



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

(11) BR 112016017589-1 B1



(22) Data do Depósito: 30/01/2015

(45) Data de Concessão: 06/09/2022

(54) Título: INDICAÇÃO DE REUSABILIDADE DE PARÂMETROS DE QUADROS PARA VETORES DE CODIFICAÇÃO

(51) Int.Cl.: G10L 19/038; G10L 19/008; H04R 5/00.

(30) Prioridade Unionista: 01/07/2014 US 62/019,663; 01/08/2014 US 62/032,440; 07/03/2014 US 61/949,591; 07/03/2014 US 61/949,583; 12/01/2015 US 62/102,243; (...).

(73) Titular(es): QUALCOMM INCORPORATED.

(72) Inventor(es): NILS GÜNTHER PETERS; DIPANJAN SEN.

(86) Pedido PCT: PCT US2015013818 de 30/01/2015

(87) Publicação PCT: WO 2015/116952 de 06/08/2015

(85) Data do Início da Fase Nacional: 28/07/2016

(57) Resumo: INDICAÇÃO DE REUSABILIDADE DE PARÂMETROS DE QUADROS PARA VETORES DE CODIFICAÇÃO Em geral, são descritas técnicas para indicar a reusabilidade de parâmetros de quadro para vetores de decodificação. Um dispositivo compreendendo um processador e uma memória pode executar as técnicas. O processador pode ser configurado para obter um fluxo de bits compreendendo um vetor representativo de um eixo espacial ortogonal em um domínio de harmônicos esféricos. O fluxo de dados pode compreender ainda um indicador quanto à reutilização, a partir de um quadro anterior, de pelo menos um elemento de sintaxe indicativo da informação utilizada quando se comprime o vetor. A memória pode ser configurada para armazenar o fluxo de bits.

"INDICAÇÃO DE REUSABILIDADE DE PARÂMETROS DE QUADROS PARA
VETORES DE CODIFICAÇÃO"

[0001] O presente pedido reivindica o benefício dos Pedidos de Patente Provisórios U.S. que seguem:

Pedido de Patente Provisório U.S. No. 61/933.706, depositado em 30 de janeiro de 2014, com o título "COMPRESSION OF DECOMPOSED REPRESENTATIONS OF A SOUND FIELD";

Pedido de Patente Provisório U.S. No. 61/933.714, depositado em 30 de janeiro de 2014, com o título "COMPRESSION OF DECOMPOSED REPRESENTATIONS OF A SOUND FIELD";

Pedido de Patente Provisório U.S. No. 61/933.731, depositado em 30 de janeiro de 2014, intitulado "INDICATING FRAME PARAMETER REUSABILITY FOR DECODING SPATIAL VECTORS";

Pedido de Patente Provisório U.S. No. 61/949.591, depositado 07 de março de 2014, intitulado "IMMEDIATE PLAY-OUT FRAME FOR SPHERICAL HARMONIC COEFFICIENTS";

Pedido de Patente Provisório U.S. No. 61/949.583, depositado em 07 de março de 2014, intitulado "FADE-IN/FADE-OUT OF DECOMPOSED REPRESENTATIONS OF A SOUND FIELD";

Pedido de Patente Provisório U.S. No. 61/994.794, depositado em 16 de maio de 2014, intitulado "CODING V-VECTORS OF A DECOMPOSED HIGHER ORDER AMBISSÔNICO (HO A) AUDIO SIGNAL";

Pedido de Patente Provisório U.S. No. 62/004.147, depositado em 28 de maio de 2014, intitulado "INDICATING FRAME PARAMETER REUSABILITY FOR DECODING SPATIAL VECTORS";

Pedido de Patente Provisório U.S. No. 62/004.067, depositado em 28 de maio de 2014, intitulado "IMMEDIATE PLAY-OUT FRAME FOR SPHERICAL HARMONIC COEFFICIENTS AND

FADE-IN/FADE-OUT OF DECOMPOSED REPRESENTATIONS OF A SOUND FIELD”;

Pedido de Patente Provisório U.S. No. 62/004.128, depositado em 28 de maio de 2014, intitulado “CODING V-VECTORS OF A DECOMPOSED HIGHER ORDER AMBISSÔNICO (HO A) AUDIO SIGNAL”;

Pedido de Patente Provisório U.S. No. 62/019.663, apresentado em 1 de julho de 2014, intitulado “CODING V-VECTORS OF A DECOMPOSED HIGHER ORDER AMBISSÔNICO (HO A) AUDIO SIGNAL”;

Pedido de Patente Provisório U.S. No. 62/027.702, depositado 22 de julho de 2014, intitulado “CODING V-VECTORS OF A DECOMPOSED HIGHER ORDER AMBISSÔNICO (HO A) AUDIO SIGNAL”;

Pedido de Patente Provisório U.S. No. 62/028.282, depositado em 23 de julho de 2014, intitulado “CODING V-VECTORS OF A DECOMPOSED HIGHER ORDER AMBISSÔNICO (HO A) AUDIO SIGNAL”;

Pedido de Patente Provisório U.S. No. 62/029.173, depositado 25 de julho de 2014, intitulado “IMMEDIATE PLAY-OUT FRAME FOR SPHERICAL HARMONIC COEFFICIENTS AND FADE-IN/FADE-OUT OF DECOMPOSED REPRESENTATIONS OF A SOUND FIELD”;

Pedido de Patente Provisório U.S. No. 62/032.440, depositado em 01 de agosto de 2014, intitulado “CODING V-VECTORS OF A DECOMPOSED HIGHER ORDER AMBISSÔNICO (HO A) AUDIO SIGNAL”;

Pedido de Patente Provisório U.S. No. 62/056.248, depositado em 26 de setembro de 2014, intitulado “SWITCHED V- VECTOR QUANTIZATION OF A HIGHER ORDER AMBISSÔNICO (HO A) AUDIO SIGNAL” e

Pedido de Patente Provisório U.S. No. 62/056.286, depositado em 26 de setembro de 2014, intitulado

"PREDICTIVE VECTOR QUANTIZATION OF A DECOMPOSED HIGHER ORDER AMBISSÔNICO (HO A) AUDIO SIGNAL"; e

Pedido de Patente Provisório U.S. No. 62/102.243, depositado em 12 de janeiro de 2015, intitulado "TRANSITIONING OF AMBIENT HIGHER-ORDER AMBISONIC COEFFICIENTS";

cada um dos Pedidos de Patente Provisórios U.S. listados acima é incorporado a título de referência como se aqui mostrado em sua respectiva totalidade.

CAMPO TÉCNICO DA INVENÇÃO

[0002] A presente invenção se refere a dados de áudio e, mais especificamente, à codificação de dados de áudio ambissônico de ordem superior.

ANTECEDENTES DA INVENÇÃO

[0003] Um sinal ambissônico de ordem superior (HOA) (muitas vezes representado por uma pluralidade de coeficientes harmônicos esféricos (SHC) ou outros elementos hierárquicos) é uma representação tridimensional de um campo de som. A representação HOA ou SHC pode representar o campo de som de uma maneira que é independente da geometria do alto-falante local utilizado para a reprodução de um sinal de áudio multicanal renderizado a partir do sinal de SHC. O sinal de SHC também pode facilitar a compatibilidade com versões anteriores uma vez que o sinal de SHC pode ser renderizado nos formatos de multicanal bem conhecidos e altamente adotados, tal como um formato de canal de áudio 5.1 ou um formato de canal de áudio 7.1. A representação de SHC pode, portanto, permitir uma melhor representação de um campo de som que também acomoda compatibilidade com versões anteriores.

SUMÁRIO

[0004] Em geral, são descritas técnicas para codificação de dados de áudio ambissônico de ordem

superior. Dados de áudio ambissônico de ordem superior podem compreender pelo menos um coeficiente harmônico esférico correspondendo a uma função com base em harmônicos esféricos tendo uma ordem maior do que um.

[0005] Em um aspecto, um método de utilização de bit eficiente compreende obtenção de um fluxo de bits que compreende um vetor representativo de um eixo ortogonal espacial em um domínio de harmônicos esféricos. O fluxo de bits compreende ainda um indicador para se reutilizar, de um quadro anterior, pelo menos um elemento de sintaxe indicativo de informação utilizada quando se comprime o vetor.

[0006] Em outro aspecto, um dispositivo configurado para executar a utilização de bit eficiente compreende um ou mais processadores configurados para obter um fluxo de bits compreendendo um vetor representativo de um eixo ortogonal espacial em um domínio de harmônicos esféricos. O fluxo de bits compreende ainda um indicador para se reutilizar, de um quadro anterior, pelo menos um elemento de sintaxe indicativo de informação utilizada quando se comprime o vetor. O dispositivo compreende também uma memória configurada para armazenar o fluxo de bits.

[0007] Em outro aspecto, um dispositivo configurado para executar utilização de bit eficiente compreende meios para obtenção de um fluxo de bit que compreende um vetor representativo de um eixo ortogonal espacial em um domínio de harmônicos esféricos. O fluxo de bits compreende, ainda, um indicador para se reutilizar, de um quadro anterior, pelo menos um elemento de sintaxe indicativo de informação utilizada quando se comprime o vetor. O dispositivo também compreende meios para armazenar o indicador.

[0008] Em outro aspecto, um meio de armazenamento legível por computador não transitório tem nele armazenadas instruções que quando executadas levam a que um ou mais processadores obtenham um fluxo de bits compreendendo um vetor representativo de um eixo espacial ortogonal em um domínio de harmônicos esféricos, em que o fluxo de bits compreende ainda um indicador para se reutilizar, de um quadro anterior, pelo menos um elemento de sintaxe indicativo de informação utilizada quando se comprime o vetor.

[0009] Os detalhes de um ou mais aspectos das técnicas são apresentados nos desenhos acompanhantes e na descrição abaixo. Outras características, objetos e vantagens das técnicas serão aparentes a partir da descrição e desenhos, e a partir das reivindicações.

BREVE DESCRIÇÃO DOS DESENHOS

[0010] A FIG. 1 é um diagrama ilustrando as funções com base em harmônicos esféricos de várias ordens e subordens.

[0011] A FIG. 2 é um diagrama que ilustra um sistema que pode executar vários aspectos das técnicas descritas na presente descrição.

[0012] A FIG. 3 é um diagrama de blocos que ilustra, em mais detalhes, um exemplo do dispositivo de codificação de áudio mostrado no exemplo da FIG. 2 que pode realizar vários aspectos das técnicas descritas na presente descrição.

[0013] A FIG. 4 é um diagrama de blocos que ilustra o dispositivo de decodificação de áudio da FIG. 2 em mais detalhes.

[0014] A FIG. 5A é um fluxograma que ilustra operação exemplificativa de um dispositivo de codificação de áudio na realização de vários aspectos das técnicas de

síntese baseadas em vetores descritas na presente descrição.

[0015] A FIG. 5B é um fluxograma que ilustra operação exemplificativa de um dispositivo de codificação de áudio na realização de vários aspectos das técnicas de codificação descritas na presente descrição.

[0016] A FIG. 6A é um fluxograma que ilustra operação exemplificativa de um dispositivo de decodificação de áudio na realização de vários aspectos das técnicas descritas na presente descrição.

[0017] A FIG. 6B é um fluxograma que ilustra operação exemplificativa de um dispositivo de decodificação de áudio na realização de vários aspectos das técnicas de codificação descritas na presente descrição.

[0018] A FIG. 7 é um diagrama que ilustra, em mais detalhes, quadros do fluxo de bits que podem especificar os componentes espaciais comprimidos.

[0019] A FIG. 8 é um diagrama que ilustra uma parte do fluxo de bits que pode especificar os componentes espaciais comprimidos em mais detalhe.

DESCRIÇÃO DETALHADA

[0020] A evolução do som surround disponibilizou muitos formatos de saída para entretenimento hoje em dia. Exemplos de tais formatos de som surround para consumidor são na sua maioria 'canal' com base em que eles especificam implicitamente alimentações para alto-falantes em determinadas coordenadas geométricas. Os formatos de som surround para consumidor incluem o popular formato 5.1 (que inclui os seis canais que seguem: efeitos frontal esquerdo (FL), frontal direito (FR), centro ou centro frontal, traseiro esquerdo ou surround esquerdo, traseiro direito ou surround direito e de baixa frequência (LFE)), o formato 7.1 crescente, vários formatos que incluem alto-falantes de

coluna tal como o formato e o formato 7.1.4 e o formato 22.2 (por exemplo, para utilização com o padrão de Televisão de Definição Ultra Alta). Formatos não consumidor podem abranger qualquer número de alto-falantes (em geometrias simétricas e não simétricas) muitas vezes denominados 'disposições surround'. Um exemplo de tal disposição inclui 32 alto-falantes posicionados em coordenadas nos ângulos de um icosaedro truncado.

[0021] A entrada para um futuro codificador MPEG é opcionalmente um dos três formatos possíveis: (i) áudio baseado em canal tradicional (como discutido acima), que é feito para ser tocado através de alto-falantes em posições preestabelecidas; (ii) áudio baseado em objeto, que envolve dados de modulação por código de pulso (PCM) discretos para objetos de áudio individuais com metadados associados contendo suas coordenadas de localização (entre outras informações); e (iii) áudio baseado em cena, que envolve representação do campo de som utilizando coeficientes de funções com base em harmônicos esféricos (também chamados "coeficientes harmônicos esféricos" ou SHC, "Ambissônico de Ordem Superior" ou HOA e "coeficientes de HOA"). O futuro codificador MPEG pode ser descrito em mais detalhes em um documento intitulado "Call for Proposals for 3D Audio", da International Organization for Standardization/International Electrotechnical Commission (ISO)/(IEC) JTC1/SC29/WG11/N13411, divulgado em janeiro 2013 em Genebra, Suíça, e está disponível em <http://mpeg.chiariglione.org/sites/default/files/files/standards/parts/docs/wl3411.zip>.

[0022] Há vários formatos baseados em canal de 'som surround' no mercado. Eles variam, por exemplo, do sistema de cinema em casa 5.1 (que tem sido o mais bem sucedido em termos de fazer incursões em salas de estar

além estéreo) ao sistema 22.2 desenvolvido pela NHK (Nippon Hoso Kyokai ou Japan Broadcasting Corporation). Os criadores de conteúdo (por exemplo, os estúdios de Hollywood) gostariam de produzir a trilha sonora para um filme uma vez, e não gastar esforço para fazer seu remix para cada configuração de alto-falante. Recentemente, Organizações de Desenvolvimento de Padrões têm considerado maneiras para fornecer uma codificação em uma corrente de bit padronizada e uma decodificação subsequente que seja adaptável e agnóstica à geometria (e número) do alto-falante e condições acústicas no local da reprodução (que envolvem um renderizador).

[0023] Para fornecer essa flexibilidade para criadores de conteúdo, um conjunto hierárquico de elementos pode ser usado para representar um campo de som. O conjunto hierárquico de elementos pode se referir a um conjunto de elementos em que os elementos são ordenados de tal forma que um conjunto básico de elementos de ordem inferior fornece uma representação completa do campo de som modelado. Como o conjunto é estendido para incluir elementos de ordem superior, a representação se torna mais detalhada, aumentando a resolução.

[0024] Um exemplo de um conjunto hierárquico de elementos é um conjunto de coeficientes harmônicos esféricos (SHC). A expressão a seguir demonstra uma descrição ou representação de um campo de som utilizando SHC:

$$p_i(t, r_r, \theta_r, \varphi_r) = \sum_{\omega=0}^{\infty} \left[4\pi \sum_{n=0}^{\infty} j_n(kr_r) \sum_{m=-n}^n A_n^m(k) Y_n^m(\theta_r, \varphi_r) \right] e^{j\omega t},$$

[0025] A expressão mostra que a pressão p_i em qualquer ponto $\{r_r, \theta_r, \varphi_r\}$ do campo de som, no tempo t ,

pode ser representada de forma única pelo SHC, $A_n^m(K)$. Aqui, $k = \frac{\omega}{c}$, c é a velocidade do som (~ 343 m/s), $\{r_r, \theta_r, \varphi_r\}$ é um ponto de referência (ou ponto de observação), $j_n(\cdot)$ é a função esférica de Bessel de ordem n e $Y_n^m(\theta_r, \varphi_r)$ são as funções com base em harmônicos esféricos de ordem n e subordem m . Pode ser reconhecido que o termo entre parênteses é uma representação no domínio da frequência do sinal (ou seja, $S(\omega, r_r, \theta_r, \varphi_r)$) que pode ser aproximada por várias transformadas tempo-frequência, tais como a transformada discreta de Fourier (DFT), transformada discreta de cosseno (DCT) ou uma transformada wavelet. Outros exemplos de conjuntos hierárquicos incluem conjuntos de coeficientes de transformada wavelet e outros conjuntos de coeficientes de funções de base de multirresolução.

[0026] A FIG. 1 é um diagrama ilustrando funções com base em harmônicos esféricos da ordem zero ($n = 0$) até a quarta ordem ($n = 4$). Como pode ser visto, para cada ordem, há uma expansão de subordens m que são mostradas, mas não observadas explicitamente, no exemplo da FIG. 1 para facilidade de fins de ilustração.

[0027] O SHC $A_n^m(K)$ pode ser ou fisicamente adquirido (por exemplo, gravado) por várias configurações de disposição de microfone ou, alternativamente, eles podem ser derivados de descrições baseadas em canal ou baseadas em objeto do campo de som. O SHC representa áudio baseado em cena, onde o SHC pode ser entrada para um codificador de áudio para obter SHC codificado que pode promover transmissão ou armazenamento mais eficiente. Por exemplo, pode ser usada uma representação de quarta ordem envolvendo coeficientes $(1+4)^2$ (25, e, portanto, de quarta ordem).

[0028] Como mencionado acima, o SHC pode ser derivado de uma gravação de microfone utilizando uma disposição de microfones. Vários exemplos de como SHC pode ser derivado de disposições de microfone estão descritos em Poletti, M., "Three-Dimensional Surround Sound Systems Based on Spherical Harmonics", *J. Audio Eng. Soc*, Vol. 53, N ° 11, 2005 novembro, pp. 1004-1025.

[0029] Para ilustrar como os SHCs podem ser derivados de uma descrição baseada em objetos, considerar a seguinte equação. Os coeficientes $A_n^m(K)$ para o campo de som correspondente a um objeto de áudio individual podem ser expressos como:

$$A_n^m(k) = g(\omega)(-4\pi i k)h_n^{(2)}(kr_s)Y_n^{m*}(\theta_s, \varphi_s),$$

em que i é $\sqrt{-1}$, $h_n^{(2)}(\cdot)$ é a função de Hankel esférica (do segundo tipo) de ordem n e $\{r_s, \theta_s, \varphi_s\}$ é a localização do objeto. Conhecer a energia da fonte de objeto $g(\omega)$ como uma função da frequência (por exemplo, usando técnicas de análise de tempo-frequência, tal como realização de uma transformada rápida de Fourier na corrente de PCM) permite converter cada objeto de PCM e a localização correspondente no SHC $A_n^m(K)$. Além disso, pode ser mostrado (uma vez que acima está uma decomposição linear e ortogonal) que os coeficientes $A_n^m(k)$ para cada objeto são aditivos. Deste modo, uma multiplicidade de objetos de PCM pode ser representada pelos coeficientes $A_n^m(k)$ (por exemplo, como uma soma dos vetores de coeficiente para os objetos individuais). Essencialmente, os coeficientes contêm informações sobre o campo de som (a pressão em função de coordenadas 3D), e o acima representa a transformação de objetos individuais para uma representação do campo de som

geral, na vizinhança do ponto de observação $\{r_r, \theta_r, \varphi_r\}$. As figuras restantes são descritas abaixo no contexto de codificação de áudio com base em objeto e com base em SHC.

[0030] A FIG. 2 é um diagrama que ilustra um sistema 10 que pode executar vários aspectos das técnicas descritas na presente descrição. Como mostrado no exemplo da FIG. 2, o sistema 10 inclui um dispositivo criador de conteúdo 12 e um dispositivo consumidor de conteúdo 14. Embora descritas no contexto do dispositivo criador de conteúdo 12 e do dispositivo consumidor de conteúdo 14, as técnicas podem ser implementadas em qualquer contexto em que SHCs (que podem também ser referidos como coeficientes de HOA) ou qualquer outra representação hierárquica de um campo de som são codificados para formar um fluxo de bit representativo dos dados de áudio. Além disso, o dispositivo criador de conteúdo 12 pode representar qualquer forma de dispositivo de computação capaz de implementar as técnicas descritas na presente descrição, incluindo um terminal móvel (ou celular), um computador tablet, um telefone inteligente ou um computador de mesa para mencionar alguns exemplos. Da mesma forma, o dispositivo consumidor de conteúdo 14 pode representar qualquer forma de dispositivo de computação capaz de implementar as técnicas descritas na presente descrição, incluindo um terminal móvel (ou celular), um computador tablet, um telefone inteligente, um conversor ou um computador de mesa para mencionar alguns exemplos.

[0031] O dispositivo criador do conteúdo 12 pode ser operado por um estúdio de cinema ou outra entidade que pode gerar conteúdo de áudio multicanal para consumo por operadores de um consumidor de conteúdo, tal como o dispositivo de consumidor de conteúdo 14. Em alguns exemplos, o dispositivo criador de conteúdo 12 pode ser

operado por um usuário individual que gostaria de comprimir coeficientes de HOA 11. Muitas vezes, o criador de conteúdo gera conteúdo de áudio em conjunto com conteúdo de vídeo. O dispositivo consumidor de conteúdo 14 pode ser operado por um indivíduo. O dispositivo consumidor de conteúdo 14 pode incluir um sistema de reprodução de áudio 16, o que pode se referir a qualquer tipo de sistema de reprodução de áudio capaz de renderizar SHC para reprodução como conteúdo de áudio multicanal.

