) é associada à terceira estrutura, uma janela de "eight_short_sequence" 546 (de tipo 318) é associada à quarta
15 estrutura 526, uma janela de "stop_start_sequence" 548 (de tipo 316) é associada à quinta estrutura, uma janela de "eight_short_sequence" 550 (de tipo 318) é associada à sexta estrutura 530 e uma janela de "long_stop_sequence" 552 (de tipo 314) é associada à sétima estrutura 532. Da mesma forma, um único
20 conjunto de 1024 coeficientes MDCT é associado à primeira estrutura 520, um outro único conjunto de 1024 coeficientes MDCT é associado à segunda estrutura 522 e ainda um outro conjunto único de 1024 coeficientes MDCT é associado à terceira estrutura 524. Entretanto, oito conjuntos de 128 coeficientes MDCT são associados
25 à quarta estrutura 526. Um único conjunto de 1024 coeficientes MDCT é associado à quinta estrutura 528.

A sequência de janelas apresentada na Figura 5 pode, por exemplo, traz um resultado de codificação

particularmente eficiente de taxa de bits, se houver um evento temporário em uma parte central da quarta estrutura 526, e se houver um outro evento temporário em uma parte central da sexta estrutura 530, enquanto o sinal for aproximadamente fixo durante o resto de tempo (por exemplo, durante a primeira estrutura 520, a segunda estrutura 522, o início da terceira estrutura 524, o centro da quinta estrutura 528 e o final da sétima estrutura 532).

Entretanto, conforme será explicado em detalhes a seguir, a presente invenção cria um conceito particularmente eficiente para codificar os tipos de janelas associados às estruturas de áudio. Em relação a essa questão, deve ser observado que um total de cinco tipos diferentes de janelas 310, 312, 314, 316, 318 é utilizado na sequência de janelas 500 da Figura 5. Da mesma forma, seria "normalmente" necessário utilizar três bits para codificar o tipo de estrutura. Ao contrário, a presente invenção cria um conceito que permite uma codificação do tipo de janela com demanda de bit reduzida.

Tendo agora como referência a Figura 6a e também as Figuras 7a, 7b e 7c, o conceito inventivo para codificar o tipo de janela será explicado. A Figura 6a apresenta uma tabela representando uma sintaxe proposta de um tipo de informação de janela, que inclui uma norma para codificar o tipo de janela. Para fins de explicação, presume-se que o tipo de informação de janela 140, que é provido ao codificador de palavra-código de extensão variável 180 pelo determinador de sequência de janelas 138, descreve o tipo de janela da estrutura atual e pode ter um dos valores "only_long_sequence", "long_start_sequence", "eight_short_sequence", "long_stop_sequence",

"stop_start_sequence" e, opcionalmente, ainda um dos valores
 "stop_1152_sequence" e "stop_start_1152_sequence". Entretanto, de
 acordo com o conceito de codificação inventivo, o codificador de
 palavra-código de extensão variável 180 provê um informação de
 5 "window_length" de 1 bit, que descreve uma extensão de uma
 inclinação de janela direita da janela associada à estrutura
 atual. Como pode ser visto na Figura 7a, um valor de "0" da
 informação de "window_length" de 1 bit pode representar uma
 extensão da inclinação de janela direita de 1024 amostras, e um
 10 valor "1" pode representar uma extensão da inclinação de janela
 direita de 128 amostras. Da mesma forma, o codificador de palavra-
 código de extensão variável 180 pode prover um valor de "0" da
 informação de "window_length" se o tipo de janela for
 "only_long_sequence" (primeiro tipo de janela 310) ou
 15 "long_start_sequence" (terceiro tipo de janela 314).
 Opcionalmente, o codificador de palavra-código de extensão
 variável 180 também pode prover uma informação de "window_length"
 de "0" para uma janela de tipo "stop_1152_sequence" (tipo de
 janela 330). Ao contrário, o codificador de palavra-código de
 20 extensão variável 180 pode prover um valor de "1" da informação de
 "window_length" para uma "long_start_sequence" (segundo tipo de
 janela 312), para uma "stop_start_sequence" (quarto tipo de janela
 316) e para uma "eight_short_sequence" (quinto tipo de janela
 318). Opcionalmente, o codificador de palavra-código de extensão
 25 variável 180 também pode prover uma informação de "window_length"
 de "1" para uma "stop_start_1152_sequence" (tipo de janela 332).
 Além disso, o codificador de palavra-código de extensão variável
 180 pode opcionalmente prover um valor de "1" da informação de

"window_length" para um ou mais dos tipos de janelas 362, 366, 368, 382.

Entretanto, o codificador de palavra-código de extensão variável 180 é configurado para prover de maneira seletiva uma outra informação de 1 bit, a saber, a denominada informação de "transform_length" da estrutura atual, dependendo do valor da informação de "window_length" de 1 bit da estrutura atual. Se a informação de "window_length" da estrutura atual tiver o valor "0" (ou seja, para os tipos de janelas "only_long_sequence", "long_stop_sequence" e, opcionalmente, "stop_1152_sequence"), o codificador de palavra-código de extensão variável 180 não provê uma informação de "transform_length" para inclusão na corrente de bits 192. Ao contrário, se a informação de "window_length" de uma estrutura atual tiver o valor "1" (ou seja, para os tipos de janelas "long_start_sequence", "stop_start_sequence", "eight_short_sequence" e, opcionalmente, "sequência_início_LPD" e "stop_start_1152_sequence") o codificador de palavra-código de extensão variável 180 provê a informação de "transform_length" de 1 bit para a inclusão na corrente de bits 192. A informação de "transform_length" é provida, se for provida, de maneira que a informação de "transform_length" represente a extensão de transformação aplicada à estrutura atual. Portanto, a informação de "transform_length" é provida para considera o primeiro valor (por exemplo, o valor de "0") para os tipos de janelas "long_start_sequence", "stop_start_sequence" e, opcionalmente, "stop_start_1152_sequence" e "sequência_parada_LPD", indicando, assim, que o tamanho de núcleo de MDCT aplicado à estrutura atual é de 1024 amostras (ou 1152

amostras). Ao contrário, a informação de "transform_length" é provida pelo codificador de palavra-código de extensão variável 180 para considerar um segundo (por exemplo, um valor de "1") se um tipo de janela de "eight_short_sequence" for associado à
 5 estrutura atual, indicando, assim, que o tamanho de núcleo de MDCT associado à estrutura atual é de 128 amostras (vide a representação de sintaxe da Figura 7b).

Para resumir, o codificador de palavra-código de extensão variável 180 provê uma palavra-código de 1 bit,
 10 compreendendo somente a informação de "window_length" de 1 bit da estrutura atual, para a inclusão na corrente de bits 192 se a inclinação de janela do lado direito da janela associada à estrutura atual for comparavelmente longa (inclinação de janela longa 310b, 314b, 330b), ou seja, para os tipos de janelas
 15 "only_long_sequence", "long_stop_sequence" e "stop_1152_sequence". Ao contrário, o codificador de palavra-código de extensão variável 180 provê uma palavra-código de 2 bits, compreendendo a informação de "window_length" de 1 bit e a informação de "transform_length" de 1 bit, para inclusão na corrente de bits 192, se a inclinação
 20 de janela do lado direito da janela associada à estrutura atual for uma inclinação de janela curta 312b, 316b, 318b, 332b, ou seja, para tipos de janelas "long_start_sequence", "eight_short_sequence", "stop_start_sequence" e, opcionalmente, "stop_start_1152_sequence". Portanto, 1 bit é preservado para o
 25 caso do tipo de janela de "only_long_sequence" e o tipo de janela de "long_stop_sequence" (e, opcionalmente, para um tipo de janela de "stop_1152_sequence").

Portanto, somente um ou dois bits, dependendo do

tipo de janela associado à estrutura atual, são necessários para codificar uma seleção de cinco (ou até mais) possíveis tipos de janelas.

Deve-se observar aqui que a Figura 6a apresenta um mapeamento de um tipo de janela, que é definido em uma coluna de tipo de janela 630, em um valor da informação de "window_length", que é apresentado em uma coluna 620 e, também, em uma condição de provisão e valor (se necessário) da informação de "transform_length", que é apresentado em uma coluna 624.

A Figura 6b apresenta uma representação gráfica de um mapeamento para derivar informação de "window_length" da estrutura atual e a informação de "transform_length" (ou uma indicação que a informação de "transform_length" é omitida da corrente de bits 192) do tipo de janela da estrutura atual. Esse mapeamento pode ser realizado pelo codificador de palavra-código de extensão variável 180, que recebe o tipo de informação de janela 140 que descreve o tipo de janela da estrutura atual e mapeia ele na informação de "window_length", conforme apresentado em uma coluna 660 da tabela da Figura 6b e em uma informação de "transform_length", conforme apresentado em uma coluna 662 da tabela da Figura 6b. Em particular, o codificador de palavra-código de extensão variável 180 pode prover a informação de "transform_length" somente se a informação de "window_length" considerar um valor predeterminado (por exemplo, de "1") e, de outra forma, omitir a provisão da informação de "transform_length", ou suprimir a inclusão da informação de "transform_length" na corrente de bits 192. Da mesma forma, diversos bits de tipo de janela incluídos na corrente de bits 192

para uma determinada estrutura podem variar, conforme indicado em uma coluna 664 de uma tabela da Figura 6b, dependendo do tipo de janela da estrutura atual.

Também se deve observar que, em algumas
5 configurações, o tipo de janela da estrutura atual pode ser adaptado ou modificado, se a estrutura atual for seguida por uma estrutura codificada no domínio de previsão linear. Entretanto, isso não afeta tipicamente o mapeamento do tipo de janela na informação de "window_length" e a informação de "transform_length"
10 provida de maneira seletiva.

Da mesma forma, o codificador de áudio 100 é configurado para prover uma corrente de bits 192, de maneira que a corrente de bits 192 execute a sintaxe, que será discutida abaixo tendo como referência as Figuras 10a a 10e.

15 Visão geral do Decodificador de áudio

A seguir, um decodificador de áudio, de acordo com uma configuração da invenção, será descrito em detalhes tendo como referência a Figura 2. A Figura 2 apresenta um diagrama esquemático de um decodificador de áudio, de acordo com uma
20 configuração da invenção. O decodificador de áudio 200 da Figura 2 é configurado para receber uma corrente de bits 210 compreendendo uma informação de áudio codificado e para prover, em sua base, uma informação de áudio decodificado 212 (por exemplo, na forma de um sinal de áudio de domínio de tempo). O decodificador de áudio 200
25 compreende um desformatador de carga útil de corrente de bits opcional 220, que é configurado para receber a corrente de bits 210 e para extrair da corrente de bits 210 uma informação de valor espectral codificado 222 e uma informação de extensão de palavra-

código variável de janela 224. O desformatador de carga útil de corrente de bits 220 pode ser configurado para extrair informações adicionais, como informações de controle, informações de ganho e informações de parâmetro de áudio adicionais, da corrente de bits 210. Entretanto, essas informações adicionais são bem conhecidas a um técnico no assunto e não são relevantes à presente invenção. Para detalhes adicionais, é feita referência, por exemplo, ao padrão internacional ISO/IEC 14496-3: 2005(E), parte 3, subparte 4.

O decodificador de áudio 200 compreende um decodificador/quantificador inverso/reescalador opcional 230 que é configurado para decodificar a informação de valor espectral codificado 222, para realizar uma quantificação inversa e também para realizar uma reescalonamento da informação de valor espectral quantificado de maneira inversa, obtendo, assim, uma informação de valor espectral decodificado 232. O decodificador de áudio 200 compreende ainda um pré-processador espectral opcional 240, que pode ser configurado para realizar uma ou mais etapas de pré-processamento espectral. Algumas das etapas de pré-processamento espectral possíveis são, por exemplo, explicadas no padrão internacional ISO/IEC 14496-3: 2005(E), parte 3, subparte 4. Da mesma forma, a funcionalidade do decodificador/quantificador inverso/reescalador e do pré-processador espectral opcional 240 resulta na provisão de uma representação "tempo frequência" 242 (decodificada e, opcionalmente, pré-processada) das informações de áudio codificado representadas pela corrente de bits 210. O decodificador de áudio 200 compreende, como um componente principal, um transformador de sinal com base em janela 250. O

transformador de sinal com base em janela 250 é configurado para transformar a representação "tempo-frequência" (decodificada) 242 em um sinal de áudio de domínio de tempo 252. Para este fim, o transformador de sinal com base em janela 250 pode ser configurado para realizar uma transformação de domínio de frequência de tempo para domínio de tempo. Por exemplo, o transformador/janelador 254 do transformador de sinal com base em janela 250 pode ser configurado para receber, como representação "tempo-frequência" 242, coeficientes da transformada discreta do cosseno modificada (coeficientes MDCT) associados temporariamente à estrutura das informações de áudio codificado sobreposta. Da mesma forma, o transformador/janelador 254 pode ser configurado para realizar a transformação sobreposta, na forma de uma transformada discreta do cosseno modificada inversa (IMDCT), para obter partes (estruturas) de domínio de tempo janeladas das informações de áudio codificado, e para sobrepor-e-adicionar as partes (estruturas) de domínio de tempo janeladas subsequentes utilizando uma operação de sobreposição e adição. Ao reconstruir o sinal de áudio de domínio de tempo 252 com base na representação "tempo-frequência" 242, ou seja, ao realizar a transformada discreta do cosseno modificada inversa em combinação com o janelamento e a operação de sobreposição-e-adição, o transformador/janelador 254 pode selecionar uma janela, de uma pluralidade de tipos de janelas disponíveis, a fim de permitir uma reconstrução apropriada e, também, a fim de evitar quaisquer artefatos de bloqueio.

O decodificador de áudio também compreende um processador posterior de domínio de tempo opcional 260, que é configurado para obter as informações de áudio decodificado 212

com base no sinal de áudio de domínio de tempo 252. Entretanto, deve ser observado que as informações de áudio decodificado 212 podem ser idênticas ao sinal de áudio de domínio de tempo 252 em algumas configurações. Além disso, o decodificador de áudio 200
5 compreende um seletor de janela 270, que é configurado para receber a informação de extensão de palavra-código variável de janela 224, por exemplo, do desformatador de carga útil de corrente de bits opcional 220. O seletor de janela 270 é configurado para prover a informação de janela 272 (por exemplo, a
10 tipo de informação de janela ou uma informação de sequência de janelas) ao transformador/janelador 254. Deve ser observado que o seletor de janela 270 pode ou não ser parte do transformador de sinal com base em janela 250 dependendo da implementação real.

Para resumir o descrito acima, o decodificador de
15 áudio 200 é configurado para prover as informações de áudio decodificado 212 com base nas informações de áudio codificado 210. O decodificador de áudio 200 compreende, como um componente principal, o transformador de sinal com base em janela 250, que é configurado para mapear uma representação "tempo-frequência" 242,
20 que é descrita pelas informações de áudio codificado 210, a uma representação nodomínio de tempo 252. O transformador de sinal com base em janela 250 é configurado para selecionar uma janela, de uma pluralidade de janelas compreendendo janelas de diferentes inclinações de transição (por exemplo, diferentes extensões de
25 inclinação de transição) e janelas de diferentes extensões de transformação, com base na informação de janela 272. O decodificador de áudio 200 compreende, como um outro componente principal, o seletor de janela 270, que é configurado para avaliar

a informação de extensão de palavra-código variável de janela 224, a fim de selecionar uma janela para um processamento de uma determinada parte da representação "tempo-frequência" 242 associada a uma determinada estrutura das informações de áudio. Os outros componentes do decodificador de áudio, a saber, o desformatador de carga útil de corrente de bits 220, o decodificador/quantificador inverso/reescalador 230, o pré-processador espectral 240 e o processador posterior de domínio de tempo 260 podem ser considerados opcionais, mas podem estar presentes em algumas implementações do decodificador de áudio 200.

A seguir, detalhes em relação à seleção da janela para transformação/janelamento realizados pelo transformador/janelador 254 serão descritos. Entretanto, em relação à importância da escolha de diferentes janelas, é feita referência às explicações acima.

O decodificador de áudio 200 é preferivelmente capaz de utilizar os tipos de janelas "only_long_sequence", "long_start_sequence", "eight_short_sequence", "long_stop_sequence" e "stop_start_sequence" descritos acima. Entretanto, o decodificador de áudio pode ser opcionalmente capaz de utilizar tipos adicionais de janelas, por exemplo, a denominada "stop_1152_sequence" e a denominada "stop_start_1152_sequence" (ambas podem ser utilizadas para uma transição de uma estrutura codificada de domínio de previsão linear para estrutura codificada de domínio de frequência). Além disso, o decodificador de áudio 200 pode ser adicionalmente configurado para utilizar tipos adicionais de janelas, como, por exemplo, os tipos de janelas 362, 366, 368, 382, que podem ser todos adaptados para uma transição de

uma estrutura codificada de domínio de frequência para uma estrutura codificada de domínio de previsão linear. Entretanto, o uso de tipos de janelas 330, 332, 362, 366, 368, 382 pode ser considerado opcional.

5 Entretanto, é uma característica importante do decodificador de áudio inventivo prover uma solução particularmente eficiente para derivar o tipo de janela adequado da informação de extensão de palavra-código variável de janela 224. Conforme discutido acima, isso será explicado adicionalmente
10 abaixo, tendo como referência as Figuras 10a a 10e.

 A informação de extensão de palavra-código variável de janela 224 compreende tipicamente 1 ou 2 bits por estrutura. Preferivelmente, a informação de extensão de palavra-código variável de janela compreende um primeiro bit que carrega a
15 informação de "window_length" da estrutura atual e um segundo bit que carrega a informação de "transform_length" da estrutura atual, onde a presença do segundo bit (bit de "transform_length") depende do valor do primeiro bit (bit de "window_length"). Portanto, o seletor de janela 270 é configurado para avaliar de maneira
20 seletiva uma ou duas informações de janela bits ("window_length" e "transform_length") para decidir sobre o tipo de janela associado à estrutura atual dependendo do valor do bit de "window_length" associado à estrutura atual. Todavia, na ausência do bit de "transform_length", o seletor de janela 270 pode presumir
25 naturalmente que o bit de "transform_length" considera um valor padrão.

 Em uma configuração preferida, o seletor de janela 270 pode ser configurado para avaliar a sintaxe, conforme

descrito acima em relação à Figura 6a, e prover a informação de janela para 272, de acordo com a referida sintaxe.

Presumindo, primeiro, que o decodificador de áudio 200 sempre opera em um modo central de domínio de frequência, ou seja, que não há alternância entre o modo central de domínio de frequência e o modo central de domínio de previsão linear pode ser suficiente para distinguir os cinco tipos de janelas descritos acima ("only_long_sequence", "long_start_sequence", "long_stop_sequence", "stop_start_sequence" e "eight_short_sequence"). Nesse caso, a informação de "window_length" da estrutura anterior, a informação de "window_length" da estrutura atual e a informação de "transform_length" da estrutura atual (se disponível) podem ser suficientes para decidir sobre o tipo de janela.

Por exemplo, presumindo a operação no modo central de domínio de frequência somente (pelo menos sobre uma sequência de três estruturas subsequentes), pode-se concluir do fato de que a informação de "window_length" da estrutura anterior indica uma inclinação de transição longa (valor "0") e que a informação de "window_length" da estrutura atual indica uma inclinação de transição longa (valor "0") que o tipo de janela "only_long_sequence" é associado à estrutura atual sem avaliar a informação de "transform_length", que não é transmitida pelo codificador nesse caso.

Novamente presumindo uma operação no modo central de domínio de frequência somente, pode-se concluir que, a parti do fato que a informação de "window_length" da estrutura anterior indica uma inclinação de transição longa (lado direito), e a

partir do fato que a informação de "window_length" da estrutura atual indica uma inclinação de transição curta (lado direito) (valor "1"), que o tipo de janela "long_start_sequence" é associado à estrutura atual, mesmo sem avaliar a informação de "transform_length" de uma estrutura atual (que pode ou não ser gerada e/ou transmitida pelo codificador nesse caso).

Novamente presumindo uma operação no modo central de domínio de frequência somente, pode-se concluir a partir do fato que a informação de "window_length" da estrutura anterior indica a presença de uma inclinação de transição curta (lado direito) (valor "1") e que a informação de "window_length" da estrutura atual indica uma inclinação de transição longa (lado direito) (valor "0"), que o tipo de janela "long_start_sequence" é associado à estrutura atual, mesmo sem avaliar a informação de "transform_length" da estrutura atual (que tipicamente não é provida pelo codificador de áudio correspondente de qualquer forma).

Se, entretanto, a informação de "window_length" da estrutura anterior indicar a presença de uma inclinação de transição curta (lado direito) e a informação de "window_length" da estrutura atual também indicar a presença de uma inclinação de transição curta (valor "1"), poderia ser necessário avaliar a informação de "transform_length" da estrutura atual. Nesse caso, se a informação de "transform_length" da estrutura atual considerar um primeiro valor (por exemplo, zero), o tipo de janela "stop_start_sequence" é associado à estrutura atual. De outra forma, ou seja, se a informação de "transform_length" da estrutura atual considerar um segundo valor (por exemplo, um), pode-se

concluir que o tipo de janela "eight_short_sequence" é associado à estrutura atual.

Para resumir o que foi descrito acima, o seletor de janela 270 é configurado para avaliar a informação de "window_length" da estrutura anterior e a informação de "window_length" da estrutura atual, a fim de determinar o tipo de janela associado à estrutura atual. Além disso, o seletor de janela 270 é configurado para levar em consideração de maneira seletiva, dependendo do valor da informação de "window_length" da estrutura atual (e, possivelmente, também dependendo de uma informação de "window_length" da estrutura anterior, ou uma informação de modo central), a informação de "transform_length" da estrutura atual para determinar o tipo de janela associado à estrutura atual. Portanto, o seletor de janela 270 é configurado para avaliar uma informação de extensão de palavra-código variável de janela, a fim de determinar o tipo de janela associado à estrutura atual.

A Figura 6c apresenta uma tabela que representa um mapeamento da informação de "transform_length" da estrutura anterior, a informação de "window_length" da estrutura atual e a informação de "transform_length" da estrutura atual em um tipo de janela da estrutura atual. A informação de "window_length" da estrutura atual e a informação de "transform_length" da estrutura atual podem ser representadas por uma informação de extensão de palavra-código variável de janela 224. O tipo de janela da estrutura atual pode ser representado por uma informação de janela 272. O mapeamento descrito pela tabela da Figura 6c pode ser realizado pelo seletor de janela 270.

Como pode ser visto, o mapeamento pode depender do modo central anterior. Se o modo central anterior for um "modo central de domínio de frequência" (abreviado por "FD"), o mapeamento pode tomar a forma, conforme discutido acima. Se, 5 entretanto, o modo central anterior for um "modo central de domínio de previsão linear" (abreviado por "LPD"), o mapeamento pode ser alterado, conforme pode ser observado nas duas últimas linhas da tabela da Figura 6c.

Além disso, o mapeamento pode ser alterado se o 10 modo central subsequente (ou seja, o modo central associado à estrutura subsequente) não for um modo central de domínio de frequência, mas um modo central de domínio de previsão linear.

O decodificador de áudio 200 pode opcionalmente compreender um separador de corrente de bits configurado para 15 separar a corrente de bits 210 representando as informações de áudio codificado e para extrair da corrente de bits uma informação de extensão de inclinação de janela de um bit (também designada aqui como informação de "window_length") e para extrair de maneira seletiva, dependendo de um valor da informação de extensão de 20 inclinação de janela de um bit, uma informação de extensão de transformação de um bit (designadas aqui como informação de "transform_length"). Nesse caso, o seletor de janela 270 é configurado para utilizar ou descuidar, de maneira seletiva, dependendo de uma informação de extensão de inclinação de janela 25 da estrutura atual, a informação de extensão de transformação, a fim de selecionar um tipo de janela para um processamento de uma determinada parte (por exemplo, estrutura) da representação "tempo-frequência" 242. O separador de corrente de bits pode, por

exemplo, ser parte do desformatador de carga útil de corrente de bits 220, e pode permitir que o decodificador de áudio 200 controle adequadamente a informação de extensão de palavra-código variável de janela, conforme discutido acima e também conforme descrito em relação às Figuras 10a a 10e.

Alternância entre modo central de domínio de frequência e modo central de domínio de tempo

Em algumas configurações, o codificador de áudio 100 e o decodificador de áudio 200 pode ser configurado para alternar entre um modo central de domínio de frequência e um modo central de domínio de previsão linear. Conforme explicado acima, presume-se que o modo central de domínio de frequência é o modo central básico, para o qual as explicações acima são mantidas. Entretanto, se o codificador de áudio é capaz de alternar entre o modo central de domínio de frequência e o modo central de domínio de previsão linear, ainda pode haver uma cross-fade (no sentido de uma operação de sobreposição-e-adição) entre estruturas codificadas no modo central de domínio de frequência e estruturas codificadas no modo central de domínio de previsão linear. Da mesma forma, janelas adequadas devem ser selecionadas a fim de garantir uma cross-fade adequada entre estruturas que são codificadas em diferentes modos centrais. Por exemplo, em algumas configurações, pode haver dois tipos de janelas, a saber, tipos de janelas 330 e 332 apresentados na Figura 2B, que são adaptadas para uma transição de um modo central de domínio de previsão linear para um modo central de domínio de frequência. Por exemplo, o tipo de janela 330 pode permitir uma transição entre uma estrutura codificada de domínio de previsão linear e uma estrutura

codificada de domínio de frequência que tem uma inclinação de transição longa do lado esquerdo, por exemplo, da estrutura codificada de domínio de previsão linear para uma estrutura codificada de domínio de frequência utilizando um tipo de janela

5 de "only_long_sequence" ou um tipo de janela de "long_start_sequence". Semelhantemente, o tipo de janela 332 pode permitir uma transição de uma estrutura codificada de domínio de previsão linear para uma estrutura codificada de domínio de frequência tendo uma inclinação de transição do lado esquerdo

10 curta (por exemplo, de uma estrutura codificada de domínio de previsão linear para uma estrutura tendo associado o tipo de janela de "eight_short_sequence" ou "long_start_sequence" ou "stop_start_sequence"). Da mesma forma, o seletor de janela 270

15 pode ser configurado para selecionar o tipo de janela 330, se for descoberto que a estrutura anterior (anterior à estrutura atual) é codificada no domínio de previsão linear, que a estrutura atual é codificada no domínio de frequência e que uma informação de "window_length" da estrutura atual indica uma inclinação de transição longa do lado direito da estrutura atual (por exemplo,

20 valor "0"). Ao contrário, o seletor de janela 270 é configurado para selecionar o tipo de janela 332 para a estrutura atual, se for descoberto que a estrutura anterior é codificada no domínio de previsão linear, que uma estrutura atual é codificada no domínio de frequência e que uma informação de "window_length" da estrutura

25 atual indica que uma inclinação de transição longa do lado direito é associada à estrutura atual (por exemplo, valor "1").