[0032] O dispositivo criador de conteúdo 12 inclui um sistema de edição de áudio 18. O dispositivo criador de conteúdo 12 obtém gravações ao vivo 7 em vários formatos (incluindo diretamente como coeficientes de HOA) e objetos de áudio 9, que o dispositivo criador de conteúdo 12 pode editar usando sistema de edição de áudio 18. O criador de conteúdo pode, durante o processo de edição, renderizar coeficientes de HOA 11 a partir de objetos de áudio 9, ouvir as alimentações do alto-falante renderizadas em uma tentativa de identificar vários aspectos do campo de som que exigem edição posterior. O dispositivo criador de conteúdo 12 pode então editar coeficientes de HOA 11 (potencialmente indiretamente, através de manipulação dos diferentes dos objetos áudio 9 dos quais os coeficientes de HOA de origem podem ser derivados da maneira descrita acima). O dispositivo criador de conteúdo 12 pode empregar o sistema de edição de áudio 18 para gerar os coeficientes de HOA 11. O sistema de edição de áudio 18 representa qualquer sistema capaz de editar dados de áudio e emitir os dados de áudio, como um ou mais coeficientes harmônicos esféricos de origem.

[0033] Quando o processo de edição está completo, o dispositivo criador de conteúdo 12 pode gerar um fluxo de bits 21 com base nos coeficientes de HOA 11. Isto é, o

dispositivo criador de conteúdo 12 inclui um dispositivo de codificação de áudio 20 que representa um dispositivo configurado para codificar ou de outro modo comprimir coeficientes de HOA 11 de acordo com vários aspectos das técnicas descritas na presente descrição para gerar o fluxo de bits 21. O dispositivo de codificação de áudio 20 pode gerar o fluxo de bits para transmissão 21, como um exemplo, através de um canal de transmissão, que pode ser um canal com fio ou sem fios, um dispositivo de armazenamento de dados ou similar. O fluxo de bits 21 pode representar uma versão codificada do coeficiente de HOA 11 e pode incluir um fluxo de bit primário e um outro de fluxo de bits lateral, que pode ser referido como a informação de canal lateral.

[0034] Embora descrito em mais detalhes abaixo, o dispositivo de codificação de áudio 20 pode ser configurado para codificar os coeficientes de HOA 11 com base em uma síntese à base de vetor ou uma síntese de base direcional. Para determinar se é necessário executar a metodologia de decomposição à base de vetor ou uma metodologia de decomposição de base direcional, o dispositivo de codificação de áudio 20 pode determinar, com base pelo menos em parte nos coeficientes de HOA 11, se os coeficientes de HOA 11 foram gerados por meio de uma gravação natural de um campo de som (por exemplo, gravação ao vivo 7) ou produzidos artificialmente (isto é, sinteticamente) a partir de, como um exemplo, objetos de áudio 9, tal como um objeto de PCM. Quando os coeficientes de HOA 11 foram gerados a partir dos objetos de áudio 9, o dispositivo de codificação de áudio 20 podem codificar os coeficientes de HOA 11 utilizando a metodologia de decomposição com base direcional. Quando os coeficientes de HOA 11 foram capturados ao vivo usando, por exemplo, um

Eigenmike, o dispositivo de codificação de áudio 20 pode codificar os coeficientes de HOA 11 com base na metodologia de decomposição baseada em vetor. A distinção acima representa um exemplo de onde metodologia de decomposição baseada em vetor ou de base direcional pode ser implantada. Pode haver outros casos em que uma ou ambas podem ser úteis para gravações naturais, conteúdo gerado artificialmente ou uma mistura dos dois (conteúdo híbrido). Além disso, também é possível utilizar simultaneamente as duas metodologias para a codificação de um único espaço de tempo de coeficientes de HOA.

[0035] Supondo para fins de ilustração que o dispositivo de codificação de áudio 20 determina que os coeficientes de HOA 11 foram capturados ao vivo ou de outro modo representam gravações ao vivo, tal como a gravação ao vivo 7, o dispositivo de codificação de áudio 20 pode ser configurado para codificar os coeficientes de HOA 11 usando uma metodologia de decomposição baseada em vetor envolvendo aplicação de uma transformada inversível linear (LIT). Um exemplo da transformada inversível linear é referido como uma "decomposição de valor singular" (ou "DVS"). Neste exemplo, o dispositivo de codificação de áudio 20 pode aplicar SVD aos coeficientes de HOA 11 para determinar uma versão decomposta dos coeficientes de HOA 11. O dispositivo de codificação de áudio 20 pode então analisar a versão decomposta dos coeficientes de HOA 11 para identificar vários parâmetros, que podem facilitar a reordenação da versão decomposta dos coeficientes de HOA 11. O dispositivo de codificação de áudio 20 pode então reordenar a versão decomposta dos coeficientes de HOA 11 com base nos parâmetros identificados, onde tal reordenação, como descrito em mais detalhes abaixo, pode melhorar a eficiência da codificação dado que a transformação pode

reordenar os coeficientes de HOA entre os quadros dos coeficientes de HOA (em que um quadro pode incluir amostras M dos coeficientes de HOA 11 e M é, em alguns exemplos, definido para 1024). Depois de reordenar a versão decomposta dos coeficientes de HOA 11, o dispositivo de codificação de áudio 20 pode selecionar a versão decomposta dos coeficientes de HOA 11 representativos dos componentes de primeiro plano (ou, em outras palavras, distintos, predominantes ou salientes) do campo de som. O dispositivo de codificação de áudio 20 pode especificar a versão decomposta dos coeficientes de HOA 11 representativos dos componentes de primeiro plano como um objeto de áudio e informação direcional associada.

[0036] O dispositivo de codificação de áudio 20 pode também executar uma análise de campo de som com respeito aos coeficientes de HOA 11 em ordem, pelo menos em parte, para identificar os coeficientes de HOA 11 representativos de um ou mais componentes de base (ou, em outras palavras, ambiente) do campo de som. O dispositivo de codificação de áudio 20 pode executar compensação de energia com relação aos componentes de base dado que, em alguns exemplos, os componentes de base podem incluir apenas um subconjunto de uma dada amostra dos coeficientes de HOA 11 (por exemplo, tais como os coeficientes de HOA 11 correspondendo a zero e funções de base esférica de primeira ordem e não os coeficientes de HOA 11 correspondendo às segundas funções de base esféricas ou de ordem superior). Quando a redução de ordem é realizada, em outras palavras, o dispositivo de codificação de áudio 20 pode aumentar (por exemplo, adicionar/subtrair energia para /a partir de) os coeficientes de HOA de base restantes dos coeficientes de HOA 11 para compensar a variação de energia total que resulta da realização da redução de ordem.

[0037] O dispositivo de codificação de áudio 20 pode, em seguida, executar uma forma de codificação psicoacústica (tais como surround MPEG, MPEG-AAC, MPEG-USAC ou outras formas conhecidas de codificação psicoacústica) com respeito a cada um dos coeficientes de HOA 11 representativos de componentes de base e cada um dos objetos de áudio de primeiro plano. O dispositivo de codificação de áudio 20 pode executar uma forma de interpolação no que diz respeito à informação direcional de primeiro plano e, em seguida, executar uma redução de ordem no que diz respeito à informação direcional de primeiro plano interpolada para gerar informação direcional de primeiro plano de ordem reduzida. O dispositivo de codificação de áudio 20 pode ainda executar, em alguns exemplos, uma quantização no que diz respeito à informação direcional de primeiro plano de ordem reduzida, emitindo informação direcional de primeiro plano codificada. Em alguns casos, a quantização pode compreender uma quantização escalar/entropia. O dispositivo de codificação de áudio 20 pode, em seguida, formar o fluxo de bits 21 para incluir os componentes de base codificados, os objetos de áudio de primeiro plano codificados e a informação direcional quantizada. O dispositivo de codificação de áudio 20 pode então transmitir ou de outro modo emitir a corrente de bits 20 para o dispositivo consumidor de conteúdo 14.

[0038] Enquanto mostrado na FIG. 2 como sendo diretamente transmitido para o dispositivo de consumidor de conteúdo 14, o dispositivo criador de conteúdo 12 pode emitir o fluxo de bits 21 para um dispositivo intermediário posicionado entre o dispositivo criador de conteúdo 12 e o dispositivo consumidor de conteúdo 14. O dispositivo intermediário pode armazenar o fluxo de bits 21 para

entrega mais tarde ao dispositivo consumidor de conteúdo 14, que pode solicitar o fluxo de bits. O dispositivo intermediário pode compreender um servidor de arquivos, servidor de web, um computador de mesa, um computador portátil, um computador tablet, um telefone celular, um telefone inteligente ou qualquer outro dispositivo capaz de armazenar o fluxo de bits 21 para posterior recuperação por um decodificador de áudio. O dispositivo intermediário pode residir em uma rede de entrega de conteúdo capaz de fazer transmissão do fluxo de bits 21 (e possivelmente em conjunto com transmissão de um fluxo de bits de dados de vídeo correspondente) para os assinantes, como o dispositivo consumidor de conteúdo 14, solicitando o fluxo de bits 21.

[0039] Alternativamente, o dispositivo criador de conteúdo 12 pode armazenar o fluxo de bits 21 em um meio de armazenamento, como um CD, um disco de vídeo digital, um disco de vídeo de alta definição ou outros meios de armazenamento, a maioria dos quais é capaz de ser lido por um computador e, portanto, podem ser referidos como meios de armazenamento legíveis por computador ou mídia de armazenamento legível por computador não transitória. Neste contexto, o canal de transmissão pode se referir aos canais pelos quais conteúdo armazenado nos meios é transmitido (e podem incluir lojas de varejo e outro mecanismo de entrega à base de loja). Em qualquer caso, as técnicas da presente descrição não devem, portanto, ser limitadas a este respeito ao exemplo da FIG. 2.

[0040] Como mostrado adicionalmente no exemplo da FIG. 2, o dispositivo de consumidor de conteúdo 14 inclui o sistema de reprodução de áudio 16. O sistema de reprodução de áudio 16 pode representar qualquer sistema de reprodução de áudio capaz de reproduzir dados de áudio multicanal. O

sistema de reprodução de áudio 16 pode incluir vários renderizadores diferentes 22. Os renderizadores 22 podem cada uma prover uma forma diferente de renderização, em que as diferentes formas de renderização podem incluir uma ou mais das várias formas de realização de panning de amplitude com base em vetor (VBAP) e/ou uma ou mais das várias formas de realizar síntese de campo de som. Tal como aqui utilizado, "A e/ou B" significa "A ou B" ou ambos "A e B".

[0041] O sistema de reprodução de áudio 16 pode ainda incluir um dispositivo de decodificação de áudio 24. O dispositivo de decodificação de áudio 24 pode representar um dispositivo configurado para decodificar coeficientes de HOA 11' a partir do fluxo de bits 21, onde os coeficientes de HOA 11' podem ser semelhantes aos coeficientes de HOA 11, mas diferem devido a operações com perdas (por exemplo, de quantização) e/ou transmissão através do canal de transmissão. Ou seja, o dispositivo de decodificação de áudio 24 pode desquantizar a informação direcional de primeiro plano especificada no fluxo de bits 21, ao mesmo tempo que executa decodificação psicoacústica no que diz respeito aos objetos de áudio em primeiro plano especificados no fluxo de bits 21 e os coeficientes de HOA codificados representativos de componentes base. O dispositivo de decodificação de áudio 24 pode executar ainda a interpolação no que diz respeito à informação direcional de primeiro plano descodificada e, em seguida, determinar os coeficientes de HOA representativos dos componentes de primeiro plano com base nos objetos de áudio decodificados de primeiro plano e a informação direcional de primeiro plano interpolada. O dispositivo de decodificação de áudio 24 pode então determinar os coeficientes de HOA 11' com base nos coeficientes de HOA

determinados representativos dos componentes de primeiro plano e nos coeficientes de HOA decodificados representativos dos componentes de base.

[0042] O sistema de reprodução de áudio 16 pode, depois de decodificar o fluxo de bits 21 para obter os coeficientes de HOA 11', renderizar os coeficientes de HOA 11' para alimentações de alto-falante 25. As alimentações de alto-falante 25 podem se dirigir a um ou mais alto-falantes (que não são mostradas no exemplo da Fig. 2 para facilidade de fins de ilustração).

[0043] Para selecionar o renderizador apropriado ou, em alguns casos, gerar um renderizador apropriado, o sistema de reprodução de áudio 16 pode obter informações de alto-falantes 13 indicativas de vários alto-falantes e/ou geometria espacial dos vários alto-falantes. Em alguns casos, o sistema de reprodução de áudio 16 pode obter a informação de alto-falante 13 usando um microfone de referência e se dirigindo aos alto-falantes de tal forma a determinar dinamicamente a informação de alto-falante 13. Em outros casos ou em conjunto com a determinação dinâmica das informações de alto-falante 13, o sistema de reprodução de áudio 16 pode solicitar que um usuário faça a interface com o sistema de reprodução de áudio 16 e insira as informações do alto-falante 13.

[0044] O sistema de reprodução de áudio 16 pode então selecionar um dos renderizadores de áudio 22 com base nas informações de alto-falante 13. Em alguns casos, o sistema de reprodução de áudio 16 pode, quando nenhum dos de renderizadores de áudio 22 está dentro de alguma medida de similaridade limiar (geometria ampla de alto-falante) para aquela especificada nas informações de alto-falante 13, gerar o um dos renderizadores de áudio 22 com base na informação do alto-falante 13. O sistema de reprodução de

áudio 16 pode, em alguns casos, gerar um dos renderizadores de áudio 22 com base na informação 13 do alto-falante sem primeiro tentar selecionar um já existente dos renderizadores de áudio 22.

[0045] A FIG. 3 é um diagrama de blocos que ilustra, em mais detalhes, um exemplo do dispositivo de codificação de áudio 20 mostrado no exemplo da FIG. 2 que pode realizar vários aspectos das técnicas descritas na presente descrição. O dispositivo de codificação de áudio 20 inclui uma unidade de análise de conteúdo 26, uma unidade de decomposição baseada em vetor 27 e uma unidade de decomposição de base direcional 28. Embora descrito rapidamente a seguir, mais informações sobre o dispositivo de codificação de áudio 20 e os vários aspectos de compressão ou de outro modo codificação de coeficientes de HOA estão disponíveis na Publicação do Pedido de Patente Internacional nº WO 2014/194099, intitulada "INTERPOLATION FOR DECOMPOSED REPRESENTATIONS OF A SOUND FIELD", depositada em 29 de maio de 2014.

[0046] A unidade de análise de conteúdo 26 representa uma unidade configurada para analisar o conteúdo dos coeficientes de HOA 11 para identificar se os coeficientes de HOA 11 representam o conteúdo gerado a partir de uma gravação ao vivo ou de um objeto de áudio. A unidade de análise de conteúdo 26 pode determinar se os coeficientes de HOA 11 foram gerados a partir de uma gravação de um campo de som real ou de um objeto de áudio artificial. Em alguns casos, quando os coeficientes de HOA 11 enquadrados foram gerados a partir de uma gravação, a unidade de análise de conteúdo 26 passa os coeficientes de HOA 11 para a unidade de decomposição baseada em vetores 27. Em alguns casos, quando os coeficientes de HOA 11 enquadrados foram gerados a partir de um objeto de áudio

sintético, a unidade de análise de conteúdo 26 passa os coeficientes de HOA 11 para a unidade de síntese de base direcional 28. A unidade de síntese de base direcional 28 pode representar uma unidade configurada para realizar uma síntese de base direcional de coeficientes de HOA 11 para gerar um fluxo de bits de base direcional.

[0047] Como mostrado no exemplo da FIG. 3, a unidade de decomposição baseada em vetores 27 pode incluir uma unidade de transformada inversível linear (LIT) 30, uma unidade de cálculo de parâmetros 32, uma unidade de reordenação 34, a unidade de seleção de primeiro plano 36, uma unidade de compensação de energia 38, uma unidade codificadora psicoacústica de áudio 40, uma unidade de geração de corrente de bits 42, uma unidade de análise de campo de som 44, uma unidade de redução de coeficiente 46, uma unidade de seleção de base (BG) 48, uma unidade de interpolação espaço-temporal 50 e uma unidade de quantização 52.

[0048] A unidade de transformada inversível linear (LIT) 30 recebe os coeficientes de HOA 11 na forma de canais de HOA, cada canal representativo de um bloco ou quadro de um coeficiente associado a uma determinada ordem, subordem das funções de base esféricas (o que pode ser chamado HOA[k], onde k pode indicar o quadro ou bloco atual de amostras). A matriz dos coeficientes de HOA 11 pode ter dimensões $D: M \times (N+1)^2$.

[0049] Ou seja, a unidade 30 de LIT pode representar uma unidade configurada para executar uma forma de análise referida como decomposição do valor singular. Embora descrito em relação a SVD, as técnicas descritas na presente descrição podem ser realizadas em relação a qualquer transformação ou decomposição similar que proveja conjuntos de saída de energia comprimida, linearmente não

relacionada. Além disso, referência a "conjuntos" na presente descrição destina-se geralmente a se referir a conjuntos diferentes de zero a não ser que especificamente indicado em contrário e não pretende se referir à definição matemática clássica de conjuntos que inclui o chamado "conjunto vazio".

[0050] Uma transformação alternativa pode compreender uma análise do componente principal, que é muitas vezes referida como "PCA." PCA se refere a um procedimento matemático que utiliza uma transformação ortogonal para converter um conjunto de observações de variáveis possivelmente correlacionadas em um conjunto de variáveis linearmente não correlacionadas referidas como componentes principais. Variáveis linearmente não correlacionadas representam variáveis que não têm uma relação linear estatística (ou dependência) uma com a outra. Os principais componentes podem ser descritos como tendo um grau pequeno de correlação estatística um com o outro. Em qualquer caso, o número dos chamados componentes principais é inferior ou igual ao número de variáveis originais. Em alguns exemplos, a transformação é definida de tal maneira que o primeiro componente principal tem a maior variância possível (ou, em outras palavras, é responsável pelo máximo possível da variabilidade nos dados), e cada componente sucessor por sua vez tem a maior variância possível sob a restrição que o componente sucessivo seja ortogonal (que pode ser reescrito como não correlacionado com) aos componentes anteriores. PCA pode executar uma forma de redução de ordem, o que em termos de coeficientes de HOA 11 pode resultar na compressão dos coeficientes de HOA 11. Dependendo do contexto, PCA pode ser referida por um número de nomes diferentes, tais como transformada de Karhunen-oeve discreta, a transformada de

Hotelling, decomposição ortogonal adequada (POD) e decomposição de autovalor (EVD), para citar alguns exemplos. Propriedades de tais operações que são favoráveis ao objetivo subjacente de compressão de dados de áudio são 'compressão de energia' e 'descorrelação' dos dados de áudio multicanal.

[0051] Em qualquer caso, supondo que a unidade de LIT 30 execute uma decomposição de valor singular (que, mais uma vez, pode ser referida como "SVD") para fins de exemplo, a unidade de LIT 30 pode transformar os coeficientes de HOA 11 em dois ou mais conjuntos de coeficientes de HOA transformados. Os "conjuntos" de coeficientes de HOA transformados podem incluir vetores de coeficientes de HOA transformados. No exemplo da FIG. 3, a unidade de LIT 30 pode executar SVD com respeito aos coeficientes de HOA 11 para gerar uma chamada matriz V, uma matriz S e uma matriz L. SVD, em álgebra linear, pode representar uma fatoração de uma matriz real ou complexa X y-em-z (onde X pode representar dados de áudio multicanal, tais como coeficientes de HOA 11) no seguinte formato:

$$X=USV^*$$

U pode representar uma matriz unitária real ou complexa y-em-y, em que as colunas y de U são conhecidas como os vetores singulares da esquerda de dados de áudio de canais múltiplos. S pode representar uma matriz diagonal rectangular y-em-Z com números reais não-negativos na diagonal, em que os valores de S diagonais são conhecidos como os valores singulares de dados de áudio de canais múltiplos. V* (que pode indicar uma transposição de conjugado de V) pode representar uma matriz unitária real ou complexa z-em-z, em que as colunas z de V* são conhecidas como os vetores singulares da direita dos dados de áudio de canais múltiplos.

[0052] Embora descritas na presente descrição como sendo aplicadas a dados de áudio de canais múltiplos que compreendem coeficientes de HOA 11, as técnicas podem ser aplicadas a qualquer forma de dados de áudio de canais múltiplos. Desta forma, o dispositivo de codificação de áudio 20 pode realizar uma decomposição de valor singular com respeito aos dados de áudio de multicanal representativos de pelo menos uma parte do campo de som para gerar uma matriz U representativa de vetores singulares da esquerda dos dados de áudio de canais múltiplos, uma matriz S representativa de valores singulares dos dados de áudio de canais múltiplos e uma matriz V representativa de vetores singulares da direita dos dados de áudio de canais múltiplos, e representando os dados de áudio de canais múltiplos, como uma função de pelo menos uma parte de uma ou mais da matriz L , da matriz S e da matriz V .