Semelhantemente, o seletor de janela 270 pode ser configurado para reagir ao fato de que a estrutura subsequente

(seguinte à estrutura atual) é codificada no domínio de previsão linear, enquanto a estrutura atual é codificada no domínio de frequência. Nesse caso, o seletor de janela 270 pode selecionar um dos tipos de janelas 362, 366, 368, 384, que são adaptados para serem seguidos por uma estrutura codificada de domínio de previsão linear, ao invés de um dos tipos de janelas 312, 316, 118, 332, que são adaptados para serem seguidos por uma estrutura codificada de domínio de frequência. Entretanto, exceto para a substituição do tipo de janela 312 pelo tipo de janela 362, a substituição do tipo de janela 318 pelo tipo de janela 368, a substituição do tipo de janela 360 pelo tipo de janela 366 e a substituição do tipo de janela 332 pelo tipo de janela 382, a seleção do tipo de janela pode ser inalterada quando comparada a uma situação na qual há somente estruturas codificadas de domínio de frequência.

Portanto, o mecanismo inventivo de utilização de uma informação de extensão de palavra-código variável de janela pode ser aplicada mesmo no caso em que as transições entre um domínio de codificação de frequência e uma codificação de previsão linear ocorrer, sem comprometer significativamente a eficiência da codificação.

Detalhes da sintaxe da corrente de bits

A seguir, os detalhes em relação à sintaxe da corrente de bits da corrente de bits 192, 210 serão discutidos, tendo como referência as Figuras 10a a 10e. A Figura 10a apresenta uma representação de sintaxe do denominado bloco de dados brutos "USAC_raw_data_block" da codificação de áudio e fala unificada ("USAC"). Como pode ser visto, o bloco de dados brutos USAC pode compreender um denominado elemento de canal único

`("single_channel_element()")` e/ou um elemento de par de canais `("channel_pair_element()")`. Entretanto, o bloco de dados brutos USAC pode compreender naturalmente mais do que um único elemento de canal e/ou mais do que um elemento de par de canais.

5 Agora, tendo como referência a Figura 10b, que apresenta uma representação de sintaxe de um elemento de canal único, alguns detalhes a mais serão explicados. Como pode ser visto na Figura 10b, um elemento de canal único pode compreender uma informação de modo central, por exemplo, na forma de um bit de

10 `"core_mode"`. A informação de modo central pode indicar se a estrutura atual é codificada em um modo central de domínio de previsão linear ou em um modo central de domínio de frequência. No caso em que a estrutura atual for codificada no modo central de domínio de previsão linear, o elemento de canal único pode

15 compreender uma corrente de canal de domínio de previsão linear `("LPD_channel_stream()")`. No caso em que estrutura atual é codificada no domínio de frequência, o elemento de canal único pode compreender uma corrente de canal de domínio de frequência `("FD_channel_stream()")`.

20 Agora, tendo como referência a Figura 10c, que apresenta uma representação de sintaxe de um elemento de par de canais, alguns detalhes adicionais serão explicados. Um elemento de par de canais pode compreender uma primeira informação de modo central, por exemplo, na forma de um bit de `"core_mode0"`, que

25 descreve um modo central do primeiro canal. Além disso, o elemento de par de canais pode compreender uma segunda informação de modo central na forma de um bit de `"core_mode1"`, que descreve um modo central do segundo canal. Portanto, modos centrais diferentes ou

idênticos podem ser selecionados para dois canais descritos por um elemento de par de canais. Opcionalmente, o elemento de par de canais pode compreender uma informação ICS comum ("ICS_info()") para ambos os canais. Essa informação ICS comum é vantajosa se a

5 configuração de dois canais descritos pelo elemento de par de canais for muito similar. Naturalmente, a informação ICS comum é somente utilizada preferivelmente se ambos os canais forem codificados no mesmo modo central.

Além disso, o elemento de par de canais

10 compreende uma corrente de canal de domínio de previsão linear ("LPD_channel_stream()") ou uma corrente de canal de domínio de frequência ("FD_channel_stream()") associada ao primeiro canal dependendo do modo central definido para o primeiro canal (por uma informação de modo central "core_mode0").

15 Também, o elemento de par de canais compreende a corrente de canal de domínio de previsão linear ("LPD_channel_stream()") ou uma corrente de canal de domínio de frequência ("FD_channel_stream()") para o segundo canal dependendo do modo central utilizado para codificar o segundo canal (que pode

20 ser sinalizado por uma informação de modo central "core_model").

Agora, tendo como referência a Figura 10d, que apresenta uma sintaxe para uma representação da informação ICS, alguns detalhes adicionais serão descritos. Deve ser observado que a informação ICS pode ser incluída no elemento de par de canais,

25 ou nas correntes de canal de domínio de frequência individuais (conforme discutido em relação à Figura 10e).

A informação ICS compreende uma informação de "window_length" de um bit (ou bit único), que descreve uma

extensão de uma inclinação de transição do lado direito da janela associada à estrutura atual, por exemplo, de acordo com a definição dada na Figura 7a. Se, e somente se, a informação de "window_length" considerar um valor predeterminado (por exemplo, "1"), a informação ICS compreende uma informação de "transform_length" de um bit (ou bit único) adicional. A informação de "transform_length" descreve um tamanho de um núcleo de MDCT, por exemplo, de acordo com a definição dada na Figura 7b. Se uma informação de "window_length" considerar um valor diferente daquele predeterminado (por exemplo, o valor "0"), a informação de "transform_length" não é incluída na (ou omitida da) informação ICS (ou na corrente de bits correspondente). Entretanto, nesse caso, um separador de corrente de bits de um decodificador de áudio pode ajustar o valor recuperado de uma "transform_length" variável de decodificador para um valor padrão (por exemplo "0").

Além disso, a informação ICS pode compreender uma denominada informação "window_shape", que pode ser uma informação de um bit (ou de bit único) que descreve uma forma de uma transição de janela. Por exemplo, a informação "window_shape" pode descrever se uma transição de janela tem uma forma de seno/cosseno ou uma forma derivada de Kaiser Bessel. Para detalhes em relação ao significado da informação "window_shape", é feita referência, por exemplo, ao padrão internacional ISO/IEC 14496-3:2005 (E), parte 3, subparte 4. Entretanto, deve ser observado que a informação "window_shape" deixa o tipo de janela básico não afetado e que as características gerais (inclinação de transição longa ou inclinação de transição curta; extensão de transformação longa ou extensão de transformação curta) são deixadas não

afetadas por uma informação "window_shape".

Portanto, nas configurações, de acordo com a invenção, a "forma de janela", ou seja, a forma das transições, é determinada separadamente do tipo de janela, ou seja, a extensão
5 geral das inclinações de transições (longas ou curtas) e a extensão de transformação (longa ou curta).

Além disso, a informação ICS pode compreender uma informação de fator de escala dependente do tipo de janela. Por exemplo, se uma informação de "window_length" e uma informação de
10 "transform_length" indicar que o tipo de janela atual é "eight_short_sequence", a informação ICS pode compreender uma informação "max_sfb" que descreve uma banda de fator de escala máxima e uma informação de "scale_factor_grouping" que descreve um agrupamento das bandas de fator de escala. Detalhes em relação a
15 essas informações são descritos, por exemplo, no padrão internacional ISO/IEC 14496-3:2005 (E), parte 3, subparte 4. De maneira alternativa, ou seja, se uma informação de "window_length" e uma informação de "transform_length" indicarem que a estrutura atual não é um tipo de janela de "eight_short_sequence", a
20 informação ICS pode compreender uma informação "max_sfb" somente (mas nenhuma informação de "scale_factor_grouping").

A seguir, alguns detalhes adicionais serão descritos tendo como referência a Figura 10e, que apresenta uma representação de sintaxe de uma corrente de canal de domínio de
25 frequência ("FD_channel_stream()"). A corrente de canal de domínio de frequência compreende uma informação de "global_gain" que descreve um ganho global associado aos valores espectrais. Além disso, a corrente de canal de domínio de frequência compreende uma

informação ICS ("ICS_info()"), a menos que essas informações já estejam incluídas em um elemento de par de canais compreendendo a corrente de canal de domínio de frequência presente. Em relação à informação ICS, foram descritos detalhes em relação à Figura 10d.

5 Além disso, a corrente de canal de domínio de frequência compreende dados de fator de escala ("ac_spectral_data scale_factor_data()"), que descreve um escalonamento a ser aplicado aos valores (ou bandas de fator de escala) da informação de valor espectral decodificado ou da
10 representação "tempo-frequência". Além disso, a corrente de canal de domínio de frequência compreende dados espectrais codificados, que podem, por exemplo, dados espectrais codificados aritmeticamente (ac_spectral_data()). Entretanto, uma codificação diferente dos dados espectrais pode ser utilizada. Em relação aos
15 dados de fator de escala e aos dados espectrais codificados, novamente é feita referência ao padrão internacional ISO/IEC 14496-3: 2005 (E), parte 3, subparte 4. Entretanto, diferentes codificações dos dados de fator de escala e dos dados espectrais podem naturalmente serem aplicadas, se desejado.

20 Conclusões e Avaliações de desempenho

A seguir, são feitas algumas conclusões, e uma avaliação de desempenho do conceito inventivo será dada. As configurações da presente invenção criam um conceito para uma redução da taxa de bits necessária, que pode ser aplicado, por
25 exemplo, em combinação com esquemas de codificação de áudio definidos no padrão internacional ISO/IEC 14496-3:2005 (E), parte 3, subparte 4. Entretanto, o conceito discutido aqui também pode ser utilizado em combinação com a denominada abordagem de

"codificação de fala e áudio unificada" (USAC). Com base nas definições de corrente de bits existentes e arquiteturas de decodificador, a presente invenção cria uma modificação da sintaxe da corrente de bits, que simplifica a sintaxe da sinalização de sequências de janelas, preserva taxa de bits sem o aumento da complexidade e não altera a forma de onda de saída do decodificador.

A seguir, o retrospecto e ideia fundamentais à presente invenção serão brevemente discutidos e resumidos. Na codificação de áudio atual, de acordo com o ISO/IEC 14496-3:2005 (E) parte 3, subparte 4, e também no projeto de trabalho USAC, uma palavra-código com uma extensão fixa de dois bits é enviada para sinalizar a sequência de janelas. Adicionalmente, a informação de sequência de janelas da estrutura anterior é, algumas vezes, necessária para determinar a sequência correta.

Entretanto, descobriu-se que ao levar em conta essas informações e ao criara a variável de extensão de palavra-código (um ou dois bits), a taxa de bits pode ser reduzida. Uma nova palavra-código tem uma extensão mínima de dois bits ("window_length" e, em alguns casos, "transform_length"). Portanto, a taxa de bits nunca é aumentada (quando comparado à abordagem convencional).

A nova palavra-código ("window_length" e, em alguns casos, "transform_length") consiste em um bit ("window_length") que indica a extensão da inclinação de janela direita e um bit ("transform_length") que indica a extensão de transformação. Em muitos casos, a extensão de transformação pode ser derivada de maneira não ambígua pela informações da estrutura

anterior, a saber, sequência de janelas e modo central. Portanto, não é necessário retransmitir essas informações. Da mesma forma, o bit de "transform_length" é omitido nesses casos, levando, assim, a uma redução da taxa de bits.

5 A seguir, alguns detalhes em relação à proposta para uma nova sintaxe da corrente de bits, de acordo com a presente invenção, serão discutidos. A nova sintaxe proposta da corrente de bits permite uma implementação mais simples e sinalização das sequências de janelas, porque transmite somente as
10 informações realmente necessárias para determinar a sequência de janelas da estrutura atual, ou seja, a inclinação de janela direita e a extensão de transformação. A inclinação de janela esquerda da estrutura atual é derivada da inclinação de janela direita da estrutura anterior.

15 A proposta (ou a nova corrente de bits proposta) separa explicitamente informações sobre a extensão da inclinação de janela (informação de "window_length") e sobre a extensão de transformação (informação de "transform_length"). A palavra-código de extensão variável é uma combinação de ambas, onde a primeira
20 "window_length" de bit determina a extensão da inclinação de janela direita (da estrutura atual) e a segunda "transform_length" de bit determina a extensão da MDCT (para a estrutura atual), de acordo com Figuras 7a e 7d. No caso da "window_length" ==0, ou seja, a inclinação de janela longa for selecionada, a transmissão
25 da "transform_length" pode ser omitida (ou ser realmente omitida), uma vez que um tamanho de núcleo de MDCT de 1024 amostras (ou 1152 amostras em alguns casos) é obrigatório.

A Figura 7c dá uma visão geral sobre todas as

combinações de "window_length" e "transform_length". Como pode ser visto, há somente três combinações significativas dos itens de informações de um bit "window_length" e "transform_length", de maneira que a transmissão da "transform_length" possa ser omitida se uma informação de "window_length" considerar o valor zero sem afetar negativamente a transmissão da informação desejada.

A seguir, o mapeamento da informação de "window_length" e uma informação de "transform_length" para uma informação de "window_sequence" (que descreve um tipo de janela a ser utilizado para a estrutura atual) será resumido. A tabela da Figura 6a apresenta como a "window_sequence" de elemento de corrente de bits da situação atual dos projetos de trabalho previstos do padrão USAC pode ser derivada de novos elementos de corrente de bits propostos. Isso demonstra que a alteração proposta é "transparente" em termos de conteúdo de informações.

Em outras palavras, a sintaxe reduzida de taxa de bits inventiva para sinalizar o tipo de janela, que tem base no uso de uma informação de extensão de palavra-código variável de janela, é capaz de carregar o conteúdo de informação "completo", que é convencionalmente transmitido utilizando uma taxa de bits maior. Também, o conceito inventivo pode ser aplicado nos codificadores e decodificadores de áudio e convencionais, por exemplo, o codificador de áudio ou decodificador de áudio, de acordo com o ISO/IEC 14496-3:2005 (E), parte 3, subparte 4 ou de acordo com o projeto de trabalho USAC atual sem qualquer modificação importante.

A seguir, uma avaliação da preservação de bits alcançável será apresentada. Entretanto, deve ser observado que em

alguns casos a preservação de bits pode ser um tanto menor do que o indicado, e que, em outros casos, a preservação de bits pode ser até significativamente maior do que a preservação de bits discutida. A "avaliação da preservação de bits" apresentado na

5 Figura 9 apresenta a avaliação da preservação de bits para uma transcodificação sem perdas, em comparação às correntes de bits que usam a nova sintaxe da corrente de bits para correntes de bits convencionais (com correntes de bits convencionais que foram submetidas para uma apresentação de propostas). Como pode ser

10 visto claramente, a transmissão do bit de "transform_length" pode ser omitida, de acordo com a invenção, em 95,67 % de todas as estruturas de domínio de frequência para mono de 12 kbps e até 95,15 % de todas as estruturas de domínio de frequência para 64 kbps.

15 Como pode ser visto da Figura 9, entre 2 e 24 bits por segundo podem ser salvos em média, sem comprometer a qualidade do conteúdo de áudio. Tendo em vista o fato de que a taxa de bits é um recurso muito importante para o armazenamento e transmissão de um conteúdo de áudio, essa melhora pode ser

20 considerada muito valiosa. Também, deve ser observado que em alguns casos, a melhora na taxa de bits pode ser significativamente maior, por exemplo, se estruturas forem escolhidas para serem comparavelmente menores.

Para resumir o exposto acima, a presente invenção

25 propõe uma nova sintaxe da corrente de bits para sinalizar as sequências de janelas. A nova sintaxe da corrente de bits preserva a taxa de dados e é mais lógica e mais flexível em comparação à sintaxe antiga. Ela é fácil de se implementar e não apresenta

inconvenientes em relação à complexidade.

Comparação com o Projeto de trabalho USAC atual

A seguir, as alterações de texto propostas para uma descrição técnica do projeto de trabalho USAC atual serão discutidas. A fim de incorporar as alterações inventivas propostas, de acordo com a presente invenção, as seguintes seções precisam ser atualizadas:

Na definição pendente de "cargas úteis para USAC do tipo de objeto de áudio", nas quais a sintaxe da denominada informação ICS é descrita, a sintaxe convencional deve ser substituída pela sintaxe apresentado na Figura 10b.

Também, o "elemento de dados" "window_sequence" deve ser substituído pela seguinte definição "window_length" e "transform_length" dos elementos de dados:

15 window_length: um campo de um bit que determina qual extensão de inclinação de janela é utilizada a parte da direção direita dessa sequência de janelas; e

transform_length: um campo de um bit que determina qual extensão de transformação é utilizada para essa sequência de janelas.

Além disso, a definição da "window_sequence" de elemento de ajuda deve ser adicionada como segue:

25 window_sequence: indica uma sequência de janelas conforme definida pela "window_length" da estrutura anterior, a "transform_length" e a "window_length" da estrutura atual e o "core_mode" da seguinte estrutura, de acordo com a tabela apresentada na Figura 8.

A Figura 8 apresenta a definição da

"window_sequence" de elemento de ajuda, que pode opcionalmente ser derivada de a informação de "window_length" da estrutura anterior, a informação de "window_length" da estrutura atual, a informação de "transform_length" da estrutura atual e a informação de "modo central" da estrutura seguinte.

Ademais, a definição convencional da "window_sequence" e da "window_shape" pode ser substituída por definições mais adequadas de "window_length", "transform_length" e "window_shape" como segue:

10 window_length: um campo de um bit que determina qual extensão de inclinação de janela é utilizada para a parte da direção direita dessa janela;

 transform_length: um campo de um bit que determina qual extensão de transformação é utilizada para essa
15 janela; e

 window_shape: um bit que indica qual função de janela é selecionada.

Método, de acordo com Figura 11

A Figura 11 apresenta uma fluxograma de um método
20 para prover uma informação de áudio codificado com base em uma informação de áudio de entrada. O método 1100, de acordo com Figura 11, compreende uma etapa 1110 para prover uma sequência de parâmetros de sinal de áudio com base em uma pluralidade de partes janeladas das informações de áudio de entrada. Ao prover uma
25 sequência de parâmetros de sinal de áudio, é realizada uma alternância entre um uso de janelas que têm a uma inclinação de transição mais longa e janelas que têm a inclinação de transição mais curta, e também entre um uso de janelas que têm associadas a

si duas ou mais extensões de transformação diferentes, a fim de adaptar um tipo de janela para obter as partes janeladas das informações de áudio de entrada dependendo das características das informações de áudio de entrada. O método 1100 também compreende
5 uma etapa 1120 para codificar uma informação de janela que descreve um tipo de janela utilizado para transformar a parte atual das informações de áudio de entrada utilizando uma palavra-chave de extensão variável.

Método de acordo com Figura 12

10 A Figura 12 apresenta uma fluxograma de um método para prover uma informação de áudio decodificado com base em uma informação de áudio codificado. O método 1200, de acordo com Figura 12, compreende uma etapa 1210 para avaliar uma informação de extensão de palavra-código variável de janela a fim de
15 selecionar uma janela, de uma pluralidade de janelas compreendendo janelas de diferentes inclinações de transição e janelas que têm associadas a si extensões de transformação diferentes, for um processamento de uma determinada parte da representação "tempo-frequência" associada a uma determinada estrutura das informações
20 de áudio. O método 1200 também compreende uma etapa 1220 para mapear a determinada parte da representação "tempo-frequência", que é descrita pelas informações de áudio codificado, para a representação nodomínio de tempo utilizando a janela selecionada.

Deve ser observado que os métodos, de acordo com
25 Figuras 11 e 12, pode ser suplementado por qualquer uma das características e funcionalidades aqui descritas em relação aos aparelhos inventivos e às características da corrente de bits inventivas.

Alternativas de implementação

Embora alguns aspectos tenham sido descritos no contexto de um aparelho, é claro que esses aspectos também representam uma descrição do método correspondente, onde um bloco ou dispositivo corresponde a uma etapa do método ou uma característica de uma etapa do método. De maneira análoga, os aspectos descritos no contexto de uma etapa do método também representam uma descrição de um bloco correspondente ou item ou característica de um aparelho correspondente.

Qualquer uma das etapas do método inventivo pode ser realizada utilizando um microprocessador, um computador programável, um fpga ou qualquer outro hardware, como, por exemplo, um hardware de processamento de dados.

O sinal de áudio codificado inventivo pode ser armazenado em um meio de armazenamento digital ou pode ser transmitido em um meio de transmissão, como um meio de transmissão wireless ou meio de transmissão com cabo, como a Internet.

Dependendo de determinadas exigências de implementação, as configurações da invenção podem ser implementadas em hardware ou em software. A implementação pode ser realizada utilizando um meio de armazenamento digital, por exemplo, um disquete, um DVD, um Blue-Ray, um CD, uma ROM, uma PROM, uma EPROM, uma EEPROM ou uma memória FLASH, tendo sinais de controle possíveis de ler eletronicamente armazenados em si, que cooperam (ou são capazes de cooperar) com um sistema de computador programável de maneira que o respectivo método seja realizado. Portanto, o meio de armazenamento digital pode ser possível de ler por computador.

Algumas configurações, de acordo com a invenção, compreendem um carregador de dados que tem sinais de controle possíveis de ler eletronicamente, que são capazes de cooperar com um sistema de computador programável, de maneira que um dos
5 métodos aqui descrito seja realizado.

De modo geral, as configurações da presente invenção podem ser implementadas como um produto de programa de computador com um código de programa, o código de programa sendo operado para realizar um dos métodos quando o produto de programa
10 de computador for executado em um computador. O código de programa pode, por exemplo, ser armazenado em um carregador possível de ler de máquina.

Outras configurações compreendem um programa de computador para realizar um dos métodos aqui descritos,
15 armazenados em um carregador possível de ler de máquina.

Em outras palavras, uma configuração do método inventivo é, portanto, um programa de computador que tem um código de programa para realizar um dos métodos aqui descritos, quando o programa de computador for executado em um computador.

20 Uma configuração adicional dos métodos inventivos é, portanto, um carregador de dados (ou a meio de armazenamento digital, ou um meio possível de ler no computador) compreendendo, registrado em si, o programa de computador para realizar um dos métodos aqui descritos.

25 Uma configuração adicional do método inventivo é, portanto, uma corrente de dados ou uma sequência de sinais que representam o programa de computador para realizar um dos métodos aqui descritos. A corrente de dados ou um sequência de sinais

pode, por exemplo, ser configurada para ser transferida por meio de uma conexão de comunicação de dados, por exemplo, por meio da Internet.

Uma configuração adicional compreende meios de
5 processamento, por exemplo, um computador ou um dispositivo de lógica programável, configurado ou adaptado para realizar um dos métodos aqui descritos.

Uma configuração adicional compreende um
computador que tem instalado em si o programa de computador para
10 realizar um dos métodos aqui descritos.

Em algumas configurações, um dispositivo de
lógica programável (por exemplo, matriz de portas de campo
programáveis) pode ser utilizado para realizar algumas ou todas as
funcionalidades dos métodos aqui descritos. Em algumas
15 configurações, a matriz de portas de campo programáveis pode
cooperar com um microprocessador a fim de realizar um dos métodos
aqui descritos. De modo geral, os métodos são preferivelmente
realizados por qualquer aparelho de hardware.

As configurações descritas acima são meramente
20 ilustrativas para os princípios da presente invenção. Entende-se
que modificações e variações das organizações e os detalhes aqui
descritos serão aparentes a outros técnicos no assunto. Pretende-
se, portanto, limitar-se somente ao escopo das reivindicações da
patente iminentes e não pelos detalhes específicos apresentados
25 para fim de descrição e explicação das configurações da presente
invenção.

REIVINDICAÇÕES

1. Decodificador de áudio (200) para prover uma informação de áudio decodificado (212) com base em uma informação de áudio codificado (210), caracterizado por o decodificador de áudio compreender:

um transformador de sinal com base em janela (250) configurado para mapear a representação "tempo-frequência" (242) das informações de áudio, que é descrita pelas informações de áudio codificado (210), para uma representação no domínio do tempo (252) das informações de áudio,

em que o transformador de sinal com base em janela

é configurado para selecionar uma janela, de uma pluralidade de janelas (310, 312, 314, 316, 318) compreendendo janelas de diferentes inclinações de transição (310a, 312a, 314a, 316a, 318a, 310b, 312b, 314b, 316b, 318b) e janelas que têm associadas nelas diferentes extensões de transformação utilizando uma informação de janela (272);

em que o decodificador de áudio (200) compreende um seletor de janela (270) configurado para selecionar uma informação de extensão variável de palavra-código de janela (224) a fim de selecionar uma janela para um processamento de uma determinada parte da representação de "tempo-frequência" [*time-frequency representation*] associada a uma determinada estrutura das informações de áudio.