[0053] Em alguns exemplos, a matriz V^* na expressão matemática SVD acima referida é indicada como a transposição conjugada da matriz V de modo a refletir que SVD pode ser aplicado a matrizes compreendendo números complexos. Quando aplicada a matrizes compreendendo apenas números reais, o conjugado complexo da matriz V (ou, em outras palavras, a matriz V^*) pode ser considerado a transposição da matriz V . Abaixo é suposto, para facilidade de propósitos de ilustração, que os coeficientes de HOA 11 compreendem números reais com o resultado que a matriz V é exibida através SVD ao invés da matriz V^* . Além disso, enquanto indicada como a matriz V na presente descrição, referência à matriz V deve ser entendida se referir à transposição da matriz V onde apropriado. Enquanto supondo ser a matriz V , as técnicas podem ser aplicadas de um modo semelhante a coeficientes de HOA 11 tendo coeficientes

complexos, onde a saída do SVD é a matriz V^* . Desta maneira, as técnicas não devem ser limitadas a este respeito apenas à provisão de aplicação de SVD para gerar uma matriz de V , mas podem incluir a aplicação de SVD a coeficientes de HOA 11 tendo componentes complexos para gerar uma matriz V^* .

[0054] Em qualquer caso, a unidade de LIT 30 pode executar uma forma em blocos de SVD com respeito a cada bloco (que pode se referir a um quadro) de dados de áudio ambissônico de ordem superior (HOA) (onde os dados de áudio ambissônico incluem blocos ou amostras dos coeficientes de HOA 11 ou qualquer outra forma de dados de áudio multicanal). Como mencionado acima, uma variável M pode ser utilizada para indicar o comprimento de um quadro de áudio em amostras. Por exemplo, quando um quadro de áudio inclui 1024 amostras de áudio, M é igual a 1024. Embora descritas com respeito ao valor típico de M , as técnicas da descrição não devem ser limitadas ao valor típico para M . A unidade de LIT 30 pode, por conseguinte, realizar um SVD em blocos no que diz respeito a um bloco de coeficientes de HOA 11 possuindo coeficientes de HOA H -em- $(N+1)^2$, onde N , mais uma vez, indica a ordem dos dados de áudio DE HOA. A unidade de LIT 30 pode gerar, por meio da execução do SVD, uma matriz V , uma matriz S e uma matriz U , onde cada uma das matrizes pode representar as respectivas matrizes V , S e U descritas acima. Desta forma, a unidade de transformada inversível linear 30 pode executar SVD com respeito aos coeficientes de HOA 11 para emitir vetores $US[k]$ 33 (que podem representar uma versão combinada dos vetores S e dos vetores U) com as dimensões de $D: M \times (N+1)^2$, e vetores $V[K]$ 35 com dimensões $D: (N+1)^2 \times (N+1)^2$. Elementos de vetor individuais na matriz $US[k]$ podem também

ser chamados $XPS(k)$, enquanto vetores individuais da matriz $V[k]$ podem também ser chamados $v(k)$.

[0055] Uma análise das matrizes L , S e V pode revelar que as matrizes carregam ou representam características espaciais e temporais do campo de som subjacente representado acima por X . Cada um dos vetores N em U (de comprimento de amostras M) pode representar sinais de áudio separados normalizados como uma função de tempo (para o período de tempo que representam as amostras M), que são ortogonais entre si e que foram dissociados de quaisquer características espaciais (que podem também ser referidas como informação direcional). As características espaciais, que representam largura de forma e posição espaciais (r , θ , ϕ), podem ao invés ser representadas por vetores individuais, vetores I -ésimo, $v^{(i)}(k)$, na matriz V (cada uma de comprimento $(N+1)^2$). Os elementos individuais de cada um dos vetores $v^{(i)}(k)$ podem representar um coeficiente de HOA descrevendo a forma e a direção do campo de som para um objeto de áudio associado. Ambos os vetores na matriz U e na matriz V são normalizados de modo que as suas energias valor quadrático médio são iguais à unidade. A energia dos sinais de áudio em U é, portanto, representada pelos elementos diagonais em S . Multiplicação de U e S para formar $US[k]$ (com elementos de vetor individual $X_{PS}(k)$), portanto, representa o sinal de áudio com energias verdadeiras. A capacidade da decomposição SVD em dissociar os sinais de tempo de áudio (em L), suas energias em (S) e suas características espaciais (em V) pode apoiar vários aspectos das técnicas descritas na presente descrição. Além disso, o modelo de síntese dos coeficientes de HOA[k] subjacentes, X , por uma multiplicação de vetor de $US[k]$ e $V[k]$ gera o termo

"decomposição baseada em vetor", que é usado em todo o presente documento.

[0056] Apesar de descrita como sendo realizada diretamente com respeito a coeficientes de HOA 11, a unidade de LIT 30 pode aplicar a transformada inversível linear a derivados dos coeficientes de HOA 11. Por exemplo, a unidade de LIT 30 pode aplicar SVD com respeito a uma matriz de densidade espectral de energia derivada dos coeficientes de HOA 11. A matriz de densidade espectral de energia pode ser indicada como PSD e obtida por meio de multiplicação de matriz da transposição de `hoaFrame` para o `hoaFrame`, como mostrado no pseudocódigo que se segue abaixo. A designação `hoaFrame` se refere a um quadro dos coeficientes de HOA 11.

[0057] A unidade de LIT 30 pode, após a aplicação do SVD (SVD) ao PSD, obter uma matriz $S[k]^2$ (`S_squared`) e uma matriz $V[k]$. A matriz $S[k]^2$ pode indicar uma matriz quadrada $S[k]$, com o que a unidade de LIT 30 pode aplicar uma operação de raiz quadrada à matriz de $S[k]^2$ para obter a matriz de $S[k]$. A unidade de LIT 30 pode, em alguns casos, executar quantização com respeito à matriz $V[k]$ para obter uma matriz $V[k]$ quantizada (que pode ser chamada matriz $V[k]$). A unidade de LIT 30 pode obter a matriz $U[k]$ primeiro multiplicando a matriz $S[k]$ pela matriz $V[k]'$ quantizada para obter uma matriz $SV[k]'$. A unidade de LIT 30 pode, em seguida, obter o pseudo-inverso (`pinv`) da matriz $SV[k]$ e então multiplicar os coeficientes de HOA 11 pelo pseudo-inverso da matriz $SV[k]'$ para obter a matriz $U[k]$. O acima pode ser representado pela pseudocódigo que se segue:

```

PSD                                =          hoaFrame*hoaFrame;
[V,S_squared]                     =          svd(PSD,'econ');

```

```

S          =          sqrt          (S_squared);
U = hoaFrame * pinv (S*V').

```

[0058] Ao realizar SVD no que diz respeito à densidade espectral de energia (PSD) dos coeficientes de HOA em vez dos próprios coeficientes, a unidade de LIT 30 pode, potencialmente, reduzir a complexidade computacional de realização do SVD em termos de um ou mais dos ciclos do processador e espaço de armazenamento, enquanto obtendo a mesma eficiência de codificação de áudio de origem como se os SVD fossem aplicados diretamente aos coeficientes de HOA. Isto é, o SVD do tipo PSD descrito acima pode ser potencialmente menos exigente computacionalmente porque o SVD é feito em uma matriz $F \times F$ (com F o número de coeficientes de HOA), em comparação com uma matriz $M \times F$ com M sendo o comprimento do quadro, ou seja, 1024 ou mais amostras. A complexidade de um SVD pode agora, através da aplicação ao PSD em vez dos coeficientes de HOA 11, ser em torno de $O(L^3)$ em comparação com $O(M \times L^2)$, quando aplicado aos coeficientes de HOA 11 (onde $O(*)$ indica a notação big-O de complexidade computacional comum às artes da ciência da computação).

[0059] A este respeito, a unidade de LIT 30 pode executar uma decomposição com relação a ou de outra forma decompor os dados de áudio ambisonic de ordem superior 11 para obter um vetor (por exemplo, o Vetor V acima) representativo de eixo espacial ortogonal em um domínio de harmônicos esféricos. A decomposição pode incluir SVD, EVD ou qualquer outra forma de decomposição.

[0060] A unidade de cálculo de parâmetros 32 representa uma unidade configurada para calcular vários parâmetros, tais como um parâmetro de correlação (R), parâmetros de propriedades direcionais (θ , ϕ , r) e uma propriedade de energia (E). Cada um dos parâmetros para o

quadro atual pode ser denominado $R[k]$, $\theta[k]$, $\phi[k]$, $r[k]$ e $e[k]$. A unidade de cálculo de parâmetros 32 pode executar uma análise e/ou correlação de energia (ou a chamada correlação cruzada) com respeito aos vetores $US[k]$ 33 para identificar os parâmetros. A unidade de cálculo de parâmetros 32 pode também determinar os parâmetros para o quadro anterior, em que os parâmetros do quadro anterior podem ser denominados $R[k-1]$, $\theta[k-1]$, $\phi[k-1]$, $r[k-1]$ e $e[k-1]$, com base no quadro anterior de vetor $US[k-1]$ e vetores $V[k-1]$. A unidade de cálculo de parâmetros 32 pode emitir os parâmetros atuais 37 e os parâmetros anteriores 39 para a unidade de reordenação 34.

[0061] A decomposição SVD não garante que o sinal/objeto de áudio representado pelo Vetor p-ésimo em vetores 33 $US[k-1]$, que pode ser denominado vetor $US[k-1][p]$ (ou, alternativamente, como $XPS(p)((k-1))$), será o mesmo sinal/objeto de áudio (progredido no tempo) representado pela Vetor p-ésimo nos vetores 33 $US[k]$, que também pode ser designado vetores 33 $US[k][p]$ (ou, alternativamente, como $US[k]$). Os parâmetros calculados pela unidade de cálculo de parâmetros 32 podem ser utilizados pela unidade de reordenamento 34 para reordenar os objetos de áudio para representar sua avaliação ou continuidade natural ao longo do tempo.

[0062] Ou seja, a unidade de reordenamento 34 pode comparar cada um dos parâmetros 37 dos primeiros vetores 33 $US[k]$ em transformada contra cada um dos parâmetros 39 para os segundos vetores 33 $US[k-1]$. A unidade de reordenação 34 pode reordenar (utilizando, como um exemplo, um algoritmo Húngaro) os vários vetores dentro da matriz 33 $US[k]$ e da matriz 35 $V[k]$ com base nos parâmetros 37 atuais e nos parâmetros 39 anteriores para produção de uma matriz 33'

US[k] reordenada (que pode ser matematicamente designada $\overline{US}[k]$) e uma matriz 35' V[k] reordenada (que pode ser matematicamente designada $\overline{V}[k]$) para uma unidade de seleção de som de primeiro plano 36 (ou som predominante - PS) ("unidade de seleção de primeiro plano 36") e uma unidade de compensação de energia 38.

[0063] A unidade de análise de campo de som 44 pode representar uma unidade configurada para realizar uma análise de campo de som com respeito aos coeficientes de HOA 11 de modo a potencialmente atingir uma taxa de bits alvo 41. A unidade de análise de campo de som 44 pode, com base na análise e/ou uma taxa de bits alvo recebida 41, determinar o número total de instanciações de codificador psicoacústicas (que pode ser uma função do número total de canais de ambiente ou de base (BG_{TOT}) e o número de canais de primeiro plano ou, em outras palavras, os canais predominantes. o número total de instanciações de codificador psicoacústicas pode ser chamado numHOATransportChannels.

[0064] A unidade de análise de campo de som 44 também pode determinar, novamente para potencialmente alcançar a taxa de bits 41 alvo, o número total de canais do primeiro plano (nFG) 45, a ordem mínima do campo de som de base (ou, em outras palavras, ambiente) (N_{BG} ou, alternativamente, MinAmbHOAorder), o número correspondente de canais reais representativos da ordem mínima de campo de som de base ($nBGa = (MinAmbHOAorder + 1)^2$), e os índices (i) de canais adicionais de HOA BG para enviar (os quais podem coletivamente ser referidas como informação de canal de base 43 no exemplo da FIG. 3). As informações do canal de base 42 podem também ser referidas como informação do canal de ambiente 43. Cada um dos canais que permanece de

numHOATransportChannels - nBGa pode tanto ser um "canal de base/ambiente adicional", um "canal predominante baseado em Vetor ativo", um "sinal predominante de base direccional ativo" ou "completamente inativo". Em um aspecto, os tipos de canais podem ser elemento de sintaxe indicado (como um "ChannelType") por dois bits (por exemplo, 00: sinal de base direccional; 01: sinal predominante baseado em vetor; 10: sinal ambiente adicional; 11: sinal inativo). O número total de sinais de base ou de ambiente, nBGa, pode ser dado por $(\text{MinAmbHOAorder} + 1)^2$ + o número de vezes que o índice 10 (no exemplo acima) aparece como um tipo de canal no fluxo de bits para este quadro.

[0065] Em qualquer caso, a unidade de análise de campo de som 44 pode seleccionar o número de canais de base (ou, em outras palavras, de ambiente) e o número de canais de primeiro plano (ou, em outras palavras, predominante) com base na taxa de bits alvo 41, seleccionando mais canais de base e/ou primeiro plano quando a taxa de bits alvo 41 é relativamente mais elevada (por exemplo, quando a taxa de bits alvo 41 é igual ou superior a 512 Kbps). Em um aspecto, os numHOATransportChannels pode ser ajustado para 8 enquanto o MinAmbHOAorder pode ser ajustado para 1 na seção de cabeçalho do fluxo de bits. Neste cenário, a cada quadro, quatro canais podem ser dedicados para representar a parte de base ou ambiente do campo de som enquanto os outros 4 canais podem, numa base de quadro-a-quadro, variar com o tipo de canal - por exemplo, seja utilizado como um canal de base/ambiente adicional ou um canal de primeiro plano/predominante. Os sinais de primeiro plano/predominantes podem ser qualquer um dos sinais baseados em vetor ou de base direccional, tal como descrito acima.

[0066] Em alguns casos, o número total de sinais predominantes com base em vetor para um quadro pode ser determinado pelo número de vezes que o índice ChannelType é 01 no fluxo de bits desta estrutura. No aspecto acima, para cada canal de base/ambiente adicional (por exemplo, correspondendo a um ChannelType de 10), a informação correspondente de qual dos coeficientes de HOA possíveis (além dos quatro primeiros) pode ser representado nesse canal. A informação, para conteúdo de HOA de quarta ordem, pode ser um índice para indicar os coeficientes de HOA 5-25. Os quatro primeiros coeficientes de HOA de ambientes 1-4 podem ser enviados o tempo todo quando minAmbHOAorder é ajustado 1, desta maneira o dispositivo de codificação de áudio pode precisar apenas indicar um do coeficiente de HOA ambiente adicional tendo um índice de 5-25. A informação poderia então ser transmitida através de um elemento de sintaxe de 5 bits (para conteúdo de 4ª ordem), o que pode ser indicado como "CodedAmbCoeffldx."

[0067] Para ilustrar, supondo que o minAmbHOAorder é ajustado para 1 e um coeficiente de HOA de ambiente adicional com um índice de seis é enviado através do fluxo de bits 21 como um exemplo. Neste exemplo, o minAmbHOAorder de 1 indica que os coeficientes de HOA de ambiente têm um índice de 1, 2, 3 e 4. O dispositivo de codificação de áudio 20 pode selecionar os coeficientes de HOA de ambiente porque os coeficientes de HOA de ambiente têm um índice inferior a ou igual a $(\text{minAmbHOAorder} + 1)^2$ ou 4 neste exemplo. O dispositivo de codificação de áudio 20 pode especificar os coeficientes de HOA de ambiente associados aos índices de 1, 2, 3 e 4 no fluxo de bits 21. O dispositivo de codificação de áudio 20 também pode especificar o coeficiente de HOA de ambiente adicional com um índice de 6 no fluxo de bits como um

additionalAmbientHOAchannel com um ChannelType de 10. O dispositivo de codificação de áudio 20 pode especificar o índice usando o elemento de sintaxe CodedAmbCoefflidx. Como uma questão prática, o elemento CodedAmbCoefflidx pode especificar todos os índices de 1-25. No entanto, devido ao fato do minAmbHOAorder ser definido para um, o dispositivo de codificação de áudio 20 pode não especificar qualquer um dos quatro primeiros índices (como os primeiros quatro índices são conhecidos ser especificados no fluxo de bits 21 através do elemento de sintaxe minAmbHOAorder). Em todo caso, devido ao dispositivo de codificação de áudio 20 especificar os cinco coeficientes de HOA de ambiente através do minAmbHOAorder (para os primeiros quatro) e do CodedAmbCoefflidx (para o coeficiente de HOA de ambiente adicional), o dispositivo de codificação de áudio 20 pode não especificar os elementos de vetor V correspondentes associados com os coeficientes de HOA de ambientes tendo um índice de 1, 2, 3, 4 e 6. Como resultado, o dispositivo de codificação de áudio 20 podem especificar o vetor V com elementos [5, 7:25].

[0068] Em um segundo aspecto, todos sinais do primeiro plano/predominantes são os sinais à base de vetor. Neste segundo aspecto, o número total de sinais de primeiro plano/predominantes pode ser dado por $nFG = \text{numHOATransportChannels} - [(\text{MinAmbHOAorder} - 1)^2 + \text{cada um dos additionalAmbientHOAchannel}]$.

[0069] A unidade de análise de campo de som 44 emite a informação de canal de base 43 e os coeficientes de HOA 11 para a unidade de seleção de base (BG) 36, a informação do canal de base 43 para unidade de redução de coeficiente 46 e a unidade de geração de fluxo de bits 42, e a NFG 45 para uma unidade de seleção de primeiro plano 36.

[0070] A unidade de seleção de base 48 pode representar uma unidade configurada para determinar os coeficientes de HOA de base ou ambiente 47 com base nas informações de canal de base (por exemplo, o campo de som de base (N_{BG}) e o número ($nBGa$) e os índices (i) de canais de HOA BG adicionais para enviar). Por exemplo, quando N_{BG} é igual a um, a unidade de seleção de base 48 pode selecionar os coeficientes de HOA 11 para cada uma das amostras do quadro de áudio tendo uma ordem igual a ou menor que um. A unidade de seleção de base 48 pode, neste exemplo, então selecionar os coeficientes de HOA de 11 tendo um índice identificado por um dos índices (i) como coeficientes de HOA BG adicionais, onde o $nBGa$ é provido para a unidade de geração de fluxo de bits 42 para ser especificado no fluxo de bits 21 de modo a permitir que o dispositivo de decodificação de áudio, tal como o dispositivo de decodificação de áudio 24 mostrado no exemplo das FIGS. 2 e 4, analise os coeficientes de HOA de base 47 dos fluxos de bits de 21. A unidade de seleção de base 48 pode então emitir os coeficientes de HOA de ambiente 47 para a unidade de compensação de energia 38. Os coeficientes de HOA 47 de ambiente podem ter dimensões $D: M \times [(N_{BG} + 1)^2 + nBGa]$. Os coeficientes de HOA de ambiente 47 podem também ser referidos como "coeficientes de HOA de ambiente 47", em que cada um dos coeficientes de HOA de ambiente 47 corresponde a um canal de HOA ambiente separado 47 a ser codificado pela unidade codificadora de áudio psicoacústica 40.

[0071] A unidade de seleção de primeiro plano 36 pode representar uma unidade configurada para selecionar a matriz $US[k]$ 33' e a matriz $V[k]$ 35' reordenada que representam os componentes de primeiro plano ou distintos do campo de som com base em nFG 45 (o que pode representar

um ou mais índices que identificam os vetores de primeiro plano). A unidade de seleção de primeiro plano 36 pode emitir sinais de nFG 49 (que podem ser chamados $US[k]_1, \dots, nFG\ 49, FG_1, \dots, nFG[k]\ 49$ ou $X_{PS}^{(1..nFG)}(k)\ 49$) para a unidade codificadora de áudio psicoacústica 40, onde os sinais nFG 49 podem ter dimensões: $D: M \times nFG$ e cada um representa objetos monoáudio. A unidade de seleção de primeiro plano 36 pode também emitir a matriz $V[k]\ 35'$ reordenada (ou $v^{(1..nFG)}(k)\ 35'$) correspondendo aos componentes de primeiro plano do campo de som para a unidade de interpolação espaço-temporal 50, em que um subconjunto da matriz $V[k]\ 35$ reordenada correspondendo aos componentes de primeiro plano pode ser chamado matriz $V[k]\ 51_k$ (que pode ser matematicamente chamada $\bar{V}_{1, \dots, nFG}[k]$) tendo as dimensões $D: (N+1)^2 \times nFG$.

[0072] A unidade de compensação de energia 38 pode representar uma unidade configurada para executar a compensação de energia no que diz respeito aos coeficientes de HOA de ambiente 47 para compensar a perda de energia devido à remoção de várias dos canais de HOA pela unidade de seleção de base 48. A unidade de compensação de energia 38 pode executar uma análise de energia em relação a um ou mais da matriz $US[k]\ 33'$ reordenada, da matriz $V[k]\ 35'$ reordenada, dos sinais nFG 49, dos vetores de base $V[k]\ 51_k$ e dos coeficientes de HOA de ambiente 47 e então realizar compensação de energia com base na análise de energias para gerar coeficientes de HOA de ambiente compensados em energia 47'. A unidade de compensação de energia 38 pode emitir os coeficientes de HOA de ambiente de energia compensada 47' para a unidade de codificação de áudio psicoacústica 40.