2. Decodificador de áudio (200), de acordo com a reivindicação 1, caracterizado por o decodificador de áudio compreender um separador de corrente de bits (220) configurado

para separar a corrente de bits (210) que representa as informações de áudio codificado e para extrair da corrente de bits (210) uma informação de extensão de inclinação de janela de um bit ("window_length") e para extrair de maneira seletiva, dependendo de um valor da informação de extensão de inclinação de janela de um bit, uma informação de extensão de transformação de um bit ("transform_length"); e

em que o seletor de janela (270) é configurado para usar ou abandonar, de maneira seletiva, dependendo de uma informação de extensão de inclinação de janela, uma informação de extensão de transformação a fim de selecionar um tipo de janela (310, 312, 314, 316, 318) para o processamento de uma determinada parte da representação "tempo-frequência" (242).

3. Decodificador de áudio (200), de acordo com uma das reivindicações 1 ou 2, caracterizado por o seletor de janela (270) ser configurado para selecionar um tipo de janela (310, 312, 314, 316, 318) para o processamento de uma parte atual da informação tempo-frequência (242), de maneira que a extensão de inclinação de janela do lado esquerdo da janela para processar a parte atual da representação "tempo-frequência" (242) seja combinada a uma extensão de inclinação de janela do lado direito de uma janela utilizado para processar uma parte anterior da representação "tempo-frequência" (242).

4. Decodificador de áudio (200), de acordo com a reivindicação 3, caracterizado por o seletor de janela (270) é configurado para selecionar entre um primeiro tipo (310) de janela e um segundo tipo (312) de janela dependendo de um valor da informação de extensão de inclinação de janela de um bit, se a

extensão de inclinação de janela do lado direito da janela para processar a parte anterior da representação tempo-frequência (242) considerar um valor longo e se uma parte anterior das informações de áudio, uma parte atual das informações de áudio e uma parte subsequente das informações de áudio, forem todas codificadas utilizando um modo central de domínio de frequência;

em que o seletor de janela (270) é configurado para selecionar um terceiro tipo (314) de janela em resposta ao primeiro valor da informação de extensão de inclinação de janela de um bit que indica uma inclinação de janela do lado direito longa, se extensão de inclinação de janela do lado direito da janela para processar uma parte anterior das informações de áudio considerar um valor curto e se a parte anterior das informações de áudio, a parte atual das informações de áudio e a parte subsequente das informações de áudio forem todas codificadas utilizando um modo central de domínio de frequência; e

em que o seletor de janela (270) é configurado para selecionar entre um quarto tipo (316) de janela e um quinto tipo (318) de janela, que define uma sequência de janelas curta (319a a 319h), dependendo de uma informação de extensão de transformação de um bit, se uma informação de extensão de inclinação de janela de um bit considerar um segundo valor que indica uma inclinação de janela do lado direito curta, se a extensão de inclinação de janela do lado direito da janela para processar a parte anterior das informações de áudio (242) considerar um valor curto e se a parte anterior das informações de áudio, a parte atual das informações de áudio e a parte subsequente das informações de áudio forem todas codificadas

utilizando um modo central de domínio de frequência;

em que o primeiro tipo (310) de janela compreende uma extensão de inclinação de janela do lado esquerdo comparavelmente longa, uma extensão de inclinação de janela do lado direito comparavelmente longa e uma extensão de transformação comparavelmente longa;

em que o segundo tipo de janela (312) compreende uma extensão de inclinação de janela do lado esquerdo comparavelmente longa, uma extensão de inclinação de janela do lado direito comparavelmente curta e uma extensão de transformação comparavelmente longa;

em que o terceiro tipo de janela (314) compreende a uma extensão de inclinação de janela do lado esquerdo comparavelmente curta, uma extensão de inclinação de janela do lado direito comparavelmente longa e uma extensão de transformação comparavelmente longa;

em que o quarto tipo de janela (316) compreende uma extensão de inclinação de janela do lado esquerdo comparavelmente curta, uma extensão de inclinação de janela do lado direito comparavelmente curta e uma extensão de transformação comparavelmente longa; e

em que a sequência de janelas (319a a 319h) do quinto tipo de janela (318) define uma superposição de uma pluralidade de janelas (319a a 319h) associada a uma única parte das informações de áudio (242), e em que cada uma das janelas (319a a 319h) da pluralidade de janelas compreende uma extensão de transformação comparavelmente curta, uma inclinação de janela do lado esquerdo comparavelmente curta e uma inclinação de janela do

lado direito comparavelmente curta.

5. Decodificador de áudio (200), de acordo com uma das reivindicações de 1 a 4, caracterizado por o seletor de janela (270) ser configurado para avaliar de maneira seletiva um bit de extensão de transformação da informação de extensão variável de palavra-código de janela (224) de uma parte atual das informações de áudio somente se um tipo de janela para o processamento de uma parte anterior das informações de áudio (242) compreender uma extensão de inclinação de janela do lado direito correspondente a uma extensão de inclinação de janela do lado esquerdo de uma sequência de janelas (318) de janelas curtas, e uma informação de extensão de inclinação de janela de um bit associada a uma parte atual da representação "tempo-frequência" (242) define uma extensão de inclinação de janela do lado direito correspondente à extensão de inclinação de janela do lado direito da sequência de janelas (318) de janelas curtas.

6. Decodificador de áudio (200), de acordo com uma das reivindicações de 1 a 5, caracterizado por o seletor de janela (270) ser configurado adicionalmente para receber uma informação de modo central anterior associada a uma estrutura anterior das informações de áudio e que descreve um modo central para codificar a estrutura anterior das informações de áudio; e

em que o seletor de janela (270) é configurado para selecionar um tipo de janela para o processamento de uma parte atual da representação tempo-frequência (242) dependendo de uma informação de modo central anterior e também dependendo de uma informação de extensão variável de palavra-código de janela (224) associada à parte atual das informações de áudio (242).

7. Decodificador de áudio (200), de acordo com uma das reivindicações de 1 a 6, caracterizado por o seletor de janela (270) ser configurado adicionalmente para receber uma informação de modo central subsequente associada a uma parte subsequente das informações de áudio (242) e que descreve um modo central para codificar a parte subsequente das informações de áudio; e

em que o seletor de janela (270) ser configurado para selecionar uma janela para o processamento de uma parte atual das informações de áudio (242) dependendo da informação de modo central subsequente e também dependendo da informação de janela de extensão variável de palavra-código (224) associada à parte atual da representação "tempo-frequência" (242).

8. Decodificador de áudio (200), de acordo com a reivindicação 7, caracterizado por o seletor de janela (270) ser configurado para selecionar janelas (362, 366, 368, 382) que têm uma inclinação do lado direito encurtada, se a informação de modo central subsequente indicar que uma parte subsequente das informações de áudio for codificada utilizando um modo central de domínio de previsão linear.

9. Codificador de áudio (100) para prover uma informação de áudio codificado (192) com base em uma informação de áudio de entrada (110), caracterizado por o codificador de áudio (100) compreender:

um transformador de sinal com base em janela (130) configurado para prover uma sequência de parâmetros de sinal de áudio (132) com base na pluralidade de partes janeladas das informações de áudio de entrada (110),

em que o transformador de sinal com base em janela (130) é configurado para adaptar tipos de janelas para obter as partes janeladas das informações de áudio de entrada dependendo das características das informações de áudio de entrada (110);

em que o transformador de sinal com base em janela (130) é configurado para alternar entre um uso de janelas (310, 312, 314, 316, 318) que têm uma inclinação de transição mais longa e janelas que têm uma inclinação de transição mais curta, e também para alternar entre um uso de janelas que têm duas ou mais extensões de transformação diferentes; e

em que o transformador de sinal com base em janela (130) é configurado para determinar um tipo de janela utilizado para transformar uma parte atual das informações de áudio de entrada dependendo de um tipo de janela utilizado para transformar uma parte anterior das informações de áudio de entrada e um conteúdo de áudio da parte atual das informações de áudio de entrada;

em que o codificador de áudio é configurado para codificar uma informação de janela (140) que descreve um tipo de janela utilizado para transformar a parte atual das informações de áudio de entrada (110) utilizando uma palavra-código de extensão variável;

em que o transformador de sinal baseado em janela (130) é configurado para transformar a informação de áudio de entrada (110), ou uma versão pré-processada do mesmo, para o domínio de "tempo-frequência"

10. Codificador de áudio (100), de acordo com a

reivindicação 9, caracterizado por o codificador de áudio ser configurado para prover a palavra-código de extensão variável de maneira que a palavra-código de extensão variável associada a uma determinada parte da representação tempo-frequência compreende a informação de bit único que descreve uma extensão de inclinação de janela de uma janela aplicada para obter a determinada parte da representação tempo-frequência (132); e

em que o codificador de áudio (100) é configurado para prover a palavra-código de extensão variável de maneira que a palavra-código de extensão variável compreende de maneira seletiva uma informação de extensão de transformação de bit único que descreve uma extensão de transformação aplicada para obter a determinada parte da representação tempo-frequência (132) se, e somente se, a informação de bit único que descreve a extensão de inclinação de janela considerar um valor predeterminado.

11. Codificador de áudio (100), de acordo com a reivindicação 9 ou com a reivindicação 10, caracterizado por o codificador de áudio ser configurado para codificar uma informação de extensão de inclinação de janela que descreve uma extensão de inclinação de janela do lado direito de uma janela aplicada para obter uma determinada parte da representação tempo-frequência e uma informação de extensão de transformação que descreve uma extensão de transformação aplicada para obter uma determinada parte da representação tempo-frequência (132) utilizando bits separados da corrente de bits (192), e para decidir sobre a presença de um bit que carrega uma informação de extensão de transformação dependendo do valor da informação de extensão de inclinação de janela.

12. Informação de áudio codificado, caracterizada por as informações de áudio codificado compreenderem:

uma representação "tempo-frequência" codificada (222) que descreve um conteúdo de áudio de uma pluralidade de partes janeladas de um sinal de áudio, caracterizado por as janelas de diferentes inclinações de transição e diferentes extensões de transformação são associadas às diferentes partes janeladas do sinal de áudio; e

uma informação de janela codificada (224) que codifica tipos de janelas utilizados para obter a representação "tempo-frequência" codificada (222) de uma pluralidade de partes janeladas do sinal de áudio,

em que uma informação de janela codificada (224) é uma informação de janela de extensão variável que codifica um ou mais tipos de janelas que utilizam um primeiro número menor de bits e que codifica um ou mais outros tipos de janelas que utilizam um segundo número maior de bits.

13. Informações de áudio codificado, de acordo com a reivindicação 12, caracterizadas por as informações de áudio codificado compreenderem unidades de informação de extensão de inclinação de janela de um bit associadas às partes janeladas correspondentes de um sinal de áudio codificado utilizando um modo central de domínio de frequência; e

unidades de informação de extensão de transformação de um bit associadas de maneira seletiva às partes janeladas do sinal de áudio para o qual uma informação de extensão de inclinação de janela de um bit considera um valor predeterminado.

14. Método (1200) para prover uma informação de áudio decodificado com base em uma informação de áudio codificado, caracterizado por o método compreender:

avaliar (12010) uma informação de extensão variável da palavra-código de janela a fim de selecionar uma janela, de uma pluralidade de janelas compreendendo janelas de diferentes inclinações de transição e janelas que têm associadas a elas diferentes extensões de transformação, para processar uma determinada parte de uma representação tempo-frequência associada a uma determinada estrutura das informações de áudio; e

mapear (1220) a determinada parte da representação tempo-frequência, que é descrita pelas informações de áudio codificado, para a representação no domínio de tempo utilizando a janela selecionada.

15. Método (1100) para prover uma informação de áudio codificado com base em uma informação de áudio de entrada, caracterizado por o método compreender:

prover (1110) uma sequência de parâmetros de sinal de áudio com base em uma pluralidade de partes janeladas das informações de áudio de entrada, caracterizado em que a informação de áudio de entrada, ou uma versão pré-processada (122) da mesma, é transformada para o domínio de "tempo-frequência", em que a alternância é realizada entre um uso de janelas que têm uma inclinação de transição mais longa e janelas que têm a inclinação de transição mais curta, e também entre um uso de janelas que têm associadas a elas duas ou mais extensões de transformação diferentes, para adaptar tipos de janelas para obter as partes janeladas das informações de áudio de entrada dependendo das

características das informações de áudio de entrada; e

codificar uma informação que descreve tipos de janelas utilizados para transformar partes das informações de áudio de entrada utilizando palavras-chave de extensão variável.