[0073] A unidade de interpolação espaço-temporal 50 pode representar uma unidade configurada para receber os vetores de primeiro plano $V[k]$ 51_k para o k-ésimo quadro e os vetores de primeiro plano $V[k-1]$ 51_{k-1} para o quadro anterior (desta maneira a notação $k-1$) e executar interpolação espaço-temporal para gerar vetores de primeiro plano $V[k]$ interpolados. A unidade de interpolação espaço-temporal 50 pode recombina os sinais nFG 49 com os vetores de primeiro plano $V[k]$ 51_k para recuperar coeficientes de HOA de primeiro plano reordenados. A unidade de interpolação espaço-temporal 50 pode então dividir os coeficientes de HOA de primeiro plano reordenados pelos vetores $V[k]$ interpolados para gerar sinais nFG interpolados 49'. A unidade de interpolação espaço-temporal 50 pode também emitir vetores $V[k]$ de primeiro plano 51_k que foram usados para gerar os vetores $V[k]$ de primeiro plano interpolados de maneira que um dispositivo de decodificação de áudio, tal como o dispositivo de decodificação de áudio 24, pode gerar os vetores $V[k]$ de primeiro plano interpolados e desta maneira recuperar os vetores $V[k]$ de primeiro plano 51_k. Os vetores $V[k]$ de primeiro plano 51_k usados para gerar os vetores $V[k]$ de primeiro plano interpolados são chamados os vetores $V[k]$ de primeiro plano restantes 53. A fim de assegurar que os mesmos $V[k]$ e $V[k-1]$ sejam usados no codificador e decodificador (para criar os vetores $V[k]$ interpolados) versões quantizadas/desquantizadas dos vetores podem ser usadas como o codificador e decodificador.

[0074] Em operação, a unidade de interpolação espaço-temporal 50 pode interpolar um ou mais subquadros de um primeiro quadro de áudio de uma primeira decomposição, por exemplo, vetores $V[k]$ de primeiro plano 51_k, de uma parte de uma primeira pluralidade dos coeficientes de HOA

11 incluídos no primeiro quadro, e uma segunda decomposição, por exemplo, vetores $V[k]$ de primeiro plano $51k-1$, de uma parte de uma segunda pluralidade dos coeficientes de HOA 11 incluídos em um segundo quadro, para gerar coeficientes harmônicos esféricos interpolados decompostos para o um ou mais subquadros.

[0075] Em alguns exemplos, a primeira decomposição compreende os vetores $V[k]$ de primeiro plano 51_k representativos de vetores direita-singulares da parte dos coeficientes de HOA 11. Do mesmo modo, em alguns exemplos, a segunda decomposição compreende os segundos vetores $V[k]$ de primeiro plano 51_k vetores representativos de vetores direita-singulares da parte dos coeficientes de HOA 11.

[0076] Em outras palavras, áudio 3D baseado em harmônicos esféricos pode ser uma representação paramétrica do campo de pressão 3D em termos de funções de base ortogonais em uma esfera. Quanto maior a ordem de N da representação, potencialmente maior a resolução espacial, e muitas vezes maior o número de coeficientes harmônicos esféricos (SH) (para um total de $(n + 1)^2$ coeficientes). Para muitas aplicações, uma compressão da largura de banda dos coeficientes pode ser necessária para ser capaz de transmitir e armazenar os coeficientes eficientemente. As técnicas dirigidas nesta descrição podem fornecer um processo de redução de dimensionalidade, com base em quadros, usando Decomposição de Valor Singular (SVD). A análise SVD pode decompor cada quadro de coeficientes em três matrizes U , S e V . Em alguns exemplos, as técnicas podem lidar com alguns dos vetores na matriz $US[k]$ como componentes de primeiro plano do campo de som subjacente. No entanto, quando tratados desta maneira, os vetores (na matriz $US[k]$) são descontínuos de um quadro para outro - mesmo que eles representem o mesmo componente de áudio

distinto. As descontinuidades podem levar a artefatos significativos quando os componentes são alimentados através de áudio-codificadores de transformada.

[0077] Em alguns aspectos, a interpolação espaço-temporal pode contar com a observação que a matriz V pode ser interpretada como eixos espaciais ortogonais no domínio de Harmônicos Esféricos. A matriz $U[k]$ pode representar uma projeção dos dados de Harmônicos Esféricos (HOA) em termos das funções de base, em que a descontinuidade pode ser atribuída a eixos ortogonais espaciais ($V[k]$) que mudam a cada quadro - e são, portanto, descontínuos por si só. Isto é ao contrário de algumas outras decomposições, como o Transformada de Fourier, onde as funções de base são, em alguns exemplos, constantes de quadro para quadro. Nestes termos, o SVD pode ser considerado um algoritmo de busca de correspondência. A unidade de interpolação espaço-temporal 50 pode executar a interpolação para potencialmente manter a continuidade entre as funções de base ($V[k]$) de quadro para quadro - por interpolação entre eles.

[0078] Como mencionado acima, a interpolação pode ser realizada em relação a amostras. O caso é generalizado na descrição acima quando os subquadros compreendem um único conjunto de amostras. Em ambos os casos de interpolação em amostras e subquadros, a operação de interpolação pode assumir a forma de equação que segue:

$$\overline{v(l)} = w(l)v(k) + (1 - w(l))v(k - 1).$$

Na equação acima, a interpolação pode ser realizada em relação ao vetor V único $v(k)$ a partir do vetor V único $v(k-1)$, que em um aspecto poderia representar vetores V de quadros k e $k-1$ adjacentes. Na equação acima, l , representa

a resolução na qual a interpolação está sendo realizada, em que l pode indicar uma amostra de número inteiro e $l = 1, \dots, T$ (em que T é o comprimento de amostras nas quais a interpolação está sendo realizada e nas quais os vetores interpolados de saída, $\overline{v(l)}$, são requeridos e também indica que a saída do processo produz l dos vetores). Alternativamente, l poderia indicar subquadros que consistem em amostras múltiplas. Quando, por exemplo, um quadro é dividido em quatro subquadros, l pode compreender valores de 1, 2, 3 e 4, para cada um dos subquadros. O valor de l pode ser assinalado como um campo chamado "CodedSpatialInterpolationTime" através de um fluxo de bits - de modo que a operação de interpolação pode ser replicada no decodificador. A $w(l)$ pode compreender valores dos pesos de interpolação. Quando a interpolação é linear, $w(l)$ pode variar linearmente e monotonicamente entre 0 e 1, como uma função de l . Em outros casos, $w(l)$ pode variar entre 0 e 1 de uma forma não-linear, mas monotônica (tal como um quarto de ciclo de cosseno elevado) como uma função de l . A função $w(l)$ pode ser indexada entre algumas possibilidades diferentes de funções e sinalizada no fluxo de bits como um campo chamado "SpatialInterpolationMethod" de tal modo que a operação de interpolação idêntica pode ser replicada pelo decodificador. Quando $w(l)$ tem um valor próximo de 0, a saída, $\overline{v(l)}$, pode ser altamente ponderada ou influenciada por $v(k-1)$. Enquanto que quando $w(l)$ tem um valor próximo de 1, assegura-se que a saída, $\overline{v(l)}$, é altamente ponderada ou influenciada por $v(k - 1)$.

[0079] A unidade de redução do coeficiente 46 pode representar uma unidade configurada para efetuar redução do coeficiente com respeito aos vetores $V[k]$ de primeiro plano restantes 53 com base nas informações de canal de base 43

para emitir vetores $V[k]$ de primeiro plano reduzidos 55 para a unidade de quantização 52. Os vetores de primeiro plano $V[k]$ reduzidos 55 podem ter dimensões D: $[(N+1)^2 - (N_{BG}+1)2 - BG_{TOT}] \times n_{FG}$.

[0080] A unidade de redução de coeficiente 46 pode, a este respeito, representar uma unidade configurada para reduzir o número de coeficientes nos vetores de primeiro plano restantes $V[k]$ 53. Em outras palavras, a unidade de redução do coeficiente 46 pode representar uma unidade configurada para eliminar os coeficientes nos vetores de primeiro plano $V[k]$ (que formam o restante dos vetores de primeiro plano $V[k]$ 53) que têm pouca ou nenhuma informação direcional. Como descrito acima, em alguns exemplos, os coeficientes dos vetores $V[k]$ distintos ou, em outras palavras, vetores de primeiro plano correspondendo a funções de base de primeira ordem e ordem zero (que podem ser chamados N_{BG}) proveem pouca ou nenhuma informação e, portanto, podem ser removidos dos vetores V de primeira ordem (através de um processo que pode ser referido como "redução de coeficiente"). Neste exemplo, maior flexibilidade pode ser provida para não apenas identificar os coeficientes que correspondem a N_{BG} , mas identificar canais de HOA adicionais (que podem ser indicados pela variável `TotalOfAddAmbHOAChan`) do conjunto de $[(N_{BG} + 1)2 + 1, (N+1)^2]$. A unidade de análise de campo de som 44 pode analisar os coeficientes de HOA 11 para determinar BG_{TOT} , que pode identificar não só a $(N_{BG}+1)^2$, mas o `TotalOfAddAmbHOAChan`, que pode ser coletivamente referida como a informação de canal de base 43. A unidade de redução de coeficiente 46 pode, em seguida, remover os coeficientes correspondentes a $(N_{BG}+1)^2$ e a `TotalOfAddAmbHOAChan` do restante dos vetores de primeiro plano $V[K]$ 53 para gerar uma matriz dimensional menor $V[K]$ 55 de tamanho $((N+1)^2 -$

(BG_{TOT}) X nFG , que também pode ser referida como os vetores $V[k]$ de primeiro plano reduzidos 55.

[0081] Em outras palavras, como observado na publicação No. WO 2014/194099, a unidade de redução de coeficiente 46 pode gerar elementos de sintaxe para a informação de canal lateral 57. Por exemplo, a unidade de redução de coeficiente 46 pode especificar um elemento de sintaxe em um cabeçalho de uma unidade de acesso (que pode incluir um ou mais quadros) indicando qual da pluralidade de modos de configuração foi selecionada. Embora descrita como sendo especificada em uma base unitária por acesso, a unidade de redução de coeficiente 46 pode especificar o elemento de sintaxe em uma base por quadro ou qualquer outra base periódica ou base não periódica (tal como uma vez para todo o fluxo de bits). Em qualquer caso, o elemento de sintaxe pode compreender dois bits indicando qual dos três modos de configuração foi selecionado para especificar o conjunto diferente de zero dos coeficientes dos vetores $V[k]$ de primeiro plano reduzidos 55 para representar os aspectos direcionais do componente distinto. O elemento de sintaxe pode ser chamado "CodedVVecLength." Desta maneira, a unidade de redução do coeficiente 46 pode sinalizar ou de outra maneira especificar no fluxo de bits quais dos três modos de configuração foi usado para especificar os vetores $V[k]$ de primeira ordem reduzidos 55 no fluxo de bits 21.

[0082] Por exemplo, três modos de configuração podem ser apresentados na tabela a sintaxe para VVecData (posteriormente referido neste documento). Nesse exemplo, os modos de configuração são os seguintes: (Modo 0), um comprimento de vetor V total é transmitido no campo VVecData; (Modo 1), os elementos do vetor V relacionado com o número mínimo de coeficientes para os coeficientes de HOA

de Ambiente e todos os elementos do vetor V que incluíam canais de HOA adicionais que não são transmitidos; e (Modo 2), os elementos do vetor V relacionado com o número mínimo de coeficientes para os coeficientes de HOA de Ambiente não são transmitidos. A tabela de sintaxe de VVecData ilustra os modos em conexão com uma mudança e exposição do caso. Embora descritas em relação a três modos de configuração, as técnicas não devem ser limitadas a três modos de configuração e podem incluir qualquer número de modos de configuração, incluindo um modo de configuração único ou uma pluralidade de modos. A Publicação No. WO 2014/194099 provê um exemplo diferente com quatro modos. A unidade de redução de coeficiente 46 pode também especificar o sinalizador 63 como um outro elemento de sintaxe na informação do canal lateral 57.

[0083] A unidade de quantização 52 pode representar uma unidade configurada para realizar qualquer forma de quantização para comprimir os vetores V[k] 55 de primeiro plano para gerar vetores V[k] de primeiro plano 57 codificados, enviando os vetores V[k] de primeiro plano 57 codificados para a unidade de geração de fluxo de bits 42. Em operação, a unidade de quantização 52 pode representar uma unidade configurada para comprimir um componente espacial do campo de som, ou seja, um ou mais dos vetores V[k] de primeiro plano 55 neste exemplo. Para fins de exemplo, os vetores V[k] de primeiro plano reduzidos 55 são assumidos incluir dois vetores em linha tendo, como resultado de redução do coeficiente, menos do que 25 elementos de cada um (o que implica uma representação de HOA de quarta ordem do campo de som). Embora descrito em relação a dois vetores em linha, qualquer número de vetores pode ser incluído nos vetores V[k] de primeiro plano reduzidos 55 até $(n+1)^2$, em que n indica a ordem da

representação de HOA do campo de som. Além disso, embora descritos abaixo como desempenhando uma quantização escalar e/ou entropia, a unidade de quantização 52 pode executar qualquer forma de quantização que resulte na compressão dos vetores $V[k]$ de primeiro plano reduzidos 55.

[0084] A unidade de quantização 52 pode receber os vetores $V[k]$ de primeiro plano reduzidos 55 e executar um esquema de compressão para gerar vetores $V[k]$ de primeiro plano 57 codificados. O esquema de compressão pode envolver qualquer esquema de compressão concebível para comprimir elementos de um vetor ou dados geralmente, e não deve ser limitado ao exemplo descrito abaixo em mais detalhes. A unidade de quantização 52 pode executar, como um exemplo, um esquema de compressão que inclui um ou mais de transformar representações de pontos flutuantes de cada elemento dos vetores $V[k]$ de primeiro plano reduzidos 55 em representações de inteiro de cada elemento dos vetores $V[k]$ de primeiro plano reduzidos 55, quantização uniforme das representações de inteiro dos vetores $V[k]$ de primeiro plano reduzidos 55 e categorização e codificação das representações de inteiros quantizadas dos vetores $V[k]$ de primeiro plano 55 restantes.

[0085] Em alguns exemplos, vários dos um ou mais processos do esquema de compressão podem ser controlados de forma dinâmica por meio de parâmetros para obter ou quase obter, como um exemplo, uma taxa de bits alvo 41 para o fluxo de bits resultante 21. Dado que cada um dos vetores $V[k]$ de primeiro plano reduzidos 55 são ortonormais um para o outro, cada um dos vetores $V[k]$ de primeiro plano reduzidos 55 pode ser codificado de forma independente. Em alguns exemplos, tal como descrito em mais detalhes abaixo, cada elemento dos vetores $V[k]$ de primeiro plano reduzidos

55 pode ser codificado utilizando o mesmo modo de codificação (definido por vários submodos).

[0086] Conforme descrito na publicação No. WO 2014/194099, a unidade de quantização 52 pode executar quantização escalar e/ou codificação de Huffman para comprimir os vetores $V[k]$ de primeiro plano 55, emitindo os vetores $V[k]$ de primeiro plano codificados 57, que também podem ser referidos como informações de canal lateral 57. As informações de canal lateral 57 podem incluir elementos de sintaxe usados para codificar o restante dos vetores $V[k]$ de primeiro plano 55.

[0087] Além disso, embora descrita em relação a uma forma de quantização escalar, a unidade de quantização 52 pode executar quantização vetorial ou qualquer outra forma de quantização. Em alguns casos, a unidade de quantização 52 pode alternar entre quantização vetorial e quantização escalar. Durante a quantização escalar descrita acima, a unidade de quantização 52 pode calcular a diferença entre dois vetores V sucessivos (sucessivos como em quadro-para-quadro) e codificar a diferença (ou, e outras palavras, residuais). Esta quantização escalar pode representar uma forma de codificação preditiva com base em um vetor anteriormente especificado e um sinal de diferença. Quantização de vetor não envolve tal codificação de diferença.

[0088] Em outras palavras, a unidade de quantização 52 pode receber um vetor V de entrada (por exemplo, um dos vetores $V[k]$ de primeiro plano reduzidos 55) e realizar tipos diferentes de quantização para selecionar um dos tipos de quantização a ser utilizada para o vetor V de entrada. A unidade de quantização 52 pode, como um exemplo, realizar quantificação vetorial, quantização escalar sem

codificação de Huffman e quantização escalar com codificação de Huffman.

[0089] Neste exemplo, a unidade de quantização 52 pode realizar quantização vetorial do vetor V de entrada de acordo com um modo de quantização vetorial para gerar um vetor V de quantização vetorial. O vetor V de quantização vetorial pode incluir valores de peso de quantização vetorial que representam o vetor V de entrada. Os valores de peso de quantização vetorial podem, em alguns exemplos, ser representados como um ou mais índices de quantificação que apontam para uma palavra de código de quantização (isto é, vetor de quantização) em um livro de códigos de quantização de palavras de código de quantização. A unidade de quantização 52 pode, quando configurada para executar quantização vetorial, decompor cada um dos vetores $V[k]$ de primeiro plano reduzidos 55 em uma soma ponderada de vetores de código com base em vetores de código 63 ("CV 63"). A unidade de quantização 52 pode gerar valores de peso para cada um dos selecionados dos vetores de codificação 63.

[0090] A unidade de quantização 52 pode, em seguida, selecionar um subconjunto dos valores de peso para gerar um subconjunto selecionado de valores de peso. Por exemplo, a unidade de quantização 52 pode selecionar os valores de peso Z de maior grandeza do conjunto de valores de peso para gerar o subconjunto selecionado dos valores de peso. Em alguns exemplos, a unidade de quantização 52 pode reordenar ainda mais os valores de peso selecionados para gerar o subconjunto selecionado de valores de peso. Por exemplo, a unidade de quantização 52 pode reordenar os valores de peso selecionados com base em grandeza começando a partir de um valor de peso de grandeza mais elevada e terminando em um valor de peso de grandeza mais baixa.

[0091] Ao realizar a quantização vetorial, a unidade de quantização 52 pode selecionar um vetor do componente Z a partir de um livro de códigos de quantização para representar os valores de peso Z. Em outras palavras, a unidade de quantização 52 pode quantizar em vetor valores de peso Z para gerar um vetor do componente Z que representa os valores de peso Z. Em alguns exemplos, Z pode corresponder ao número de valores de peso selecionados pela unidade de quantização 52 para representar um único vetor V. A unidade de quantização 52 pode gerar dados indicativos do vetor do componente Z selecionado para representar os valores de peso Z, e fornecem estes dados para a unidade de geração de fluxo de bits 42 como os pesos codificados 57. Em alguns exemplos, o livro de códigos de quantização pode incluir uma pluralidade de vetores do componente Z que são indexados, e os dados indicativos do vetor do componente Z podem ser um valor do índice no livro de códigos para quantização que aponta para o vetor selecionado. Nestes exemplos, o decodificador pode incluir um livro de códigos de quantização indexada de modo semelhante para decodificar o valor de índice.

[0092] Matematicamente, cada um dos vetores $V[k]$ 55 de primeiro plano reduzidos podem ser representados com base na seguinte expressão:

29

$$V \approx \sum_{j=1}^J \omega_j \Omega_j$$

(I)

em que Ω_j representa o vetor de código j -ésimo em um conjunto de vetores de código ($\{\Omega_j\}$), ω_j representa o peso j -ésimo em um conjunto de pesos ($\{\omega_j\}$), V corresponde ao

vetor V que está sendo representado, decomposto, e / ou codificado pela unidade de codificação do vetor V 52, e J representa o número de pesos e o número de vetores de código utilizados para representar V . O lado direito da expressão (1) pode representar uma soma ponderada dos vetores de código que inclui um conjunto de pesos ($\{\omega_j\}$) e um conjunto de vetores de código ($\{\Omega_j\}$).

[0093] Em alguns exemplos, a unidade de quantização 52 pode determinar os valores de peso com base na seguinte equação:

$$\omega_k = V \Omega_k^T \quad (2)$$

em que Ω_k^T representa uma transposição do vetor de código k -ésimo em um conjunto de vetores de código ($\{\Omega_k\}$), V corresponde ao vetor V que está sendo representado, decomposto e / ou codificado pela unidade de quantização 52, e Ω_k representa o peso k -ésimo em um conjunto de pesos ($\{\omega_k\}$).

[0094] Considere um exemplo em que 25 pesos 25 e vetores de código são usados para representar um vetor V , V_{FG} . Tal decomposição de V_{FG} pode ser escrita como:

$$V_{FG} \approx \sum_{j=1}^{25} \omega_j \Omega_j \quad (3)$$

em que Ω_j representa o vetor de código j -ésimo em um conjunto de vetores de código ($\{\Omega_j\}$), ω_j representa o j -ésimo peso em um conjunto de pesos ($\{\omega_j\}$), e V_{FG} corresponde ao vetor V que está sendo representado, decomposto e / ou codificado pela unidade de quantização 52.

[0095] Nos exemplos onde o conjunto de vetores de código ($\{\Omega_j\}$) é ortonormal, a expressão a seguir pode ser aplicadas:

$$\Omega_j \Omega_k^T = \begin{cases} 1 & \text{for } j = k \\ 0 & \text{for } j \neq k \end{cases} \quad (4)$$

Nestes exemplos, o lado direito da equação (3) pode simplificar como se segue:

$$V_{FG} \Omega_k^T \approx \left(\sum_{j=1}^{25} \omega_j \Omega_j \right) \Omega_k^T = \omega_k \quad (5)$$

em que ω_k corresponde ao k -ésimo peso M_i na soma ponderada dos vetores de código.

[0096] Para a soma ponderada exemplificativa de vetores de código utilizados na equação (3), a unidade de quantização 52 pode calcular os valores de peso para cada um dos pesos na soma ponderada de vetores de código utilizando a equação (5) (semelhante à equação (2)) e os pesos resultantes podem ser representados como:

$$\{\omega_k\}_{k=1, \dots, 25} \quad (6)$$

Considerar um exemplo onde a unidade de quantização 52 seleciona os cinco valores de peso máximo (isto é, pesos com valores maiores ou valores absolutos). O subconjunto dos valores de peso a ser quantizado pode ser representado como:

$$\{\bar{\omega}_k\}_{k=1, \dots, 5} \quad (7)$$

O subconjunto dos valores de peso em conjunto com os seus vetores de código correspondentes pode ser utilizado para formar uma soma ponderada dos vetores de código que estima o vetor V , como se mostra na seguinte expressão:

$$\bar{V}_{FG} \approx \sum_{j=1}^5 \bar{\omega}_j \Omega_j \quad (8)$$

em que Ω_j representa o j -ésimo vetor de código em um subconjunto dos vetores de código ($\{\Omega_j\}$), $\bar{\omega}_j$ representa o

peso j -ésimo em um subconjunto de pesos $(\{\bar{\omega}_j\})$, e $\bar{V}_{FG}$ corresponde a um vetor V estimado que corresponde ao vetor V sendo decomposto e / ou codificado pela unidade de quantização 52. O lado direito da expressão (1) pode representar uma soma ponderada de vetores de código que inclui um conjunto de pesos $(\{\bar{\omega}_j\})$ e um conjunto de vetores de código $(\{\Omega_j\})$.

[0097] A unidade de quantização 52 pode quantizar o subconjunto dos valores de peso para gerar valores de peso quantizados que podem ser representados como:

$$\{\hat{\omega}_k\}_{k=1,\dots,5} \quad (9)$$

Os valores de peso quantizados em conjunto com os seus vetores de código correspondentes podem ser utilizados para formar uma soma ponderada dos vetores de código que representa uma versão quantizada do vetor V estimado, como mostrado na expressão a seguir:

$$\hat{V}_{FG} \approx \sum_{j=1}^5 \hat{\omega}_j \Omega_j \quad (10)$$

em que Ω_j representa o vetor de código j -ésimo em um subconjunto dos vetores de código $(\{\Omega_j\})$, $\hat{\omega}_j$ representa o peso j -ésimo em um subconjunto de pesos $(\{\hat{\omega}_j\})$, e $\hat{V}_{FG}$ corresponde a um vetor V estimado que corresponde ao vetor V sendo decomposto e / ou codificado pela unidade de quantização 52. O lado direito da expressão (1) pode representar uma soma ponderada de um subconjunto dos vetores de código que inclui um conjunto de pesos $(\{\hat{\omega}_j\})$ e um conjunto de vetores de código $(\{\Omega_j\})$.

[0098] Uma reafirmação alternativa do exposto (que é substancialmente equivalente ao descrito acima) pode

ser como segue. Os vetores V podem ser codificados com base em um conjunto predefinido de vetores de código. Para codificar os vetores V , cada vetor V é decomposto em uma soma ponderada dos vetores de código. A soma ponderada de vetores de código é constituída por pares k de vetores de código predefinidos e pesos associados:

$$V \approx \sum_{j=0}^k \omega_j \Omega_j \quad (11)$$

em que Ω_j representa o j -ésimo vetor de código em um conjunto de vetores de código predefinido $(\{\Omega_j\})$, ω_j representa o j -ésimo peso de valor real em um conjunto de pesos predefinidos $(\{\omega_j\})$, k corresponde ao índice de adendos, que pode ser de até 7, e V corresponde ao vetor V que está sendo codificado. A escolha de k depende do codificador. Se o codificador escolhe uma soma ponderada de dois ou mais vetores de código, o número total de vetores de código predefinido que o codificador pode escolher de é $(N + 1)^2$, cujos vetores de código predefinidos são derivados como coeficientes de expansão HOA nas Tabelas F.3 a F.7 do padrão de Áudio 3D, intitulado "Information Technology - High Efficiency Coding and Media Delivery in Heterogeneous Environments - Part 3: 3D áudio", por ISSO/IEC JTC 1/SC 29/WG 11, datado de 25/07/2014, e identificado pelo documento número ISSO/IEC DIS 23008-3. Quando N é 4, a tabela no Anexo F.5 do padrão de Áudio 3D acima referenciado com 32 direções predefinidas é usada. Em todos os casos, os valores de ponderação absolutos ω são de quantização vetorial em relação aos valores de ponderação predefinidos $\hat{\omega}$ encontrado nas primeiras $k + 1$ colunas da tabela na tabela F.12 do padrão de Áudio 3D acima referenciado e assinalado com o índice de número da fila associado.

[0099] Os sinais numéricos dos pesos ω são codificados separadamente como:

$$s_j = \begin{cases} 1, & \omega_j \geq 0 \\ 0, & \omega_j < 0. \end{cases} \quad (12)$$

[0100] Em outras palavras, após a sinalização do valor k , um vetor V é codificado com índices $k + 1$ que apontam para os $k + 1$ vetores de código predefinidos $\{\Omega_j\}$, um índice que aponta para os pesos quantizados k $\{\hat{\omega}_k\}$ no livro de códigos de ponderação predefinida, e $k + 1$ valores de sinal numérico s_j :

$$\hat{V} = \sum_{j=0}^k (2s_j - 1) \hat{\omega}_j \Omega_j \quad (13)$$

Se o codificador seleciona uma soma ponderada de um vetor de código, um livro de códigos derivado da tabela F.8 da norma de Áudio 3D acima referida é utilizada em combinação com os valores de ponderação absolutos $\hat{\omega}$ na tabela da tabela F11 do padrão de Áudio 3D acima referenciado, em que ambas as tabelas são mostradas abaixo. Além disso, o sinal numérico do valor de ponderação ω pode ser codificado separadamente. A unidade de quantização 52 pode sinalizar quais das tabelas de códigos precedentes estabelecidas nas tabelas acima mencionadas F.3 até F.12 são usadas para codificar o vetor V de entrada usando um elemento de sintaxe de índice do livro de códigos (o qual pode ser indicado como "Codebkldx" abaixo). A unidade de quantização 52 também pode realizar quantização escalar do vetor V de entrada para gerar um vetor V de quantização escalar de saída sem codificação de Huffman do vetor V de quantização escalar. A unidade de quantização 52 pode ainda realizar a quantização escalar do vetor V de entrada de acordo com um modo de quantização escalar e codificação de Huffman para gerar um vetor V de quantização da codificação de Huffman.

Por exemplo, a unidade de quantização 52 pode realizar a quantização escalar do vetor V de entrada para gerar um vetor de V de quantização escalar e codificação de Huffman o vetor V de quantização escalar para gerar um vetor V de quantização escalar e codificação de Huffman.

[0101] Em alguns exemplos, a unidade de quantização 52 pode executar uma forma de quantização vetorial predita. A unidade de quantização 52 pode identificar se a quantização vetorial é predita ou não pela especificação de um ou mais bits (por exemplo, o elemento de sintaxe PFlag) no fluxo de bits 21 que indica se a predição é realizada por quantização vetorial (como identificado por um ou mais bits, por exemplo, o elemento de sintaxe NbitsQ, indicando um modo de quantização).

[0102] Para ilustrar a quantização vetorial predita, a unidade de quantização 42 pode ser configurada para receber valores de peso (por exemplo, grandezas de valor de peso) que correspondem a uma decomposição baseada no vetor de código de um vetor (por exemplo, um vetor v), para gerar valores de peso preditivos com base nos valores de peso recebidos e com base nos valores de peso reconstruídos (por exemplo, valores de peso reconstruídos a partir de um ou mais quadros de áudio anteriores ou posteriores), e para quantização vetorial dos conjuntos de valores de peso preditivos. Em alguns casos, cada valor de peso em um conjunto de valores de peso preditivos pode corresponder a um valor de peso incluído em uma decomposição baseada no vetor de código de um único vetor.

[0103] A unidade de quantização 52 pode receber um valor de peso e um valor de peso reconstruída ponderada a partir de uma codificação anterior ou posterior de um vetor. A unidade de quantização 52 pode gerar um valor de peso preditivo com base no valor de peso e o valor do peso

reconstruído ponderado. A unidade de quantização 42 pode subtrair o valor de peso reconstruído ponderado do valor de peso para gerar o valor de peso preditivo. O valor de peso preditivo pode ser alternativamente referido como, por exemplo, um resíduo, um resíduo de predição, um valor de peso residual, uma diferença do valor de peso, um erro ou um erro de predição.

[0104] O valor de peso pode ser representado como $|w_{i,j}|$, que é uma grandeza (ou valor absoluto) do valor de peso correspondente, $|w_{i,j}|$. Como tal, o valor de peso pode ser alternativamente referido como uma grandeza do valor de peso ou como uma grandeza de um valor de peso. O valor de peso, $|w_{i,j}|$, corresponde ao valor de peso j -ésimo a partir de um subconjunto ordenado de valores de peso para o i -ésimo quadro de áudio. Em alguns exemplos, o subconjunto ordenado de valores de peso pode corresponder a um subconjunto dos valores de peso em uma decomposição baseada em vetor de código do vetor (por exemplo, vetor v) que são ordenados com base na grandeza dos valores de peso (por exemplo, ordenados a partir da maior grandeza para a menor grandeza).

[0105] O valor do peso reconstruído ponderado pode incluir um termo $|\hat{w}_{i-1,j}|$, que corresponde a uma grandeza (ou um valor absoluto) do valor de peso reconstruído correspondente. O valor de peso reconstruído, $\hat{w}_{i-1,j}$, corresponde ao j -ésimo valor de peso reconstruído a partir de um subconjunto ordenado de valores de peso reconstruídos para o $(i-1)$ -ésimo quadro de áudio. Em alguns exemplos, o subconjunto ordenado (ou conjunto) de valores de peso reconstruídos pode ser gerado com base nos valores de peso

predictivos quantizados que correspondem aos valores de peso reconstruídos.

[0106] A unidade de quantização 42 também inclui um fator de ponderação, α_j . Em alguns exemplos, $\alpha_j = 1$, caso em que o valor de peso reconstruído ponderado pode reduzir para $|\hat{w}_{i-1,j}|$. Em outros exemplos, $\alpha_j \neq 1$. Por exemplo, α_j pode ser determinado com base na seguinte equação:

$$\alpha_j = \frac{\sum_{i=1}^I w_{i,j} w_{i-1,j}}{\sum_{i=1}^I w_{i-1,j}^2}$$

em que I corresponde ao número de quadros de áudio usados para determinar α_j . Como mostrado na equação anterior, o fator de ponderação, em alguns exemplos, pode ser determinado com base em uma pluralidade de diferentes valores de peso a partir de uma pluralidade de diferentes quadros de áudio.

[0107] Além disso, quando configurado para executar a quantização vetorial predita, a unidade de quantização 52 pode gerar o valor de peso preditivo com base na seguinte equação:

$$e_{i,j} = |w_{i,j} - \alpha_j \hat{w}_{i-1,j}|$$

em que $e_{i,j}$ corresponde ao valor de peso preditivo para o j -ésimo valor de peso a partir de um subconjunto ordenado de valores de peso para o i -ésimo quadro de áudio.

[0108] A unidade de quantização 52 gera um valor de peso preditivo peso quantizado com base no valor de peso preditivo e um livro de códigos da codificação vetorial predita (PVQ). Por exemplo, a unidade de quantização 52 pode realizar a quantização vetorial do valor de peso preditivo em combinação com outros valores de peso

predictivos gerados para o vetor a ser codificado ou para o quadro a ser codificado, a fim de gerar o valor de peso preditivo quantizado.

[0109] A unidade de quantização 52 pode realizar a quantização vetorial do valor de peso preditivo 620 com base no livro de códigos PVQ. O livro de códigos PVQ pode incluir uma pluralidade de vetores de quantização candidatos do componente M, e a unidade de quantização 52 pode selecionar um dos vetores de quantização candidatos para representar os valores de peso preditivos Z. Em alguns exemplos, a unidade de quantização 52 pode selecionar um vetor de quantização candidato a partir do livro de códigos PVQ que minimiza a um erro de quantização (por exemplo, minimiza um erro de quadrados mínimos).

[0110] Em alguns exemplos, o livro de códigos PVQ pode incluir uma pluralidade de entradas, em que cada uma das entradas inclui um índice do livro de códigos de quantização e um vetor de quantização candidato do componente M correspondente. Cada um dos índices do livro de códigos de quantização pode corresponder a um respectivo de uma pluralidade de vetores de quantização candidatos do componente M.

[0111] O número de componentes em cada um dos vetores de quantização pode ser dependente do número de pesos (isto é, Z) que são selecionados para representar um único vetor V. Em geral, para um livro de códigos com vetores de quantização candidatos do componente Z, a unidade de quantização 52 pode realizar a quantização vetorial de valores de peso preditivos Z de cada vez para gerar um único vetor quantizado. O número de entradas no livro de códigos de quantização pode ser dependente da taxa de bits utilizada para realizar a quantização vetorial dos valores de peso.

[0112] Quando o vetor de unidade de quantização 52 quantiza o valor de peso preditivo, a unidade de quantização 52 pode selecionar um vetor do componente Z a partir do livro de códigos PVQ para ser o vetor de quantização que representa os valores de peso preditivos Z. O valor de peso preditivo quantizado pode ser denotado como $\hat{e}_{i,j}$, o qual pode corresponder ao j -ésimo componente do vetor de quantização do componente Z para o i -ésimo quadro de áudio, o qual pode ainda corresponder a uma versão de quantização vetorial do j -ésimo valor preditivo para o i -ésimo quadro de áudio.

[0113] Quando configurada para executar a quantização vetorial predita, a unidade de quantização 52 também pode gerar um valor de peso reconstruído com base no valor de peso preditivo quantizado e o valor de peso reconstruído ponderado. Por exemplo, a unidade de quantização 52 pode adicionar o valor de peso reconstruído ponderado para o valor de peso preditivo quantizado para gerar o valor de peso reconstruído. O valor de peso reconstruído ponderado pode ser idêntico ao valor de peso reconstruído ponderado, que é descrito acima. Em alguns exemplos, o valor de peso reconstruído ponderado pode ser uma versão ponderada e retardada do valor do peso reconstruído.

[0114] O valor de peso reconstruído pode ser representado como $|\hat{w}_{i-1,j}|$ que corresponde a uma grandeza (ou um valor absoluto) do valor de peso reconstruído correspondente, $\hat{w}_{i-1,j}$. O valor de peso reconstruído, $\hat{w}_{i-1,j}$, corresponde ao j -ésimo valor de peso reconstruído a partir de um subconjunto ordenado de valores de peso reconstruídos para o $(i-1)$ -ésimo quadro de áudio. Em alguns exemplos, a unidade de quantização 52 pode codificar separadamente os

dados indicativos do sinal de um valor do peso que é codificado de forma preditiva, e o decodificador pode usar esta informação para determinar o sinal do valor de peso reconstruído.

[0115] A unidade de quantização 52 pode gerar o valor de peso reconstruído com base na seguinte equação:

$$|\hat{w}_{i,j}| = \hat{e}_{i,j} + \alpha_j |\hat{w}_{i-1,j}|$$

em que $\hat{e}_{i,j}$ corresponde a um valor de peso preditivo quantizado para o j -ésimo valor de peso a partir de um subconjunto ordenado de valores de peso (por exemplo, o j -ésimo componente de um vetor de quantização do componente M) para o i -ésimo quadro de áudio, $|\hat{w}_{i-1,j}|$, corresponde a uma grandeza de um valor de peso reconstruído para o j -ésimo valor de peso a partir de um subconjunto ordenado de valores de peso para o $(i-1)$ -ésimo quadro de áudio, e α_j corresponde a um fator de ponderação para o j -ésimo valor de peso a partir de um subconjunto ordenado de valores de peso.

[0116] A unidade de quantização 52 pode gerar um valor de peso reconstruído retardado com base no valor de peso reconstruído. Por exemplo, a unidade de quantização 52 pode atrasar o valor de peso reconstruído por um quadro de áudio para gerar o valor de peso reconstruído retardado.

[0117] A unidade de quantização 52 também pode gerar o valor de peso reconstruído ponderado com base no valor de peso reconstruído retardado e o fator de ponderação. Por exemplo, a unidade de quantização 52 pode multiplicar o valor de peso reconstruído retardado pelo fator de ponderação para gerar o valor de peso reconstruído ponderado.

[0118] Da mesma forma, a unidade de quantização 52 gera o valor de peso reconstruído ponderado com base no valor de peso reconstruído retardado e o fator de ponderação. Por exemplo, a unidade de quantização 52 pode multiplicar o valor de peso reconstruído retardado pelo fator de ponderação para gerar o valor de peso reconstruído ponderado.

[0119] Em resposta à seleção de um vetor do componente Z a partir do livro de códigos PVQ ser um vetor de quantização para valores de peso preditivos Z, a unidade de quantização 52 pode, em alguns exemplos, codificar o índice (a partir do livro de códigos PVQ) que corresponde ao vetor do componente Z selecionado, em vez de codificar o próprio vetor do componente Z selecionado. O índice pode ser indicativo de um conjunto de valores de peso preditivos quantizados. Nestes exemplos, o decodificador 24 pode incluir um livro de códigos semelhante ao do livro de códigos PVQ, e pode decodificar o índice indicativo dos valores de peso preditivos quantizados mapeando o índice para um vetor do componente Z correspondente no livro de códigos do decodificador. Cada um dos componentes no vetor do componente Z pode corresponder a um valor de peso preditivo quantizado.

[0120] A quantização escalar de um vetor (por exemplo, um vetor V) pode envolver a quantização de cada um dos componentes do vetor de forma individual e / ou independente dos outros componentes. Por exemplo, considerar o seguinte exemplo do vetor V:

$$V = [0,23 \ 0,31 \ - \ 0,47 \ \dots \ 0,85]$$

Para quantização escalar deste vetor V exemplificativo, cada um dos componentes pode ser quantizado individualmente

(isto é, por quantização escalar). Por exemplo, se a etapa de quantização é de 0,1; então o componente 0,23 pode ser quantizado a 0,2; o componente 0,31 pode ser quantizado a 0,3; etc. Os componentes de quantização escalar podem coletivamente formar um vetor V de quantização escalar.

[0121] Em outras palavras, a unidade de quantização 52 pode executar quantização escalar uniforme no que diz respeito a todos os elementos do dado um dos vetores $V[k]$ de primeiro plano reduzidos 55. A unidade de quantização 52 pode identificar um tamanho de etapa de quantização com base em um valor, o qual pode ser indicado como um elemento de sintaxe N_{bitsQ} . A unidade de quantização 52 pode determinar dinamicamente este elemento de sintaxe N_{bitsQ} com base na taxa de bits alvo 41. O elemento de sintaxe N_{bitsQ} também pode identificar o modo de quantização como observado na tabela sintaxe $ChannelSidelInfoData$ reproduzida a seguir, ao mesmo tempo também identificando para fins de quantização escalar o tamanho da etapa. Isto é, a unidade de quantização 52 pode determinar o tamanho da etapa de quantização como uma função deste elemento de sintaxe N_{bitsQ} . Como um exemplo, a unidade de quantização 52 pode determinar o tamanho da etapa de quantização (indicado como "delta" ou " Δ " na presente descrição) como igual a $2^{16-N_{bitsQ}}$. Neste exemplo, quando o valor do elemento de sintaxe N_{bitsQ} é igual a seis, delta é igual a 2^{10} e existem 2^6 níveis de quantização. A este respeito, para um elemento de vetor v , o elemento de vetor quantizado v_q é igual a $[v/\Delta]$ e $-2^{N_{bitsQ}-1} < v_q < 2^{N_{bitsQ}-1}$.

[0122] A unidade de quantização 52 pode, em seguida, executar a categorização e codificação residual dos elementos de vetor quantizados. Como um exemplo, a unidade de quantização 52 pode, para um dado elemento de

vetor quantizado v_q identificar uma categoria (por determinação de um identificador de categoria cid) à qual corresponde este elemento utilizando a seguinte equação:

$$cid = \begin{cases} 0, & \text{if } v_q = 0 \\ \lfloor \log_2 |v_q| \rfloor + 1, & \text{if } v_q \neq 0 \end{cases}$$

A unidade de quantização 52 pode, em seguida, realizar a codificação de Huffman neste índice de categoria cid , ao mesmo tempo também identificando um bit de sinal que indica se v_q é um valor positivo ou um valor negativo. A unidade de quantização 52 pode a seguir identificar um resíduo nesta categoria. Como um exemplo, a unidade de quantização 52 pode determinar este resíduo de acordo com a seguinte equação:

$$resíduo = |v_q| - 2^{cid-1}$$

A unidade de quantização 52 pode então codificar por bloco este resíduo com $cid-1$ bits.

[0123] A unidade de quantização 52 pode, em alguns exemplos, selecionar diferentes tabelas de códigos de Huffman para diferentes valores do elemento de sintaxe $NbitsQ$ quando codificando o cid . Em alguns exemplos, a unidade de quantização 52 pode proporcionar uma tabela de codificação de Huffman diferente para valores de elementos de sintaxe $NbitsQ$ 6, ..., 15. Além disso, a unidade de quantização 52 pode incluir cinco diferentes tabelas de códigos de Huffman para cada um dos diferentes valores de elementos de sintaxe $NbitsQ$ que variam de 6, ..., 15 para um total de 50 livros de códigos de Huffman. A este respeito, a unidade de quantização 52 pode incluir uma pluralidade de diferentes livros de códigos de Huffman para acomodar a codificação do cid em uma série de diferentes contextos estatísticos.