CODIFICADOR DE ÁUDIO

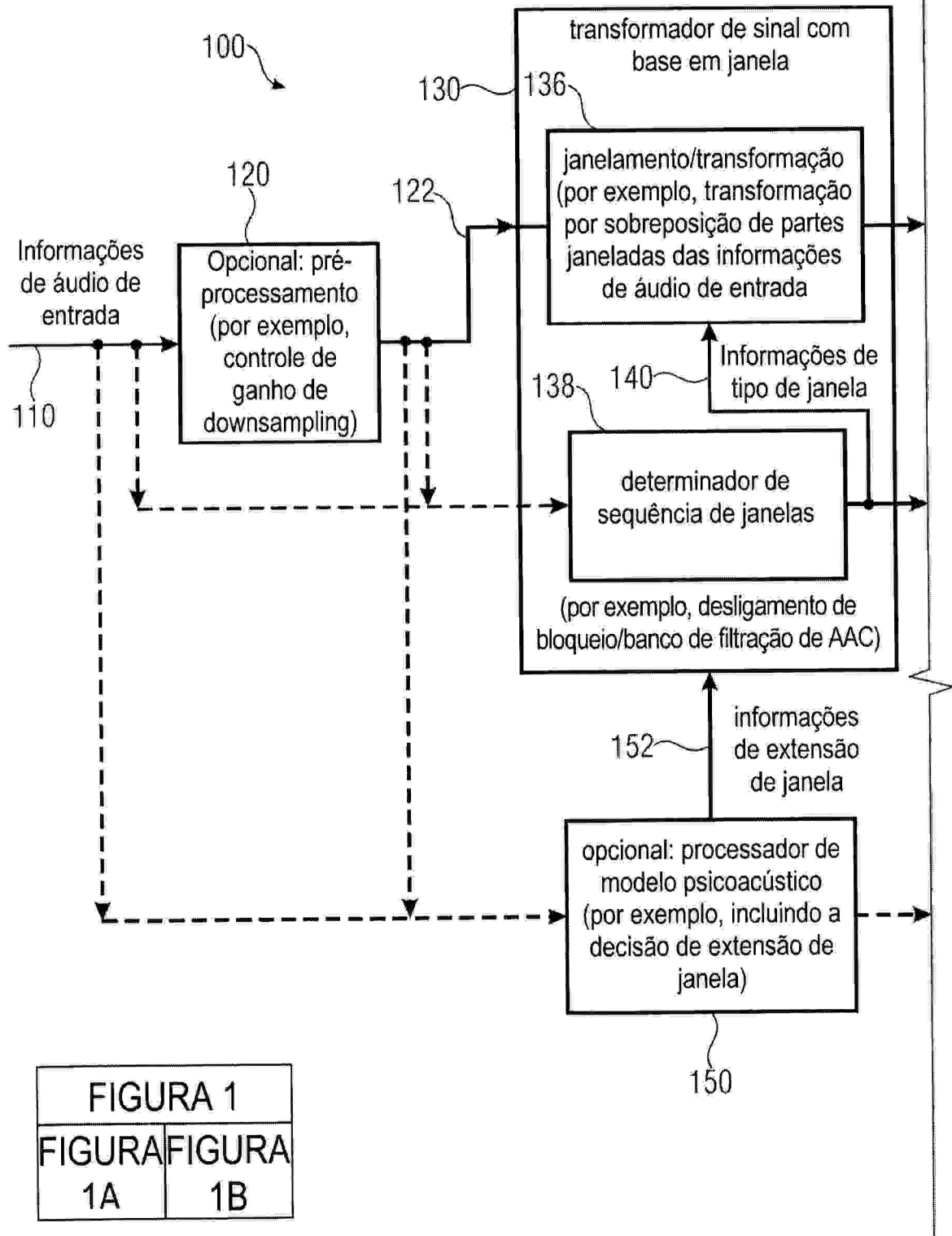


FIGURA 1A

CODIFICADOR DE ÁUDIO

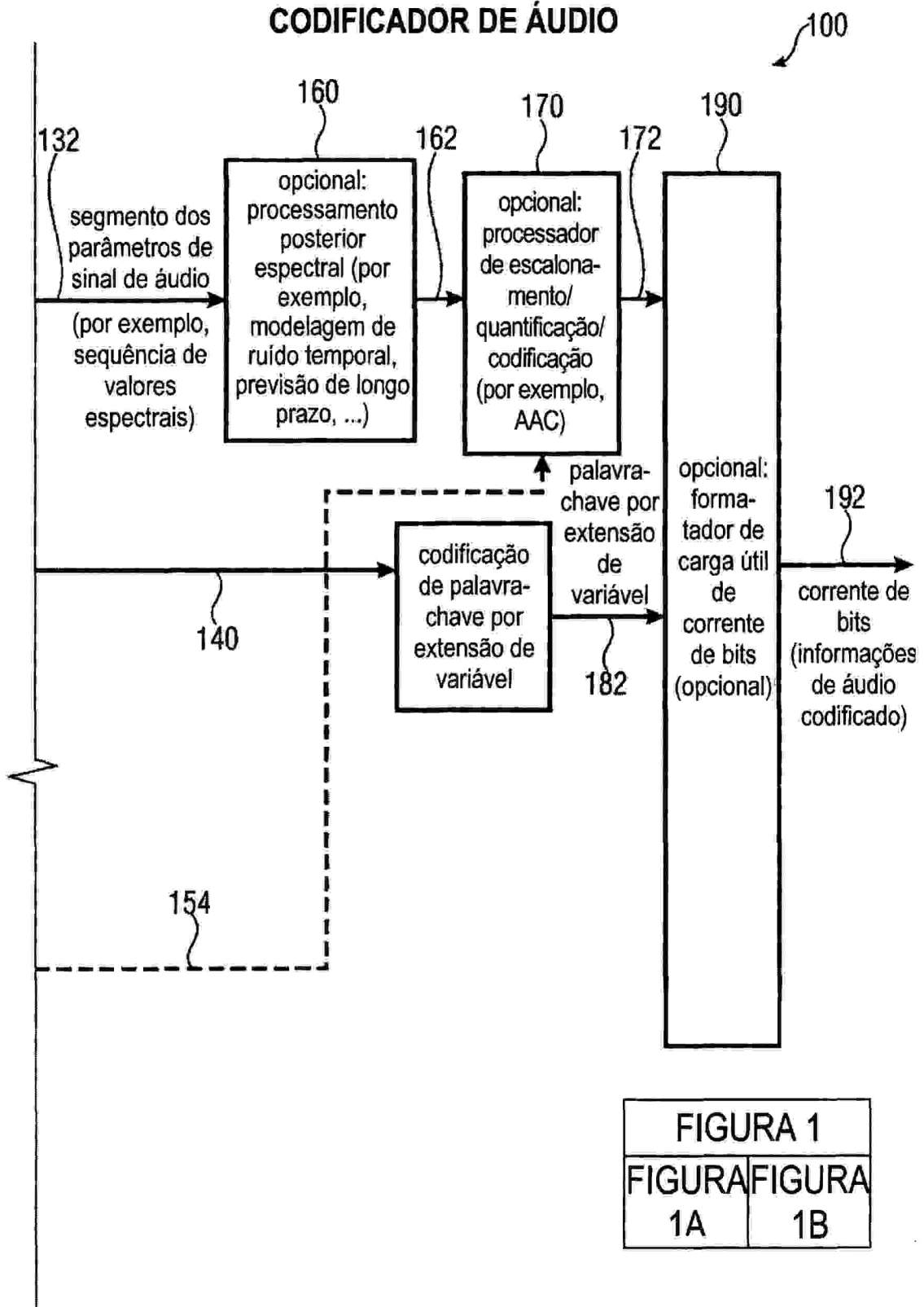


FIGURA 1	
FIGURA 1A	FIGURA 1B

FIGURA 1B

CODIFICADOR DE ÁUDIO

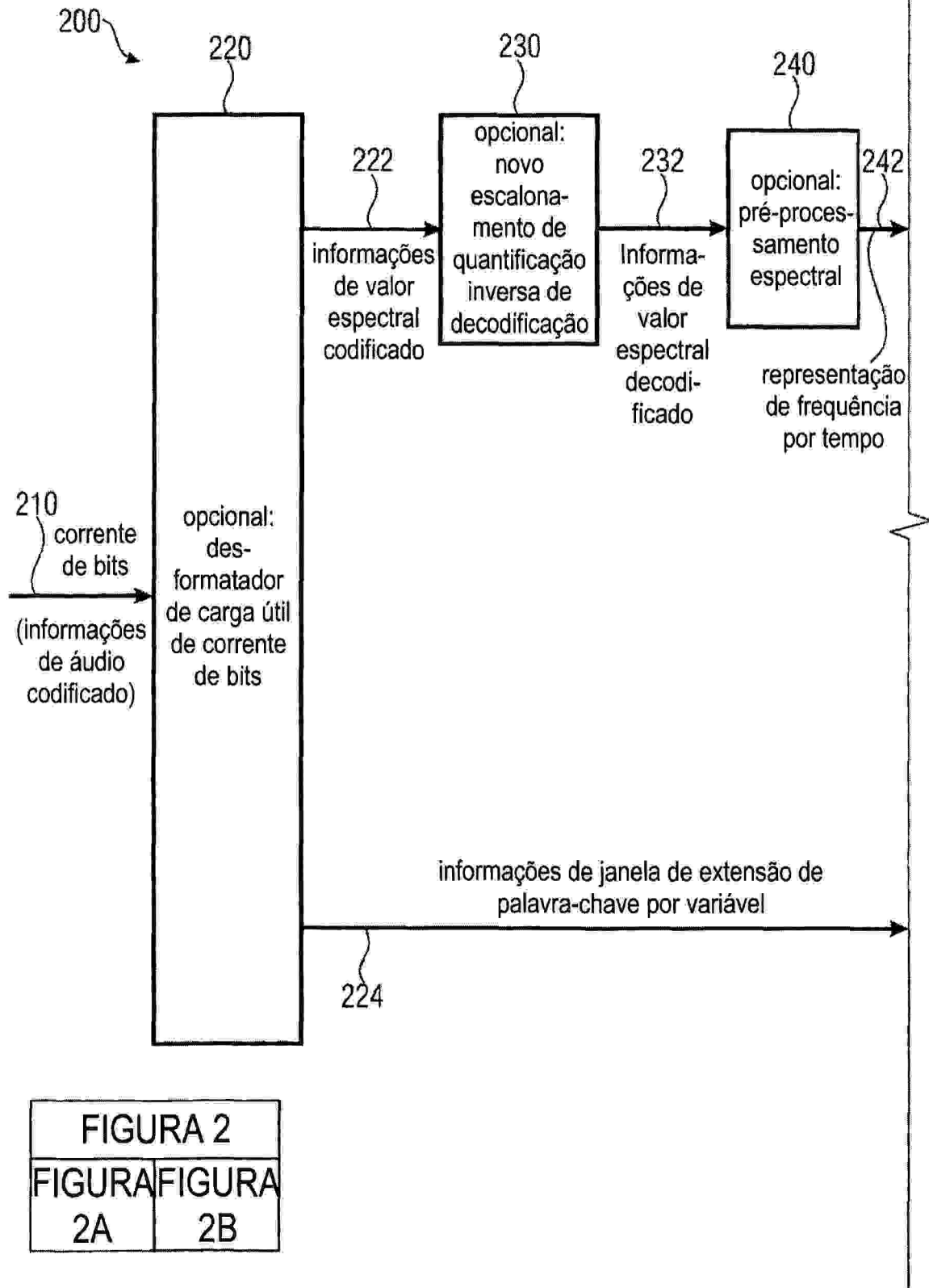


FIGURA 2A

CODIFICAD

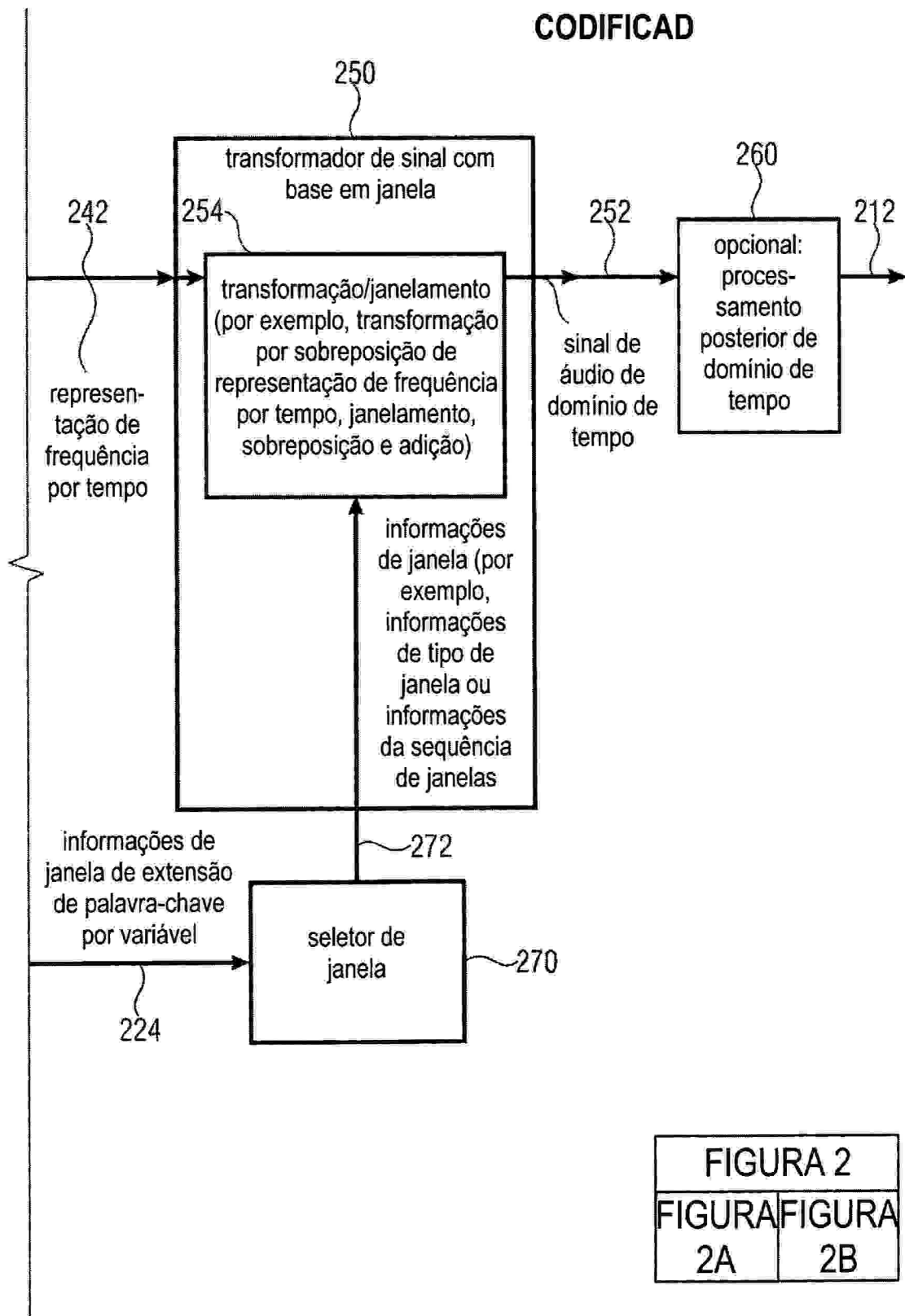


FIGURA 2	
FIGURA 2A	FIGURA 2B

FIGURA 2B

Sequências de Janelas e janelas de transformação

Valor	Janela	nº de coeficientes
0	ONLY_LONG_SEQUENCE =LONG_WINDOW	1024/960
1	LONG_START_SEQUENCE =LONG_START_WINDOW	1024/960
2	EIGHT_SHORT_SEQUENCE =8*SHORT_WINDOW	8*(128/120)
3	LONG_STOP_SEQUENCE =LONG_STOP_WINDOW	1024/960
1	STOP_START_SEQUENCE =STOP_START_WINDOW	1024/960
3	STOP_1152_SEQUENCE =STOP_WINDOW_1152 optional	1152/1080
1	STOP_START_1152_SEQUENCE =STOP_START_WINDOW_1152 optional	1152/1080

FIGURA 3	
FIGURA 3A	FIGURA 3B

FIGURA 3A

Sequências de Janelas e janelas de transformação

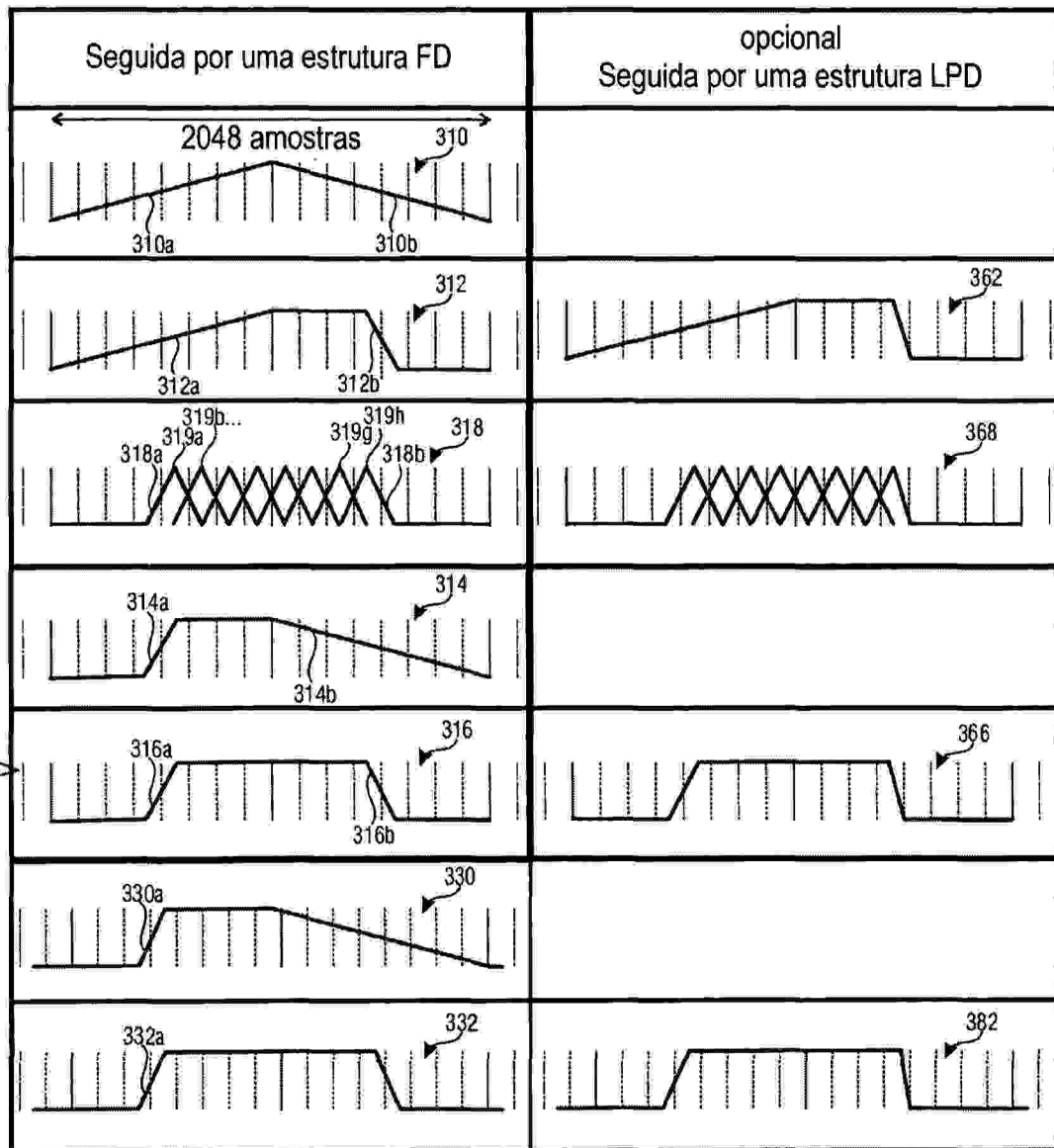


FIGURA 3	
FIGURA 3A	FIGURA 3B

FIGURA 3B

- Sequências de Janelas Permitidas

sequência de janelas para → de ↓	ONLY_LONG_SEQUENCE	LONG_START_SEQUENCE	EIGHT_SHORT_SEQUENCE	LONG_STOP_SEQUENCE	STOP_START_SEQUENCE	LPD_SEQUENCE optional	STOP_1152_SEQUENCE optional	STOP_START_1152_SEQUENCE optional