[0124] Para ilustrar, a unidade de quantização 52 pode, para cada um dos valores de elementos de sintaxe N_{bitsQ} , incluir um primeiro livro de códigos de Huffman para codificar elementos de vetor de um a quatro, um segundo livro de códigos de Huffman para codificar elementos de vetor de cinco a nove anos, um terceiro livro de códigos de Huffman para codificar elementos de vetor nove e acima. Estas três primeiras tabelas de códigos de Huffman podem ser usadas quando o um dos vetores $V[k]$ de primeiro plano reduzidos 55 a serem comprimidos não é predito a partir de um vetor temporalmente subsequente correspondente dos vetores $V[k]$ de primeiro plano reduzidos 55 e não é representativo de informação espacial de um objeto de áudio sintético (um definido, por exemplo, originalmente por um objeto de áudio modulado por código de impulso (PCM)). A unidade de quantização 52 pode adicionalmente incluir, para cada um dos valores de elementos de sintaxe N_{bitsQ} , um quarto livro de códigos de Huffman para codificar o um dos vetores $V[k]$ de primeiro plano reduzidos 55 quando este um dos vetores $V[k]$ de primeiro plano reduzidos 55 é predito a partir de um vetor temporalmente subsequente correspondente dos vetores $V[k]$ de primeiro plano reduzidos 55. A unidade de quantização 52 também pode incluir, para cada um dos valores de elementos de sintaxe N_{bitsQ} , um quinto livro de códigos de Huffman para codificar o um dos vetores $V[k]$ de primeiro plano reduzidos 55 quando este um dos vetores $V[k]$ de primeiro plano reduzidos 55 é representativa de um objeto de áudio sintético. Os vários livros de códigos de Huffman podem ser desenvolvidos para cada um desses diferentes contextos estatísticos, isto é, o contexto não predito e não sintético, o contexto preditivo e o contexto sintético neste exemplo.

[0125] A tabela a seguir ilustra a seleção da tabela de Huffman e os bits a serem especificados no fluxo de bits para permitir que a unidade de descompressão selecione a tabela de Huffman apropriada:

Modo pred.	Info HT	Tabela HT
0	0	HT5
0	1	HT{1,2,3}
1	0	HT4
1	1	HT4

Na tabela anterior, o modo de predição ("Modo pred") indica se a predição foi realizada para o vetor corrente, enquanto a Tabela de Huffman ("Info HT") indica informações do livro de códigos de Huffman adicional (ou tabela) utilizadas para selecionar uma das tabelas de Huffman um a cinco. O modo de predição pode também ser representado como o elemento de sintaxe PFlag discutido abaixo, enquanto as informações da HT podem ser representadas pelo elemento de sintaxe CbFlag discutido abaixo.

[0126] A tabela a seguir ilustra adicionalmente esse processo de seleção da tabela de Huffman dado vários contextos ou cenários estatísticos.

	Gravação	Sintético
Sem Pred	HT{1,2,3}	HT5
Com Pred	HT4	HT5

Na tabela anterior, a coluna "Registro" indica o contexto de codificação quando o vetor é representativo de um objeto de áudio que foi gravado, ao passo que a coluna "Sintética" indica um contexto de codificação para quando o vetor é representativo de um objecto de áudio sintético. A linha "Sem Pred" linha indica o contexto de codificação

quando a previsão não é executada no que diz respeito aos elementos de vetor, ao passo que a linha "Com Pred" indica o contexto de codificação quando a predição é realizada com respeito aos elementos de vetor. Como se mostra nesta tabela, a unidade de quantização 52 seleciona HT{1,2,3} quando o vetor é representativo de um objeto de áudio gravado e a predição não é executada no que diz respeito aos elementos de vetor. A unidade de quantização 52 seleciona HT5 quando o objeto de áudio é representativo de um objeto de áudio sintético e a predição não é executada no que diz respeito aos elementos de vetor. A unidade de quantização 52 seleciona HT4 quando o vetor é representativo de um objeto de áudio gravado e a predição é realizada com respeito aos elementos de vetor. A unidade de quantização 52 seleciona HT5 quando o objeto de áudio é representativo de um objeto de áudio sintético e a predição é realizada com respeito aos elementos de vetor.

[0127] A unidade de quantização 52 pode selecionar um dos vetor V de quantização vetorial não predito, vetor V de quantização vetorial predito, vetor V de quantização escalar de não codificação de Huffman e o vetor V de quantização escalar de codificação de Huffman para usar como o vetor V quantizado-comutado de saída com base em qualquer combinação dos critérios discutidos nesta invenção. Em alguns exemplos, a unidade de quantização 52 pode selecionar um modo de quantização a partir de um conjunto de modos de quantização que inclui um modo de quantização vetorial e um ou mais modos de quantização escalar, e quantizar um vetor V de entrada com base em (ou segundo) o modo selecionado. A unidade de quantização 52 pode, em seguida, fornecer o um vetor selecionado do vetor N de quantização vetorial não predito (por exemplo, em termos de valores de peso ou seus bits indicativos), o

vetor V de quantização vetorial predito (por exemplo, em termos de valores de erro ou seus bits indicativos), o vetor V de quantização escalar de não codificação de Huffman e o vetor V de quantização escalar de codificação de Huffman para a unidade de geração de fluxo de bits 52 como os vetores $V[k]$ de primeiro plano codificados 57. A unidade de quantização 52 também pode fornecer os elementos de sintaxe indicativos do modo de quantização (por exemplo, o elemento de sintaxe NbitsQ) e quaisquer outros elementos de sintaxe usados para dequantizar ou de outra forma reconstruir o vetor V tal como discutido em mais detalhes abaixo em relação ao exemplo das Figs. 4 e 7.

[0128] A unidade de codificador de áudio psicoacústica 40 incluída dentro do dispositivo de codificação de áudio 20 pode representar vários casos de um codificador de áudio psicoacústico, cada um deles é usado para codificar um objeto de áudio diferente ou um canal de HOA de cada um dos coeficientes de HOA de ambiente compensado de energia 47' e os sinais nFG interpolados 49' para gerar coeficientes de HOA de ambiente codificado 59 e sinais nFG codificados 61. A unidade de codificador de áudio psicoacústica 40 pode emitir os coeficientes de HOA de ambiente codificado 59 e os sinais nFG codificados 61 para a unidade de geração de fluxo de bits 42.

[0129] A unidade de geração de fluxo de bits 42 incluída dentro do dispositivo de codificação de áudio 20 representa uma unidade que formata os dados para estar em conformidade com um formato conhecido (que pode referir-se a um formato conhecido por um dispositivo de decodificação), gerando deste modo o fluxo de bits com base em vetor 21. O fluxo de bits 21 pode, em outras palavras, representar dados de áudio codificados, tendo sido codificados do modo descrito acima. A unidade de geração de

fluxo de bits 42 pode representar um multiplexador em alguns exemplos, os quais podem receber os vetores $V[k]$ de primeiro plano codificados 57, os coeficientes de HOA de ambiente codificado 59, os sinais nFG codificados 61 e as informações de canal de base 43. A unidade de geração de fluxo de bits 42 pode, em seguida, gerar um fluxo de bits 21 com base nos vetores $V[k]$ de primeiro plano codificados 57, os coeficientes de HOA de ambiente codificado 59, os sinais nFG codificados 61 e as informações de canal de base 43. Deste modo, a unidade de geração de fluxo de bits 42 pode, assim, especificar os vetores 57 no fluxo de bits 21 para obter o fluxo de bits 21 como descrito abaixo em mais detalhes no que diz respeito ao exemplo da FIG. 7. O fluxo de bits 21 pode incluir um fluxo de bits primário ou principal e um ou mais fluxos de bits de canal secundários.

[0130] Embora não esteja representado no exemplo da FIG. 3, o dispositivo de codificação de áudio 20 pode também incluir uma unidade de saída de fluxo de bits que comuta a saída do fluxo de bits a partir do dispositivo de codificação de áudio 20 (por exemplo, entre o fluxo de bits de base direcional 21 e o fluxo de bits de base vetorial 21) com base em se um quadro corrente deve ser codificado usando a síntese de base direcional ou a síntese de base vetorial. A unidade de saída de fluxo de bits pode realizar a comutação com base no elemento de sintaxe emitido pela unidade de análise de conteúdo 26 indicando se foi realizada uma síntese de base direcional (como um resultado da detecção de que os coeficientes de HOA 11 foram gerados a partir de um objeto de áudio sintético) ou se foi realizada uma síntese de base vetorial (como um resultado da detecção de que os coeficientes de HOA foram registados). A unidade de saída de fluxo de bits pode especificar a sintaxe correta do cabeçalho para indicar a comutação ou

codificação corrente usada para o quadro corrente, juntamente com o respectivo um dos fluxos de bits 21.

[0131] Além disso, como mencionado acima, a unidade de análise de campo sonoro 44 pode identificar coeficientes de HOA de ambiente BG_{TOT} 47, que podem mudar em uma base de quadro-a-quadro (embora, por vezes, BG_{TOT} possa permanecer constante ou o mesmo em dois ou mais quadros adjacentes (no tempo)). A mudança em BG_{TOT} pode resultar em alterações para os coeficientes expressos nos vetores $V[k]$ de primeiro plano reduzidos 55. A mudança em BG_{TOT} pode resultar em coeficientes de HOA de base (que também podem ser referidos como coeficientes de HOA "ambiente") que mudam em uma base quadro-a-quadro (embora, novamente, às vezes BG_{TOT} possa permanecer constante ou o mesmo em dois ou mais quadros adjacentes (no tempo)). As alterações resultam muitas vezes em uma variação de energia para os aspectos do campo de som representado pela adição ou remoção dos coeficientes de HOA de ambiente adicionais e a correspondente remoção de coeficientes a partir de ou além de coeficientes para os vetores $V[k]$ de primeiro plano reduzidos 55.

[0132] Como resultado, a unidade de análise do campo de som 44 pode ainda determinar quando os coeficientes de HOA de ambiente mudam de quadro a quadro e geram uma bandeira ou outro elemento de sintaxe indicativo da mudança para o coeficiente de HOA ambiente em termos de sendo utilizado para representar os componentes do ambiente do campo de som (em que a alteração pode também ser referida como uma "transição" do coeficiente de HOA ambiente ou como uma "transição" do coeficiente de HOA ambiente). Em particular, a unidade de redução do coeficiente 46 pode gerar a bandeira (que pode ser indicada como uma bandeira `AmbCoeffTransition` ou uma bandeira

AmbCoeffIdxTransition), proporcionando a bandeira para a unidade de geração de fluxo de bits 42, de modo que a bandeira pode ser incluída no fluxo de bits 21 (possivelmente como parte da informação do canal secundário).

[0133] A unidade de redução de coeficiente 46 pode, além de especificar a bandeira de transição de coeficiente de ambiente, também modificar a forma como são gerados os vetores $V[k]$ de primeiro plano reduzidos 55. Em um exemplo, após a determinação de que um dos coeficientes ambiente HOA ambiente está em transição durante o quadro corrente, a unidade de redução do coeficiente 46 pode especificar, um coeficiente de vetor (que também pode ser referido como um "elemento de vetor" ou "elemento") para cada um dos vetores V dos vetores $V[k]$ de primeiro plano reduzidos 55 que corresponde ao coeficiente de HOA ambiente em transição. Mais uma vez, o coeficiente de HOA ambiente em transição pode adicionar ou remover a partir do número total BG_{TOT} de coeficientes de base. Por conseguinte, a alteração resultante no número total de coeficientes de base afeta se o coeficiente de HOA ambiente está incluído ou não incluído no fluxo de bits, e se o elemento correspondente dos vetores V está incluído para os vetores V especificados no fluxo de bits nos segundo e terceiro modos de configuração descritos acima. Mais informações sobre a forma como a unidade de redução de coeficiente 46 pode especificar os vetores $V[k]$ de primeiro plano reduzidos 55 para superar as mudanças em energia são fornecidas no pedido US n° de Série 14/594.533, intitulado "TRANSITIONING OF AMBIENT HIGHER_ORDER AMBISONIC COEFFICIENTS", depositado em 12 de janeiro de 2015.

[0134] A FIG. 4 é um diagrama em blocos que ilustra o dispositivo de decodificação de áudio 24 da FIG.

2 em maiores detalhes. Como mostrado no exemplo da FIG. 4, o dispositivo de decodificação de áudio 24 pode incluir uma unidade de extração 72, uma unidade de reconstrução à base de direcionalidade 90 e uma unidade de reconstrução de base vetorial 92. Embora descrito abaixo, mais informações sobre o dispositivo de decodificação de áudio 24 e os vários aspectos da descompressão ou outra forma de decodificação dos coeficientes de HOA estão disponíveis na Publicação do Pedido de Patente Internacional n° WO 2014/194099, intitulado "INTERPOLATION FOR DECOMPOSED REPRESENTATIONS OF A SOUND FIELD", depositado em 29 de maio de 2014.

[0135] A unidade de extração 72 pode representar uma unidade configurada para receber o fluxo de bits 21 e extrair as várias versões codificadas (por exemplo, uma versão codificada de base direcional ou uma versão codificada de base vetorial) dos coeficientes de HOA 11. A unidade de extração 72 pode determinar a partir do elemento de sintaxe acima referido indicativo de se os coeficientes de HOA 11 foram codificados por meio das diversas versões de base direcional ou de base vetorial. Quando uma codificação de base direcional foi realizada, a unidade de extração 72 pode extrair a versão de base direcional dos coeficientes de HOA 11 e os elementos de sintaxe associados com a versão codificada (que são designados como informações de base direcional 91 no exemplo da FIG. 4), que passa as informações de base direcional 91 para a unidade de reconstrução de base direcional 90. A unidade de reconstrução de base direcional 90 pode representar uma unidade configurada para reconstruir os coeficientes de HOA sob a forma de coeficientes de HOA 11' com base nas informações de base direcional 91. O fluxo de bits e o arranjo dos elementos de sintaxe dentro do fluxo de bits

são descritos abaixo em maiores detalhes com relação ao exemplo das Figs. 7A-7J.

[0136] Quando o elemento de sintaxe indica que os coeficientes de HOA 11 foram codificados utilizando uma síntese de base vetorial, a unidade de extração 72 pode extrair os vetores $V[k]$ de primeiro plano codificados 57 (que podem incluir pesos codificados 57 e / ou índices 63 ou vetores V de quantização escalar), os coeficientes de HOA de ambiente codificados 59 e os objetos de áudio correspondentes 61 (que também podem ser referidos como os sinais nFG codificados 61). Os objetos de áudio 61 correspondem, cada um, a um dos vetores 57. A unidade de extração 72 pode passar os vetores $V[k]$ de primeiro plano codificados 57 para a unidade de reconstrução do vetor V 74 e os coeficientes de HOA de ambiente codificados 59, juntamente com os sinais nFG codificados 61 para a unidade de decodificação psicoacústica 80.

[0137] Para extrair os vetores $V[k]$ de primeiro plano codificados 57, a unidade de extração 72 pode extrair os elementos de sintaxe de acordo com a tabela de sintaxe ChannelSideInfoData (CSID) a seguir.

TABELA - SINTAXE DE CHANNELSIDEINFODATA (I)

Sintaxe	Nº de bits	Mnemonic
ChannelSideInfoData(i)		
{		
ChannelType[i]	2	uimsbf
comutar ChannelType[i]		
{		
caso 0:		
ActiveDirsllds[i];	NumOfBits PerDirldx	uimsbf
pausa;		

caso 1:		
bA	1	bslbf
bB;	1	bslbf
se ((bA + bB) == 0){		
NbitsQ(k)[i] = NbitsQ(k-		
1)[i];		
PFlag(k)[i] = PFlag(k-1)[i];		
CbFlag(k)[i] = Codebkldx(k-		
1)[i];		
Codebkldx(k)[i] =		
Codebkldx(k-1)[i];		
NumVecIndices(k)[i] =		
NumVecIndices[k-1][i];		
}		
else{		
NbitsQ(k)[i] =	2	uimsbf
(8*bA)+(4*bB)+uintC;		
se (NbitsQ(k)[i] == 4){		
PFlag(k)[i];	1	bslbf
Codebkldx(k)[i];	3	uimsbf
NumVecIndices(k)[i]++;	NumVVec	uimsbf
	VqElemen	
	tsBits	
}		
elseif (NbitsQ(k)[i];>=6){		
PFlag(k)[i];	1	bslbf
CbFlag(k)[i];	1	bslbf
}		
}		
}		
pausa;		
caso 2:		

AddAmbHoaInfoChannel(i);		
pausa;		
padrão:		
}		
}		

[0138] A semântica para a tabela anterior é como a seguir. Esta carga útil contém as informações secundárias para o i -ésimo canal. O tamanho e os dados da carga útil dependem do tipo de canal.

ChannelType[i]: este elemento armazena o tipo do i -ésimo canal que é definido na Tabela 95.

ActiveDirsIds[i]: este elemento indica a direção do sinal direcional ativo usando um índice dos 900 pontos predefinidos distribuídos uniformemente do Anexo F.7. A palavra de código 0 é usada para sinalizar o fim de um sinal direcional.

PFLAG[i]: a bandeira predição ~~usada para a decodificação de Huffman do vetor V de quantização escalar~~ associada com o sinal de base vetorial do i -ésimo canal.

CbFlag[i]: a bandeira do livro de códigos usada para a decodificação de Huffman do vetor V de quantização escalar associado com o sinal de base vetorial do i -ésimo canal.

CodebkIdx[i]: sinaliza o livro de códigos específico utilizado para dequantizar o vetor V de quantização vetorial associado com o sinal de base vetorial do i -ésimo canal.

NbitsQ[i]: este índice determina a tabela de Huffman utilizada para a decodificação de Huffman dos dados associados com o sinal de base vetorial do i -ésimo canal. A palavra de código 5 determina a utilização de um dequantizador de 8 bits uniforme. Os dois MSBs 00

determinam a reutilização dos dados $NbitsQ[i]$, $PFlag[i]$ e $CbFlag[i]$ do quadro anterior ($k-1$).

bA ou bB : o msb (bA) e o segundo msb (bB) do campo $NbitsQ[i]$.

$uintC$: a palavra de código dos restantes dois bits do campo $NbitsQ[i]$.

$NumVecIndices$: o número de vetores utilizados para dequantizar um vetor V de quantização vetorial.

$AddAmbHoaInfoChannel(i)$: esta carga útil contém as informações para coeficientes de HOA de ambiente adicionais.

[0139] De acordo com a tabela de sintaxe CSID, a unidade de extração 72 pode primeiro obter um elemento de sintaxe $ChannelType$ indicativo do tipo de canal (por exemplo, onde um valor de zero sinaliza um sinal de base direcional, um valor de 1 sinaliza um sinal de base vetorial e um valor de 2 sinaliza um sinal HOA ambiente adicional). Com base no elemento de sintaxe $ChannelType$, a unidade de extração 72 pode alternar entre os três casos.

[0140] Concentrando no caso 1 para ilustrar um exemplo das técnicas descritas nesta invenção, a unidade de extração 72 pode obter o bit mais significativo do elemento de sintaxe $NbitsQ$ (isto é, o elemento de sintaxe bA na tabela sintaxe CSID exemplificativa acima) e o segundo bit mais significativo do elemento de sintaxe $NbitsQ$ (isto é, o elemento de sintaxe bB na tabela sintaxe CSID exemplificativa acima). O $(k)[i]$ de $NbitsQ(k)[i]$ pode indicar que o elemento de sintaxe $NbitsQ$ é obtido para o k -ésimo quadro do i -ésimo canal de transporte. O elemento de sintaxe $NbitsQ$ pode representar um ou mais bits indicativos de um modo de quantização utilizado para quantizar o componente espacial do campo sonoro representado pelos coeficientes de HOA 11. O componente espacial também pode

ser referido como um vetor V na presente descrição ou como os vetores de primeiro plano codificados $V[K]$ 57.

[0141] Na tabela de sintaxe CSID exemplificativa acima, o elemento de sintaxe NbitsQ pode incluir quatro bits para indicar um dos 12 modos de quantização (visto que um valor de zero a três para o elemento de sintaxe NbitsQ é reservado ou não utilizado) utilizados para comprimir o vetor especificado no campo VVecData correspondente. Os 12 modos de quantização incluem o seguinte indicado abaixo:

- 0-3: Reservado
- 4: Quantização Vetorial
- 5: Quantização Escalar sem Codificação de Huffman
- 6: Quantização Escalar de 6 bits com Codificação de Huffman
- 7: Quantização Escalar de 7 bits com Codificação de Huffman
- 8: Quantização Escalar de 8 bits com Codificação de Huffman
-
- 16: Quantização Escalar de 16 bits com Codificação de Huffman

No exemplo acima, o valor do elemento de sintaxe NbitsQ de 6-16 indica não apenas que a quantização escalar deve ser realizada com a codificação de Huffman, mas também o tamanho da etapa de quantização da quantização escalar. A este respeito, o modo de quantização pode compreender um modo de quantização vetorial, um modo de quantização escalar sem codificação de Huffman e um modo de quantização escalar com codificação de Huffman.

[0142] Voltando à tabela de sintaxe CSID exemplificativa acima, a unidade de extração 72 pode combinar o elemento de sintaxe bA com o elemento de sintaxe bB, em que esta combinação pode ser uma adição, como

mostrado na tabela sintaxe CSID exemplificativa acima. O elemento de sintaxe bA / bB combinado pode representar um indicador quanto à reutilização, a partir de um quadro anterior, de pelo menos um elemento de sintaxe indicativo de informações utilizadas quando se comprime o vetor. A unidade de extração 72 seguinte compara o elemento de sintaxe bA / bB combinado a um valor de zero. Quando o elemento de sintaxe bA / bB combinado tem um valor de zero, a unidade de extração 72 pode determinar que as informações do modo de quantização para o k-ésimo quadro corrente do i-ésimo canal de transporte (isto é, o elemento de sintaxe NbitsQ indicativo do modo de quantização na tabela de sintaxe CSID exemplificativa acima) são as mesmas que as informações do modo de quantização do k-1-ésimo quadro do i-ésimo canal de transporte. Em outras palavras, o indicador, quando configurado a um valor de zero, indica reutilizar o, pelo menos, um elemento de sintaxe do quadro anterior.

[0143] A unidade de extração 72 de modo semelhante determina que as informações de predição para o k-ésimo quadro corrente do i-ésimo canal de transporte (isto é, o elemento de sintaxe PFlag indicativo de se a predição é realizada durante ou quantização vetorial ou quantização escalar no exemplo) são as mesmas que as informações de predição do k-1-ésimo quadro do i-ésimo canal de transporte. A unidade de extração 72 pode também determinar que as informações do livro de códigos de Huffman para o quadro k-ésimo corrente do i-ésimo canal de transporte (isto é, o elemento de sintaxe CbFlag indicativo de um livro de códigos de Huffman utilizado para reconstruir o vetor V) são as mesmas que as informações do livro de códigos de Huffman do quadro k-1-ésimo do i-ésimo canal de transporte. A unidade de extração 72 pode também

determinar que as informações de quantização vetorial para o k-ésimo quadro corrente do i-ésimo canal de transporte (isto é, o elemento de sintaxe `CodebkIdx` indicativo de um livro de códigos de quantização vetorial utilizado para reconstruir o vetor V e o elemento de sintaxe `NumVecIndices` indicativo de uma série de vetores de código utilizados para reconstruir o vetor V) são as mesmas que as informações de quantização vetorial do k-1-ésimo quadro do i-ésimo canal de transporte.

[0144] Quando o elemento de sintaxe `bA / bB` combinado não tem um valor de zero, a unidade de extração 72 pode determinar que as informações do modo de quantização, as informações de predição, as informações do livro de códigos de Huffman e as informações de quantização vetorial para o k-ésimo quadro do i-ésimo canal de transporte não são as mesmas que as do k-1-ésimo quadro do i-ésimo canal de transporte. Como resultado, a unidade de extração 72 pode obter os bits menos significativos do elemento de sintaxe `NbitsQ` (ou seja, o elemento de sintaxe `uintC` na tabela de sintaxe `CSID` exemplificativa acima), combinando os elementos de sintaxe `bA`, `bB` e `uintC` para obter o elemento de sintaxe `NbitsQ`. Com base neste elemento de sintaxe `NbitsQ`, a unidade de extração 72 pode obter ou, quando o elemento de sintaxe `NbitsQ` sinaliza quantização vetorial, os elementos de sintaxe `PFlag`, `CodebkIdx` e `NumVecIndices`; ou, quando o elemento de sintaxe `NbitsQ` sinaliza quantização escalar com codificação de Huffman, os elementos de sintaxe `PFlag` e `CbFlag`. Desta forma, a unidade de extração 72 pode extrair os elementos de sintaxe precedentes utilizados para reconstruir o vetor V , passando estes elementos de sintaxe para a unidade de reconstrução de base vetorial 72.

[0145] A unidade de extração 72 pode a seguir extrair o vetor V do k-ésimo quadro do i-ésimo canal de transporte. A unidade de extração 72 pode obter um recipiente HOADecoderConfig, que inclui o elemento de sintaxe denotado CodedVVecLength. A unidade de extração 72 pode analisar o CodedVVecLength do recipiente HOADecoderConfig. A unidade de extração 72 pode obter o vetor V de acordo com a seguinte tabela de sintaxe VVecData.

Sintaxe	Nº de bits	Mnemonic
VVectorData(i)		
{		
se (NbitsQ(k)[i] == 4) {		
se (NumVecIndices(k)[i] == 1) {		
VecIdx[0] = VecIdx + 1;	10	uimsbf
WeightVal[0] = ((SgnVal*2)-1);	1	uimsbf
} else {		
WeightIdx	nbitsW	uimsbf
nbitsIdx =		
ceil(log2(NumOfHoaCoeffs));		
para (j=0;		
j<NumVecIndices(k)[i]; ++j) {		
VecIdx[j] = VecIdx + 1;	nbitsIdx	uimsbf
se (PFlag[i] == 0) {		
tmpWeightVal(k)[j] =		
WeightValCdbk[CodebkIdx(k)[i]][WeightIdx][j];		
else {		
tmpWeightVal(k)[j] =		
WeightValPredCdbk[CodebkIdx(k)[i]][Wei		

ghtIdx][j] + WeightValAlpha[j]*tmpWeightVal(k- 1)[j];		
}		
WeightVal[j] = ((SgnVal*2)-1)*	1	uimsbf
}		
}		
}		
elseif (NbitsQ(k)[i] == 5){		
para (m=0; m< VVecLength; ++m)		
aVal[i][m] = (VecVal / 128.0)-1,0;	8	uimsbf
}		
elseif (NbitsQ(k)[i] >= 6){		
para (m=0; m< VVecLength; ++m){		
huffIdx = huffSelect(VVecCoeffId[m], PFlag[i], CbFlag[i]);		
cid = huffDecode(NbitsQ[i], huffIdx, huffVal);	dinâmico	huffDecode
aVal[i][m] = 0,0;		
se (cid > 0){		
aVal[i][m] = sgn = (SgnVal*2) - 1;	1	uimsbf
se (cid > 1){		
aVal[i][m] = sgn*(2.0^(cid-1) + intAddVal);	cid-1	uimsbf
}		
}		
}		
}		
Nota: Veja Erro! Fonte de referência		

não encontrada para computação de VVecLength		
--	--	--

VVec(k)[i]: este é o vetor V para o k-ésimo HOAframe() para o i-ésimo canal.

VVecLength: esta variável indica o número de elementos de vetor para ler.

VVecCoeffId: este vetor contém os índices dos coeficientes do vetor V transmitidos.

VecVal: um valor inteiro entre 0 e 255.

Aval: uma variável temporária usada durante a decodificação do VVetorData.

huffVal: uma palavra de código de Huffman a ser decodificada por Huffman.

SgnVal: este é o valor do sinal codificado usado durante a decodificação.

intAddVal: este é o valor inteiro adicional utilizado durante a decodificação.

NumVecIndices: o número de vetores usados para dequantizar um vetor V quantizado.

WeightIdx: o índice em WeightValCdbk usado para dequantizar um vetor V de quantização vetorial.

nBitsW: tamanho do campo para ler WeightIdx para decodificar um vetor V de quantização vetorial.

WeightValCbk: livro de códigos que contém um vetor de coeficientes de ponderação com valores reais positivos. Apenas necessário se NumVecIndices é > 1. O WeightValCdbk com 256 entradas é fornecido.

WeightValPredCdbk: livro de códigos que contém um vetor de coeficientes de ponderação preditiva. Apenas necessário se NumVecIndices é > 1. O WeightValPredCdbk com 256 entradas é fornecido.

WeightValAlpha: coeficientes de codificação preditiva que são utilizados para o modo de codificação preditiva da quantização do vetor V.

VvecIdx: Um índice para VecDict, usado para dequantizar um vetor V por quantização vetorial.

nbitsIdx: tamanho do campo para ler VvecIdx para decodificar um vetor V de quantização vetorial.

WeightVal: um coeficiente de ponderação de valor real para decodificar uma vetor V de quantização vetorial.

[0146] Na tabela de sintaxe que precede, a unidade de extração 72 pode determinar se o valor do elemento de sintaxe NbitsQ é igual a quatro (ou, em outras palavras, sinaliza que a dequantização vetorial é usada para reconstruir o vetor V). Quando o valor do elemento de sintaxe NbitsQ é igual a quatro, a unidade de extração 72 pode comparar o valor do elemento de sintaxe NumVecIndices a um valor de um. Quando o valor do NumVecIndices é igual a um, a unidade de extração 72 pode obter um elemento de sintaxe VecIdx. O elemento de sintaxe VecIdx pode representar um ou mais bits indicativos de um índice para uma VecDict usado para dequantizar um vetor V de quantização vetorial. A unidade de extração 72 pode justificar uma matriz VecIdx, com o zero-ésimo elemento ajustado ao valor do elemento de sintaxe VecIdx mais um. A unidade de extração 72 também pode obter um elemento de sintaxe SgnVal. O elemento de sintaxe SgnVal pode representar um ou mais bits indicativos de um valor de sinal codificado usado durante a decodificação do vetor V. A unidade de extração 72 pode justificar uma matriz WeightVal, definindo o zero-ésimo elemento como uma função do valor do elemento de sintaxe SgnVal.

[0147] Quando o valor do elemento de sintaxe NumVecIndices não é igual a um valor de um, a unidade de

extração 72 pode obter um elemento de sintaxe `WeightIdx`. O elemento de sintaxe `WeightIdx` pode representar um ou mais bits indicativos de um índice na matriz `WeightValCdbk` usada para dequantizar um vetor `V` de quantização vetorial. A matriz `WeightValCdbk` pode representar um livro de códigos que contém um vetor de coeficientes de ponderação de valores reais positivos. A unidade de extração 72 pode, em seguida, determinar um `nbitsIdx` como uma função de um elemento de sintaxe `NumOfHoaCoeffs` especificado no recipiente `HOAConfig` (especificado como um exemplo, no início do fluxo de bits 21). A unidade de extração 72 pode então percorrer o `NumVecIndices`, obter um elemento de sintaxe `VecIdx` a partir do fluxo de bits 21 e definir os elementos da matriz `VecIdx` com cada elemento de sintaxe `VecIdx` obtido.

[0148] A unidade de extração 72 não executa a comparação de sintaxe de `PFLAG` a seguir, que envolve determinar os valores das variáveis `tmpWeightVal` que não estão relacionados com a extração de elementos de sintaxe a partir do fluxo de bits 21. Como tal, a unidade de extração 72 pode, em seguida, obter o elemento de sintaxe `SgnVal` para usar na determinação de um elemento de sintaxe `WeightVal`.

[0149] Quando o valor do elemento de sintaxe `NbitsQ` é igual a cinco (sinalizando que a dequantização escalar sem decodificação de Huffman é usada para reconstruir o vetor `V`), a unidade de extração 72 repete de 0 ao `VVecLength`, ajustando a variável `Aval` ao elemento de sintaxe `VecVal` obtido a partir do fluxo de bits 21. O elemento de sintaxe `VecVal` pode representar um ou mais bits indicativos de um número inteiro entre 0 e 255.

[0150] Quando o valor do elemento de sintaxe `NbitsQ` é igual a ou maior do que seis (sinalizar que a

dequantização escalar de NbitsQ-bit com decodificação de Huffman é usada para reconstruir o vetor V), a unidade de extração 72 repete de 0 ao VVecLength, obtendo um ou mais dos elementos de sintaxe huffVal, SgnVal e intAddVal. O elemento de sintaxe huffVal pode representar um ou mais bits indicativos de uma palavra de código Huffman. O elemento de sintaxe intAddVal pode representar um ou mais bits indicativos de valores inteiros adicionais usados durante a decodificação. A unidade de extração 72 pode fornecer esses elementos de sintaxe para a unidade de reconstrução de base vetorial 92.

[0151] A unidade de reconstrução baseada em vetor 92 pode representar uma unidade configurada para executar operações recíprocas àquelas descritas acima em relação à unidade de síntese baseada em vetor 27, de modo a reconstruir os coeficientes de HOA 11'. A unidade de reconstrução baseada em vetor 92 pode incluir uma unidade de reconstrução do vetor V 74, uma unidade de interpolação espaço-temporal 76, uma unidade de formulação de primeiro plano 78, uma unidade de decodificação psicoacústica 80, uma unidade de formulação de coeficiente de HOA 82, uma unidade de fade 770 e uma unidade de reordenamento 84. As linhas tracejadas da unidade de fade 770 indicam que a unidade de fade 770 pode ser uma unidade opcional em termos de ser incluída na unidade de reconstrução baseada em vetor 92.

[0152] A unidade de reconstrução do vetor V 74 pode representar uma unidade configurada para reconstruir os vetores V a partir dos vetores V[K] de primeiro plano codificados 57. A unidade de reconstrução do vetor V 74 pode operar de um modo recíproco ao da unidade de quantização 52.

[0153] A unidade de reconstrução do vetor V 74 pode, em outras palavras, operar de acordo com o seguinte pseudocódigo para reconstruir os vetores V:

```

se (NbitsQ(k)[i] == 4){
  se (NumVvecIndicies == 1){
    para (m = 0; m< VVecLength; m++){
      idx = VVecCoeffID[m];
       $v^{(i)}_{VVeccoeffId[m]}(K) = WeightVal[0]*VecDict[900].$ 
      [VecIdx[0]][idx];
    }
  } ou {
    cdbLen = 0;
    se (N == 4)
      cdbLen = 32;
    se
      para (M = 0; M< 0; ++m){
        TmpVVec[M] = 0;
        para (j=0; j< NumVecIndecies; ++j){
          TmpVVec[m] +=
eightVal[j]*VecDict[cdbLen].[VecIdx[j]][m];
        }
      }
      FNorm = 0,0;
      para (m=0; M < 0; ++ m) {
        FNorm += TmpVVecfm]*TmpVVecfm];
      }
      FNorm = (N+1) / sqrt (FNorm);
      para (m = 0; m< VVecLength; ++m){
        idx = VVecCoeffID[m];
         $v^{(i)}_{VVecCoeffId[m]}(k) = TmpVVec[idx]*FNorm;$ 
      }
    }
  }
}

```

```

elseif (NbbitsQ(k)[i] == 5){
    para (m=0; m< VVecLength; ++m){
         $v_{VVecCoeffId[m]}^{(i)}(k) = (N+1) * aVal[i][m];$ 
    }
}
elseif (NbbitsQ(k)[i] >= 6){
    para (m=0; m< VVecLength; ++m){
         $v_{VVecCoeffId[m]}^{(i)}(k) = (N+1) * (2^{(16 -$ 
NbbitsQ(k)[i]) * aVval[i][m]) / 2^15;
        if (PFlag(k)[i] == 1){
             $v_{VVecCoeffId[m]}^{(i)}(k) += v_{VVecCoeffId[m]}^{(i)}(k - 1);$ 
        }
    }
}

```

[0154] De acordo com o pseudocódigo que precede, a unidade de reconstrução do vetor V 74 pode obter o elemento de sintaxe NbbitsQ para o k-ésimo quadro do i-ésimo canal de transporte. Quando o elemento de sintaxe NbbitsQ é igual a quatro (que, mais uma vez, sinaliza que a quantização vetorial foi realizada), a unidade de reconstrução do vetor V 74 pode comparar os elementos de sintaxe NumVecIndicies a um. O elemento de sintaxe NumVecIndicies pode, como acima descrito, representar um ou mais bits indicativos de uma série de vetores utilizados para dequantizar um vetor V de quantização vetorial. Quando o valor do elemento de sintaxe NumVecIndicies é igual a um, a unidade de reconstrução do vetor V 74 pode, em seguida, repetir de zero até o valor do elemento de sintaxe VVecLength, ajustando a variável idx ao VVecCoeffId e o VVecCoeffId-ésimo elemento do vetor V ($v_{VVecCoeffId[m]}^{(i)}(k)$) para o WeightVal multiplicado pela entrada VecDict identificada pelo [900] [VecIdx[0]][idx]. Em outras palavras, quando o valor de NumVvecIndicies é igual a um, os coeficientes de expansão

HOA do livro de códigos do Vetor derivados da tabela F.8 em conjunto com um livro de códigos de 8x1 valores de ponderação mostrados na tabela F.11.

[0155] Quando o valor do elemento de sintaxe NumVecIndicies não é igual a um, a unidade de reconstrução do vetor V 74 pode definir a variável cdbLen a 0, que é uma variável que denota o número de vetores. O elemento de sintaxe cdbLen indica o número de entradas no dicionário ou livro de códigos de vetores de código (onde este dicionário é denotado como "VecDict" no pseudocódigo anterior e representa um livro de códigos com entradas do livro de códigos cdbLen contendo vetores de coeficientes de expansão HOA, usados para decodificar um vetor de de quantização vetorial). Quando o pedido (indicado por "N") dos coeficientes de HOA 11 é igual a quatro, a unidade de reconstrução do vetor V 74 pode definir a variável cdbLen a 32. A unidade de reconstrução do vetor V 74 pode, em seguida, repetir de zero a 0, definindo um matriz TmpVVec a zero. Durante estas repetições, a unidade de reconstrução do vetor V 74 também pode repetir de zero ao valor do elemento de sintaxe NumVecIndecies, definindo a m^{th} entrada da matriz TempVVec para ser igual ao j -ésimo WeightVal multiplicado pela entrada [cdbLen][VecIdx[j]][m] do VecDict.

[0156] A unidade de reconstrução do vetor V 74 pode derivar do WeightVal de acordo com o seguinte pseudocódigo:

```

para (j=0; j< NumVecIndices(k)[i]; ++j){
    se (PFlag[i] == 0){
        tmpWeightVal(k)[J] =
WeightValCdbk[CodebkIdx(k)[i]][WeightIdx][j];
    else {

```

```

        tmpWeightVal(k)[J] =
WeightValPredCdbk[CodebkIdx(k)[i]][WeightIdx][j];
    + WeightValAlpha[j]*tmpWeightVal(k-1)[j];
    }
WeightVal[j] = ((SgnVal*2)-1)*tmpWeightVal(k)[j];

```

No pseudocódigo que precede, a unidade de reconstrução do vetor V 74 pode repetir de zero até o valor do elemento de sintaxe NumVecIndices, primeiro determinando se o valor do elemento de sintaxe PFlag é igual a zero. Quando o elemento de sintaxe PFlag é igual a zero, a unidade de reconstrução do vetor V 74 pode determinar uma variável tmpWeightVal, definindo a variável tmpWeightVal igual à entrada [CodebkIdx][WeightIdx] do livro de códigos WeightValCdbk. Quando o valor do elemento de sintaxe PFlag não é igual a zero, a unidade de reconstrução do vetor V 74 pode definir a variável tmpWeightVal igual à entrada [CodebkIdx][WeightIdx] do livro de códigos WeightValPredCdbk mais a variável WeightValAlpha multiplicada por tmpWeightVal do k-1-ésimo quadro do i-ésimo canal de transporte. A variável WeightValAlpha pode referir-se ao valor de alfa observado acima, o que pode ser definido estaticamente nos dispositivos de codificação e decodificação de áudio 20 e 24. A unidade de reconstrução do vetor V 74 pode, em seguida, obter o WeightVal como uma função do elemento de sintaxe SgnVal obtido pela unidade de extração 72 e a variável tmpWeightVal.

[0157] A unidade de reconstrução do vetor V 74 pode, em outras palavras, derivar o valor de peso para cada vetor de código correspondente utilizado para reconstruir o vetor V com base em um livro de códigos de valor de peso (indicado como "WeightValCdbk" para quantização vetorial não predita e "WeightValPredCdbk" para quantização vetorial

predita, ambos os quais podem representar uma tabela multidimensional indexada com base em um ou mais de um índice do livro de códigos (indicado elemento de sintaxe "CodebkIdx" na tabela de sintaxe `VVetorData(i)` acima exposta) e um índice de peso (indicado elemento de sintaxe "WeightIdx" na tabela de sintaxe `VVectorData(i)` acima exposta). Este elemento de sintaxe `CodebkIdx` pode ser definido em uma parte das informações de canal secundário, como mostrado na tabela de sintaxe `ChannelSideInfoData(i)`. abaixo

[0158] A parte de quantização vetorial restante do pseudocódigo acima se refere ao cálculo de um `FNorm` para normalizar os elementos do vetor V , seguido por um cálculo do elemento do vetor ($v^{(i)}_{VVecCoeffId[m]}(k)$) como sendo igual a `TmpVVec[idx]` multiplicado por `FNorm`. A unidade de reconstrução do vetor V 74 pode obter a variável `idx` como uma função para o `VVecCoeffID`.

[0159] Quando `NbitsQ` é igual a 5, uma dequantização escalar de 8 bits uniforme é realizada. Em contraste, um valor de `NbitsQ` maior ou igual a 6 pode resultar em aplicação da decodificação de Huffman. O valor `cid` referido acima pode ser igual aos dois bits menos significativos do valor `NbitsQ`. O modo de predição é indicado como o `PFlag` na tabela de sintaxe acima, enquanto o bit info da tabela de Huffman é indicado como o `CbFlag` na tabela de sintaxe acima. A sintaxe restante especifica como a decodificação ocorre de um modo substancialmente semelhante ao descrito acima.

[0160] A unidade de decodificação psicoacústica 80 pode funcionar de um modo recíproco à unidade de codificador de áudio psicoacústica 40 mostrada no exemplo da FIG. 3, de modo a decodificar os coeficientes de HOA de ambiente codificados 59 e os sinais `nFG` codificados 61 e,

assim, gerar coeficientes de HOA de ambiente compensados de energia 47' e os sinais nFG interpolados 49' (que também podem ser referidos como objetos de áudio nFG interpolados 49'). A unidade de decodificação psicoacústica 80 pode passar os coeficientes de HOA de ambiente compensados de energia 47' para a unidade de fade 770 e os sinais nFG 49' para a unidade de formulação de primeiro plano 78.