ONLY_LONG_SEQUENCE	☒	☒						
LONG_START_SEQUENCE			☒	☒	☒	☒		
EIGHT_SHORT_SEQUENCE			☒	☒	☒	☒		
LONG_STOP_SEQUENCE	☒	☒						
STOP_START_SEQUENCE			☒	☒	☒	☒		
LPD_SEQUENCE optional						☒	☒	☒
STOP_1152_SEQUENCE optional	☒	☒						
STOP_START_1152_SEQUENCE optional			☒	☒	☒	☒		

FIGURA 4

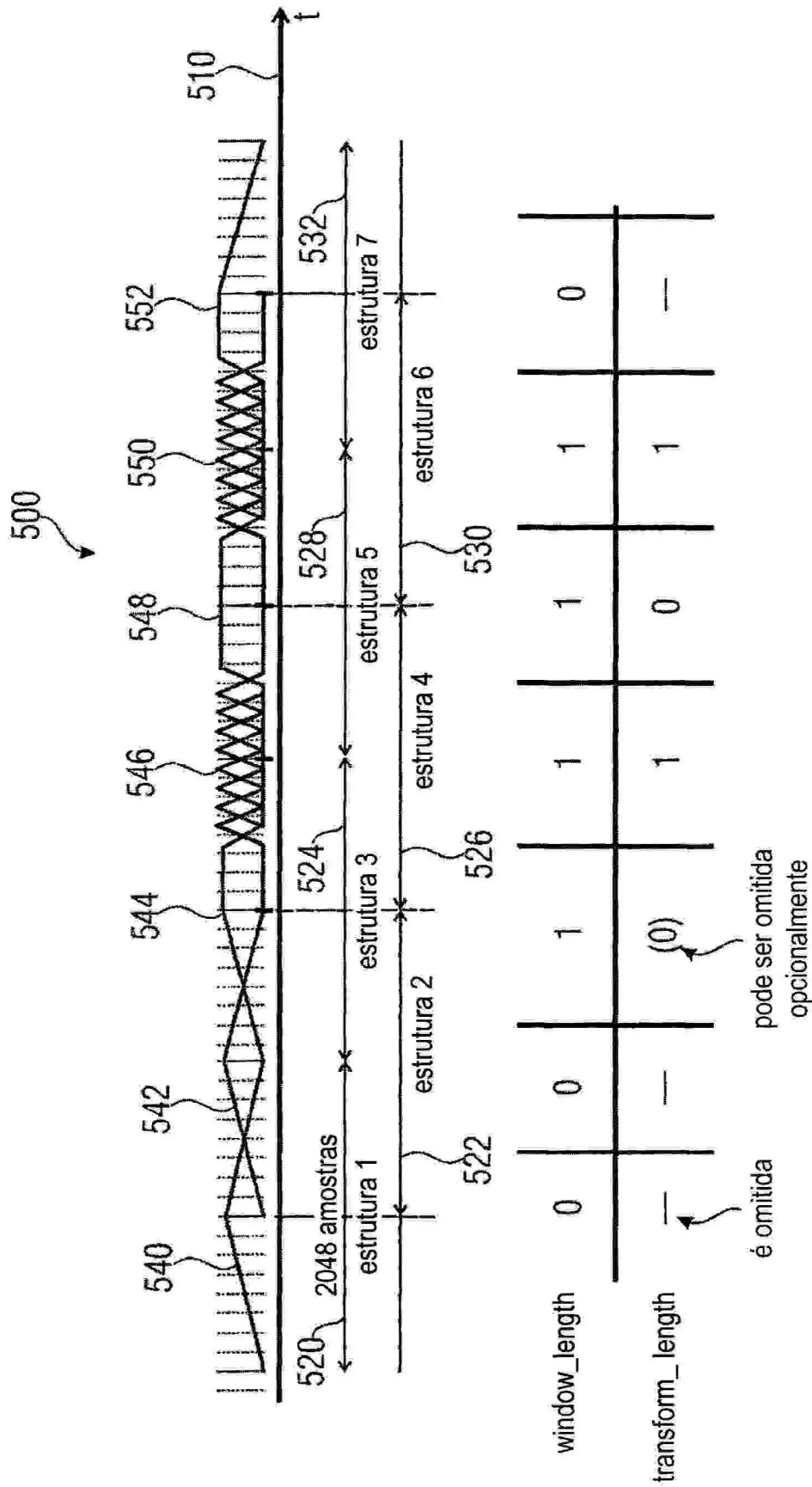


FIGURA 5

opcional	620	624	tabela	630	opcional
		sintaxe proposta		sintaxe atual	
o modo central anterior é LPD	window_length da estrutura anterior	window_length da estrutura atual	transform_length da estrutura atual	core_mode da estrutura seguinte	nome da window_sequence em USAC
	window_sequence da estrutura anterior	window_sequence da estrutura atual			
0	0	0	não transmitida (=0)	0 ou 3	ONLY_LONG_SEQUENCE
0	0	1	0 (ou não transmitida)	0 ou 3	LONG_START_SEQUENCE
0	0	1	0 (ou não transmitida)	0 ou 3	LPD_START_SEQUENCE
0	1	0	não transmitida (=0)	1 ou 2	LONG_STOP_SEQUENCE
0	1	1	0	1 ou 2	STOP_START_SEQUENCE
0	1	1	1	1 ou 2	EIGHT_SHORT_SEQUENCE
1	não transmitida (=1)	0	não transmitida (=0)	LPD	STOP_1152_SEQUENCE
1	não transmitida (=1)	1	0	LPD	STOP_START_1152_SEQUENCE
opcional					opcional

FIGURA 6A

660 window_type da estrutura atual	660 window_length da estrutura atual	662 Transform_length da estrutura atual	664 número de bits de window_type incluídos na corrente de bits
ONLY_LONG_SEQUENCE	0	não transmitida	1
LONG_START_SEQUENCE*	1	0 (ou não transmitida)	2 (opcionalmente:1)
LONG_STOP_SEQUENCE	0	não transmitida	1
STOP_START_SEQUENCE*	1	0	2
EIGHT_SHORT_SEQUENCE*	1	1	2
STOP_1152_SEQUENCE	0	não transmitida	1
STOP_START_1152_SEQUENCE*	1	0 (ou não transmitida)	2 (opcionalmente:1)

opcional

opcional

*pode ser opcionalmente adaptada, caso a estrutura atual for seguida pela estrutura LPD

FIGURA 6B

modo central anterior	window_length da estrutura anterior	window_length da estrutura atual	transform_length da estrutura atual	window_length da estrutura atual
FD	0	0	não avaliada (ou ajustada ao valor padrão)	ONLY_LONG_SEQUENCE
FD	0	1	não avaliada (ou ajustada ao valor padrão)	LONG_START_SEQUENCE*
FD	1	0	não avaliada (ou ajustada ao valor padrão)	LONG_STOP_SEQUENCE
FD	1	1	0	STOP_START_SEQUENCE*
FD	1	1	1	EIGHT_SHORT_SEQUENCE*
LPD	não avaliada (ou ajustada ao valor padrão)	0	não avaliada (ou ajustada ao valor padrão)	STOP_1152_SEQUENCE
LPD	não avaliada (ou ajustada ao valor padrão)	1	não avaliada (ou ajustada ao valor padrão)	STOP_START_1152_SEQUENCE*

*pode ser opcionalmente adaptada, caso a estrutura atual for seguida pela estrutura LPD

opcional

FIGURA 6C

Sintaxe da window_length

window_length	extensão da inclinação da janela direita
0	1024
1	128

FIGURA 7A

Sintaxe da transform_length

Transform_length	tamanho de núcleo MDCT
0	1024/1152
1	128

FIGURA 7B

Transição e sintaxe de nova corrente de bits

window_length	extensão da inclinação da janela direita	transform_length	extensão da transformação
0	1024	0 (não transmitida)	
0	1024	1 (não possível)	
1	128	0	
1	128	1	128

FIGURA 7C

Visão geral sobre todas as combinações de
window_length e transform_length

window_length da estrutura anterior	window_length da estrutura atual	transform_length da estrutura atual	core_mode da estrutura seguinte	valor da window_sequence
0	0	0	—	0
—	1	0	0	1
—	1	0	1	3
—	1	1	—	2
1	0	0	—	3

FIGURA 8

Avaliação de bits salvos

taxa de bits nominal	Nº de bits salvos de transform_length/nº total de estruturas FD	Nº de bits salvos de transform_length/nº total de todas as estruturas	redução de taxa de bits total
todos_mono.wav			
12 kbps	95,67	35,12	4,41
16 kbps	96,54	35,49	5,01
20 kbps	97,18	54,78	8,60
24 kbps	96,86	54,58	9,14
todos_estéreo.wav			
16 kbps	95,89	36,39	5,15
20 kbps	97,23	55,76	8,76
24 kbps	96,70	55,40	9,29
32 kbps	95,65	54,80	10,75
64 kbps	95,15	95,15	22,41

FIGURA 9

```

usac_raw_data-block ()
{
    single_channel_element (); and/or
    channel_pair_element ();
}

```

FIGURA 10A

Sintaxe de single_channel_element

Sintaxe	N° de bits	Mnemônico
<pre> single_channel_element() { core_mode if (core_mode == 1) { lpd_channel_stream(); } else{ fd_channel_stream(); } } </pre>	1	uimbsf

FIGURA 10B

Sintaxe de channel_pair_element()

Sintaxe	Nº de bits	Mnemônico
channel_pair_element()		
{		
core_mode0	1	uimsbf
core_mode1	1	uimsbf
ics_info(); opcional: ics_info comum para dois canais		
if (core_mode0 == 1) {		
lpd_channel_stream();		
}		
else{		
fd_channel_stream();		
}		
if (core_mode1 == 1) {		
lpd_channel_stream();		
}		
else{		
fd_channel_stream();		
}		
}		

FIGURA 10C

Sintaxe de ics_info()

Sintaxe	N° de bits	Mnemônico
ics_info() {		
window_legth:	1	uimsbf
if (window_length != 0) {		
transform_legth:	1	uimsbf
}		
else{		
transform_length=0;		
}		
window_shape:	1	uimsbf
if (window_length != 0 && transform_length != 0) {		
max_sfb;	4	uimsbf
scale_factore_grouping;	7	uimsbf
}		
else{		
max_sfb;	6	uimsbf
}		
}		

opcional

FIGURA 10D

Sintaxe de fd_channel_stream()

Sintaxe	N° de bits	Mnemônico
fd_channel_stream() { global_gain; ics_info (); scale_factor_data (); ac_spectral_data (); }	8	uimsbf

FIGURA 10E

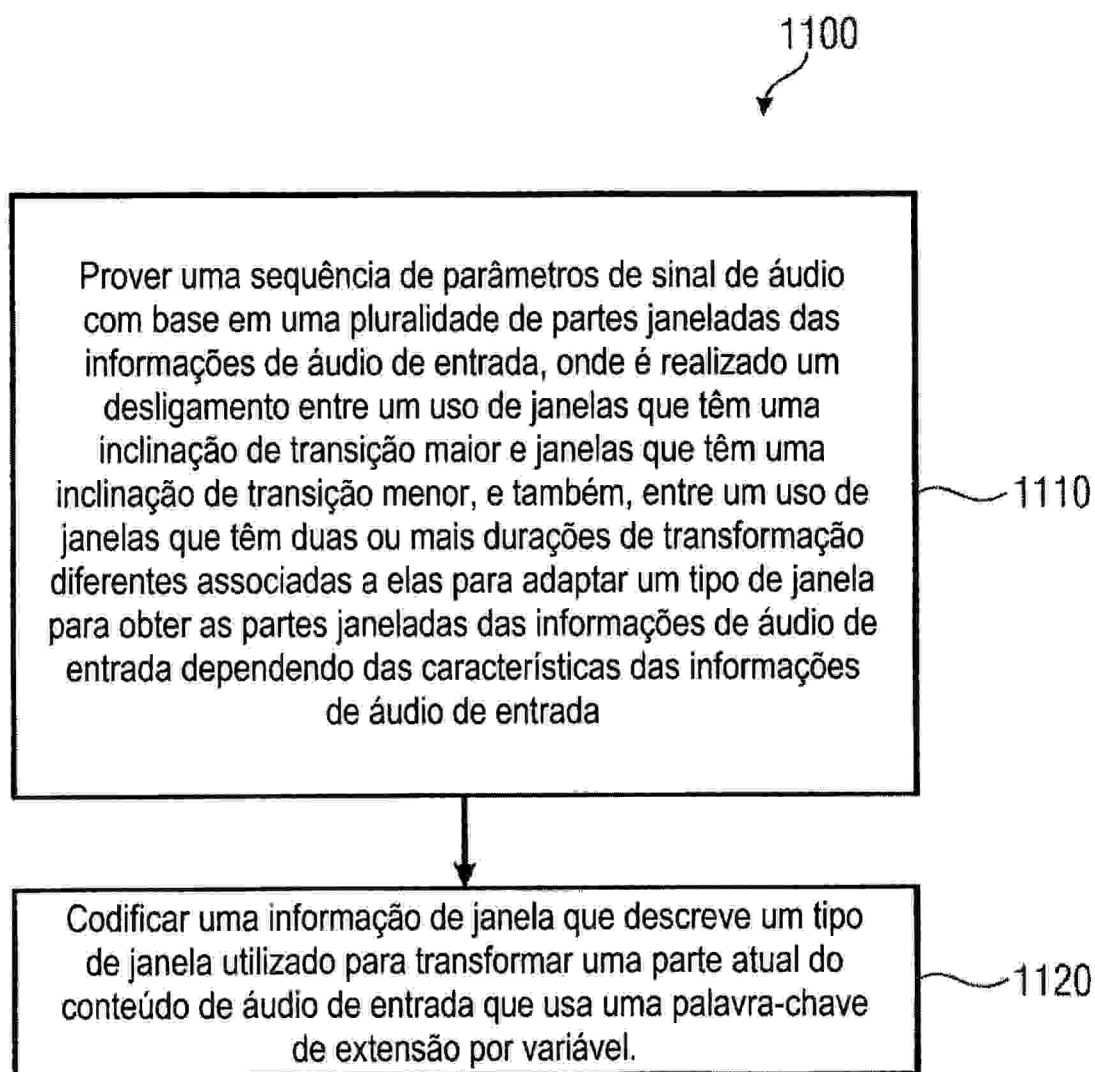


FIGURA 11

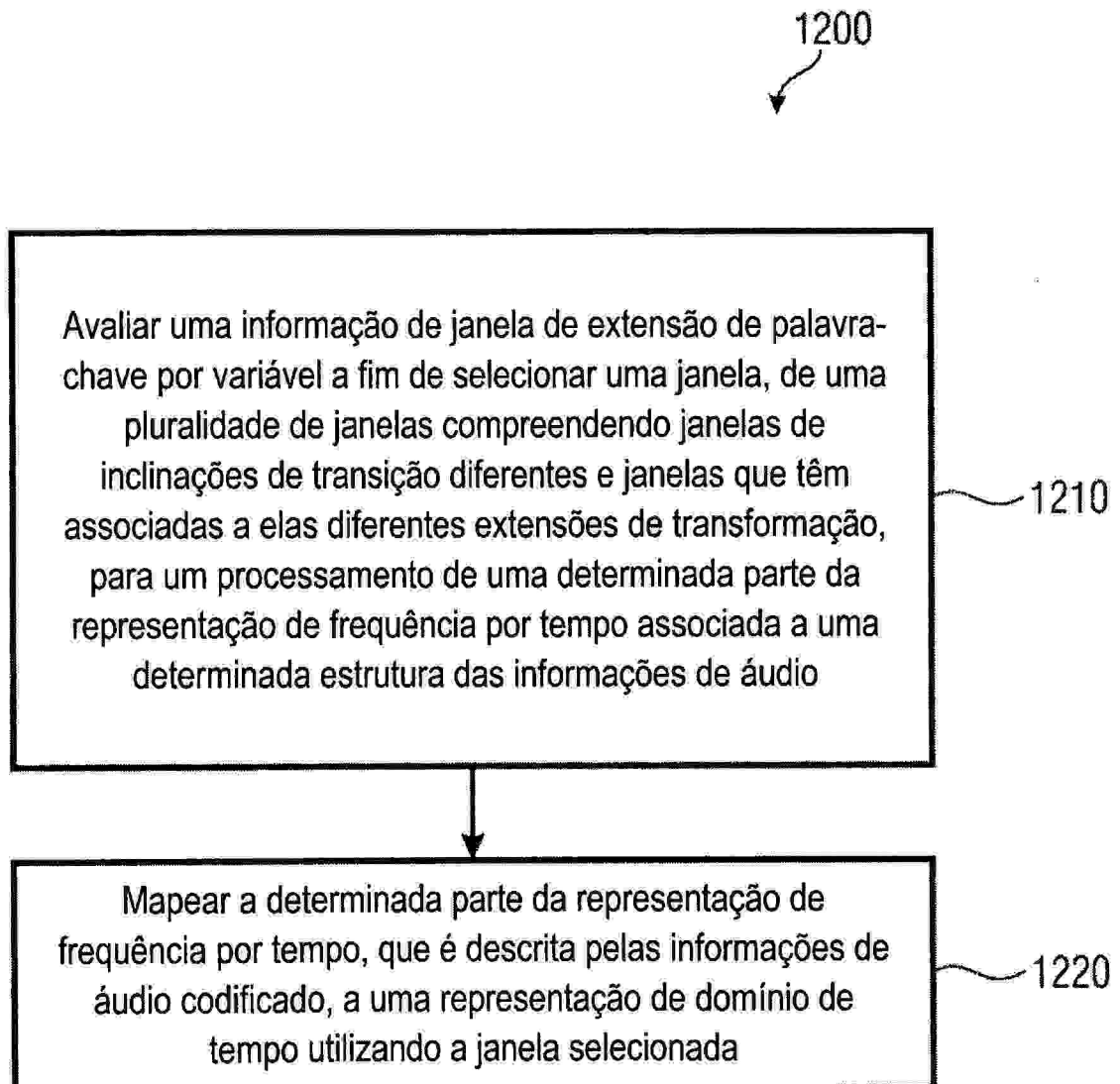


FIGURA 12