[0161] A unidade de interpolação espaço-temporal 76 pode funcionar de um modo semelhante ao descrito acima em relação à unidade de interpolação espaço-temporal 50. A unidade de interpolação espaço-temporal 76 pode receber os vetores $V[k]$ de primeiro plano reduzidos 55_k e realizar a interpolação espaço-temporal com respeito aos vetores $V[k]$ de primeiro plano 55_k e os vetores $V[k-1]$ de primeiro plano reduzidos 55_{k-1} para gerar vetores $V[k]$ de primeiro plano interpolados $55_k''$. A unidade de interpolação espaço-temporal 76 pode transmitir os vetores $V[k]$ de primeiro plano interpolados $55_k''$ para a unidade de fade 770.

[0162] A unidade de extração 72 pode também enviar um sinal 757 indicativo de quando um dos coeficientes de HOA de ambiente está em transição para a unidade de fade 770, que pode, então, determinar quais dos SHC_{BG} 47' (onde o SHC_{BG} 47' pode também ser denotado como "canais HOA ambiente 47'" ou "coeficientes de HOA de ambiente 47'") e os elementos dos vetores $V[k]$ de primeiro plano interpolados $55_k''$ devem ser aumentados (faded-in) ou diminuídos (faded-out). Em alguns exemplos, a unidade de fade 770 pode operar em oposição a cada um dos coeficientes de HOA de ambiente 47' e os elementos dos vetores $V[k]$ de primeiro plano interpolados $55_k''$. Ou seja, a unidade de fade 770 pode realizar um aumento ou diminuição, ou tanto um aumento quanto uma diminuição em relação ao coeficiente correspondente dos coeficientes de HOA de ambiente 47',

durante a execução de um aumento ou diminuição ou ambos um aumento e uma diminuição, em relação ao correspondente dos elementos dos vetores $V[k]$ de primeiro plano interpolados $55_k''$. A unidade de fade 770 pode emitir coeficientes de HOA de ambiente ajustados $47''$ para a unidade de formulação de coeficiente de HOA 82 e vetores $V[k]$ de primeiro plano ajustados $55_k'''$ para a unidade de formulação de primeiro plano 78. A este respeito, a unidade de fade 770 representa uma unidade configurada para executar uma operação de fade com respeito a vários aspectos dos coeficientes de HOA ou os seus derivados, por exemplo, sob a forma dos coeficientes de HOA de ambiente $47'$ e os elementos dos vetores $V[k]$ de primeiro plano interpolados $55_k''$.

[0163] A unidade de formulação de primeiro plano 78 pode representar uma unidade configurada para executar a multiplicação de matrizes com respeito aos vetores $V[k]$ de primeiro plano ajustados $55_k'''$ e os sinais nFG interpolados $49'$ para gerar os coeficientes de HOA de primeiro plano 65. A este respeito, a unidade de formulação de primeiro plano 78 pode combinar os objetos de áudio $49'$ (que é outro meio pelo qual indicar os sinais nFG interpolados $49'$) com os vetores $55_k'''$ para reconstruir os aspectos de primeiro plano ou, em outras palavras, predominantes dos coeficientes de HOA $11'$. A unidade de formulação de primeiro plano 78 pode realizar uma multiplicação de matrizes dos sinais nFG interpolados $49'$ pelos vetores $V[k]$ de primeiro plano ajustados $55_k'''$.

[0164] A unidade de formulação do coeficiente de HOA 82 pode representar uma unidade configurada para combinar os coeficientes de HOA de primeiro plano 65 aos coeficientes de HOA de ambiente ajustados $47''$, de modo a obter os coeficientes de HOA $11'$. A notação privilegiada reflete que os coeficientes de HOA $11'$ podem ser

semelhantes, mas não os mesmos que os coeficientes de HOA 11. As diferenças entre os coeficientes de HOA 11 e 11' podem resultar de perda devido à transmissão através de um meio de transmissão com perdas, quantização ou outras operações com perdas.

[0165] A FIG. 5A é um fluxograma que ilustra a operação exemplificativa de um dispositivo de codificação de áudio, tal como o dispositivo de codificação de áudio 20 mostrado no exemplo da FIG. 3, na realização de vários aspectos das técnicas de síntese de base vetorial descritas nesta invenção. Inicialmente, o dispositivo de codificação de áudio 20 recebe os coeficientes de HOA 11 (106). O dispositivo de codificação de áudio 20 pode invocar a unidade de LIT 30, o qual pode aplicar uma LIT com respeito aos coeficientes de HOA para emitir coeficientes de HOA transformados (por exemplo, no caso de SVD, os coeficientes de HOA transformados podem compreender os vetores $US[k]$ 33 e os vetores $V[k]$ 35) (107).

[0166] O dispositivo de codificação de áudio 20 pode, em seguida, invocar a unidade de cálculo de parâmetros 32 para executar a análise descrita acima com respeito a qualquer combinação dos vetores $US[k]$ 33, vetores $US[k-1]$ 33, os vetores $V[K]$ e / ou $V[k-1]$ 35 para identificar vários parâmetros da maneira descrita acima. Isto é, a unidade de cálculo de parâmetros 32 pode determinar pelo menos um parâmetro com base na análise dos coeficientes de HOA transformados 33/35 (108).

[0167] O dispositivo de codificação de áudio 20 pode então invocar a unidade de reordenamento 34, o que pode reordenar os coeficientes de HOA transformados (que, novamente no contexto da SVD, podem se referir aos vetores $US[k]$ 33 e os vetores $V[k]$ 35) com base no parâmetro para gerar coeficientes de HOA transformados reordenados 33'/35'

(ou, em outras palavras, os vetores $US[k]$ 33' e os vetores $V[K]$ 35'), como acima descrito (109). O dispositivo de codificação de áudio 20 pode, durante qualquer uma das operações anteriores ou operações subsequentes, também invocar a unidade de análise de campo de som 44. A unidade de análise de campo de som 44 pode, tal como descrito acima, executar uma análise do campo de som com respeito aos coeficientes de HOA 11 e / ou os coeficientes de HOA transformados 33/35 para determinar o número total de canais de primeiro plano (n_{FG}) 45, a ordem do campo de som de base (N_{BG}) e o número (n_{BGa}) e os índices (i) de canais HOA BG adicionais a enviar (que podem ser coletivamente indicados como informações de canal de base 43 no exemplo da FIG. 3) (109).

[0168] O dispositivo de codificação de áudio 20 pode também invocar a unidade de seleção de base 48. A unidade de seleção de base 48 pode determinar coeficientes de HOA de base ou ambiente 47 com base nas informações de canal de base 43 (110). O dispositivo de codificação de áudio 20 pode ainda invocar a unidade de seleção de primeiro plano 36, que pode selecionar os vetores $US[k]$ reordenados 33' e os vetores $V[k]$ reordenados 35' que representam os componentes de primeiro plano ou distintos do campo de som com base em n_{FG} 45 (o que pode representar um ou mais índices que identificam os vetores de primeiro plano) (112).

[0169] O dispositivo de codificação de áudio 20 pode invocar a unidade de compensação de energia 38. A unidade de compensação de energia 38 pode executar a compensação de energia com respeito aos coeficientes de HOA de ambiente 47 para compensar a perda de energia devido à remoção de vários dos coeficientes de HOA pela unidade de

selecção de base 48 (114) e, assim, gerar coeficientes de HOA de ambiente compensados em energia 47'.

[0170] O dispositivo de codificação de áudio 20 também pode invocar a unidade de interpolação espaço-temporal 50. A unidade de interpolação espaço-temporal 50 pode executar a interpolação espaço-temporal com respeito aos coeficientes de HOA transformados reordenados 33'/35' para obter os sinais de primeiro plano interpolados 49' (que também podem ser referidos como os sinais nFG interpolados 49''') e as informações direcionais de primeiro plano restantes 53 (que também podem ser referidas como vetores "V[k] 53") (116). O dispositivo de codificação de áudio 20 pode então invocar a unidade de redução de coeficiente 46. A unidade de redução de coeficiente 46 pode realizar a redução do coeficiente em relação aos vetores de primeiro plano V[k] restantes 53 com base nas informações do canal de base 43 para obter reduzida informações direcionais de primeiro plano reduzidas 55 (que também podem ser referidas como os vetores V[K] de primeiro plano reduzidos 55) (118).

[0171] O dispositivo de codificação de áudio 20 pode, em seguida, invocar a unidade de quantização 52 para comprimir, na maneira descrita acima, os vetores V[K] de primeiro plano reduzidos 55 e gerar vetores V[K] de primeiro plano codificados 57 (120).

[0172] O dispositivo de codificação de áudio 20 pode também invocar a unidade de codificador de áudio psicoacústica 40. A unidade de codificador de áudio psicoacústica 40 pode codificar de forma psicoacústica cada vetor dos coeficientes de HOA de ambiente compensados em energia 47' e os sinais nFG interpolados 49' para gerar coeficientes de HOA de ambiente codificados 59 e sinais nFG codificados 61. O dispositivo de codificação de áudio pode,

em seguida, invocar a unidade de geração de fluxo de bits 42. A unidade de geração de fluxo de bits 42 pode gerar o fluxo de bits 21 com base nas informações direcionais de primeiro plano codificadas 57, os coeficientes de HOA de ambiente codificados 59, os sinais nFG codificados 61 e as informações de canal fundo43.

[0173] A FIG. 5B é um fluxograma que ilustra a operação exemplificativa de um dispositivo de codificação de áudio na execução das técnicas de codificação descritas nesta invenção. A unidade de geração de fluxo de bits 42 do dispositivo de codificação de áudio 20 mostrado no exemplo da FIG. 3 pode representar uma unidade de exemplo configurado para executar as técnicas descritas na presente invenção. A unidade de geração de fluxo de bits 42 pode determinar se o modo de quantização do quadro é o mesmo que o modo de quantização de um quadro temporalmente anterior (o qual pode ser indicado como um "segundo quando") (314). Embora descritas em relação a um quadro anterior, as técnicas podem ser realizadas com respeito aos quadros temporais subsequentes. O quadro pode incluir uma parte de um ou mais canais de transporte. A parte do canal de transporte pode incluir um ChannelSideInfoData (formado de acordo com a tabela de sintaxe ChannelSideInfoData) juntamente com alguma carga útil (por exemplo, os campos VVectorData 156 no exemplo da FIG. 7). Outros exemplos de cargas úteis podem incluir campos AddAmbientHOACoeffs.

[0174] Quando os modos de quantização são os mesmos ("SIM" 316), a unidade de geração de fluxo de bits 42 pode especificar uma parte do modo de quantização no fluxo de bits 21 (318). A parte do modo de quantização pode incluir o elemento de sintaxe bA e o elemento de sintaxe bB, mas não o elemento de sintaxe uintC. O elemento de sintaxe bA pode representar um bit indicativo do bit mais

significativo do elemento de sintaxe NbitsQ. O elemento de sintaxe bB pode representar um bit indicativo do segundo bit mais significativo do elemento de sintaxe NbitsQ. A unidade de geração de fluxo de bits 42 pode definir o valor de cada um dos elementos de sintaxe bA e elemento de sintaxe bB a zero, assim indicando que o campo de modo de quantização no fluxo de bits 21 (isto é, o campo NbitsQ como um exemplo) não inclui o elemento de sintaxe uintC. Esta sinalização do elemento de sintaxe bA e o elemento de sintaxe bB de valor zero também indica que o valor de NbitsQ, o valor de PFlag, o valor de CbFlag e o valor de CodebkIdx a partir do quadro anterior deve ser usado como os valores correspondentes para os mesmos elementos de sintaxe do quadro corrente.

[0175] Quando os modos de quantização não são os mesmos ("NÃO" 316), a unidade de geração de fluxo de bits 42 pode especificar um ou mais bits indicativos do modo de quantização inteiro no fluxo de bits 21 (320). Ou seja, a unidade de geração de fluxo de bits 42 especifica os elementos de sintaxe bA, Bb e uintC no fluxo de bits 21. A unidade de geração de fluxo de bits 42 também pode especificar informações de quantização com base no modo de quantização (322). Estas informações de quantização podem incluir quaisquer informações relacionadas com quantização, tais como as informações de quantização vetorial, as informações de predição e as informações do livro de códigos de Huffman. As informações de quantização vetorial podem incluir, como um exemplo, um ou ambos o elemento de sintaxe CodebkIdx e o elemento de sintaxe NumVecIndices. As informações de predição podem incluir, como um exemplo, o elemento de sintaxe PFlag. As informações do livro de códigos de Huffman podem incluir, como um exemplo, o elemento de sintaxe CbFlag.

[0176] A este respeito, as técnicas podem permitir que o dispositivo de codificação de áudio 20 seja configurado para obter um fluxo de bits 21 que compreende uma versão comprimida de um componente espacial de um campo de som. O componente espacial pode ser gerado através da realização de uma síntese de base vetorial em relação a uma pluralidade de coeficientes harmônicos esféricos. O fluxo de bits pode compreender ainda um indicador quanto à reutilização de um ou mais bits de um campo de cabeçalho, a partir de um quadro anterior, que especifica as informações usadas quando se comprime o componente espacial.

[0177] Em outras palavras, as técnicas podem permitir que o dispositivo de codificação de áudio 20 seja configurado para obter um fluxo de bits 21 que compreende um vetor 57 representativo de um eixo espacial ortogonal em um domínio de harmônicos esféricos. O fluxo de bits 21 pode ainda compreender um indicador (por exemplo, elementos de sintaxe bA / bB do elemento de sintaxe NbitsQ) quanto à reutilização, a partir de um quadro anterior, pelo menos um elemento de sintaxe indicativo das informações utilizadas quando comprimindo (por exemplo, quantizando) o vetor.

[0178] A FIG. 6A é um fluxograma que ilustra a operação exemplificativa de um dispositivo de decodificação de áudio, tal como o dispositivo de decodificação de áudio 24 mostrado na FIG. 4, na realização de vários aspectos das técnicas descritas nesta invenção. Inicialmente, o dispositivo de decodificação de áudio 24 pode receber o fluxo de bits 21 (130). Ao receber o fluxo de bits, o dispositivo de decodificação de áudio 24 pode invocar a unidade de extração 72. Supondo que, para fins de discussão, o fluxo de bits 21 indica que a reconstrução de base vetorial deve ser realizada, a unidade de extração 72 pode analisar o fluxo de bits para recuperar as informações

acima notadas, passando as informações para a unidade de reconstrução de base vetorial 92.

[0179] Em outras palavras, a unidade de extração 72 pode extrair as informações direcionais de primeiro plano codificadas 57 (que, mais uma vez, podem também ser referidas como os vetores $V[K]$ de primeiro plano codificados 57), os coeficientes de HOA de ambiente codificados 59 e os sinais de primeiro plano codificados (que também podem ser referidos como os sinais nFG de primeiro plano codificados 59 ou os objetos de áudio de primeiro plano codificados 59) a partir do fluxo de bits 21 da maneira descrita acima (132).

[0180] O dispositivo de decodificação de áudio 24 pode ainda invocar a unidade de dequantização 74. A unidade de dequantização 74 pode decodificar por entropia e dequantizar as informações direcionais de primeiro plano codificadas 57 para obter informações direcionais de primeiro plano reduzidas 55_k (136). O dispositivo de decodificação de áudio 24 também pode invocar a unidade de decodificação psicoacústica 80. A unidade de decodificação de áudio psicoacústica 80 pode decodificar os coeficientes de HOA de ambiente codificados 59 e os sinais de primeiro plano codificados 61 para obtenção de coeficientes de HOA de ambiente compensados em energia 47' e os sinais de primeiro plano interpolados 49' (138). A unidade de decodificação psicoacústica 80 pode passar os coeficientes de HOA de ambiente compensados em energia 47' para a unidade de fade 770 e os sinais nFG 49' para a unidade de formulação de primeiro plano 78.

[0181] O dispositivo de decodificação de áudio 24 pode, em seguida, invocar a unidade de interpolação espaço-temporal 76. A unidade de interpolação espaço-temporal 76 pode receber as informações direcionais de primeiro plano

reordenadas 55_k e realizar a interpolação espaço-temporal com respeito às informações direcionais de primeiro plano reduzidas $55_k / 55_{k-1}$ para gerar as informações direcionais de primeiro plano interpoladas $55_k''$ (140). A unidade de interpolação espaço-temporal 76 pode transmitir os vetores $V[k]$ de primeiro plano interpolados $55_k''$ para a unidade de fade 770.

[0182] O dispositivo de decodificação de áudio 24 pode invocar a unidade de fade 770. A unidade de fade 770 pode receber ou de outra forma obter elementos de sintaxe (por exemplo, a partir da unidade de extração 72) indicativos de quando os coeficientes de HOA de ambiente compensados em energia 47' estão em transição (por exemplo, o elemento de sintaxe AmbCoeffTransition). A unidade de fade 770 pode, com base nos elementos de sintaxe de transição e as informações de estado de transição mantidas, diminuir ou aumentar os coeficientes de HOA de ambiente compensados em energia 47' que emitem coeficientes de HOA de ambiente ajustados 47'' para a unidade de formulação de coeficiente de HOA 82. A unidade de fade 770 também pode, com base nos elementos de sintaxe e as informações de estado de transição mantidas, aumentar ou diminuir os correspondentes um ou mais elementos dos vetores $V[k]$ de primeiro plano interpolados $55_k''$ que emitem os vetores $V[K]$ de primeiro plano ajustados $55_k'''$ para a unidade de formulação de primeiro plano 78 (142).

[0183] O dispositivo de decodificação de áudio 24 pode invocar a unidade de formulação de primeiro plano 78. A unidade de formulação de primeiro plano 78 pode realizar a multiplicação de matriz dos sinais nFG 49' pelas informações direcionais de primeiro plano ajustadas $55_k'''$ para obter os coeficientes de HOA de primeiro plano 65 (144). O dispositivo de decodificação de áudio 24 pode

também invocar a unidade de formulação de coeficientes de HOA 82. A unidade de formulação de coeficiente de HOA 82 pode adicionar os coeficientes de HOA de primeiro plano 65 para ajustar os coeficientes de HOA de ambiente ajustados 47'', de modo a obter os coeficientes de HOA 11' (146).

[0184] A FIG. 6B é um fluxograma que ilustra a operação exemplificativa de um dispositivo de decodificação de áudio na execução das técnicas de codificação descritas na presente invenção. A unidade de extração 72 do dispositivo de codificação de áudio 24 mostrada no exemplo da FIG. 4 pode representar uma unidade exemplificativa configurada para executar as técnicas descritas na presente invenção. A unidade de extração de fluxo de bits 72 pode obter bits indicativos de se o modo de quantização do quadro é o mesmo que o modo de quantização de um quadro temporalmente anterior (o qual pode ser indicado como um "segundo quadro") (362). Mais uma vez, embora descritos em relação a um quadro anterior, as técnicas podem ser realizadas no que diz respeito a quadros temporalmente subsequentes.

[0185] Quando os modos de quantização são os mesmos ("SIM" 364), a unidade de extração 72 pode obter uma parte do modo de quantização a partir do fluxo de bits 21 (366). A parte do modo de quantização pode incluir o elemento de sintaxe bA e o elemento de sintaxe bB, mas não o elemento de sintaxe uintC. A unidade de extração 42 pode também definir os valores do valor de NbitsQ, o valor de PFlag, o valor de CbFlag, o valor de CodebkIdx e o valor de NumVertIndices para o quadro corrente como sendo os mesmos que os valores do valor de NbitsQ, o valor de PFlag, o valor de CbFlag, o valor CodebkIdx e o valor de NumVertIndices definidos para o quadro anterior (368).

[0186] Quando os modos de quantização não são os mesmos ("NÃO" 364), a unidade de extração 72 pode obter um ou mais bits indicativos do modo de quantização inteiro a partir do fluxo de bits 21. Isto é, a unidade de extração 72 obtém os elementos de sintaxe bA, bB e uintC do fluxo de bits 21 (370). A unidade de extração 72 também pode obter um ou mais bits indicativos das informações de quantização com base no modo de quantização (372). Como observado acima, com respeito à FIG. 5B, as informações de quantização podem incluir quaisquer informações relacionadas com quantização, tais como as informações de quantização vetorial, as informações de predição e as informações do livro de códigos de Huffman. As informações de quantização vetorial podem incluir, como um exemplo, um ou ambos os elementos de sintaxe CodebkIdx e elemento de sintaxe NumVecIndices. As informações de predição podem incluir, como um exemplo, o elemento de sintaxe PFlag. As informações do livro de códigos de Huffman podem incluir, como um exemplo, o elemento de sintaxe CbFlag.

[0187] A este respeito, as técnicas podem permitir que o dispositivo de decodificação de áudio 24 seja configurado para obter um fluxo de bits 21 que compreende uma versão comprimida de um componente espacial de um campo de som. O componente espacial pode ser gerado através da realização de uma síntese de base vetorial em relação a uma pluralidade de coeficientes harmônicos esféricos. O fluxo de bits pode compreender ainda um indicador quanto à reutilização de um ou mais bits de um campo de cabeçalho, a partir de um quadro anterior, que especifica as informações usadas quando se comprime o componente espacial.

[0188] Em outras palavras, as técnicas podem permitir que o dispositivo de decodificação de áudio 24

seja configurado para obter um fluxo de bits 21 que compreende um vetor 57 representativo de um eixo espacial ortogonal em um domínio de harmônicos esféricos. O fluxo de bits 21 pode ainda compreender um indicador (por exemplo, os elementos de sintaxe bA / bB do elemento de sintaxe NbitsQ) quanto à reutilização, a partir de um quadro anterior, de pelo menos um elemento de sintaxe indicativo das informações utilizadas quando comprimindo (por exemplo, quantizando) o vetor.

[0189] A FIG. 7 é um diagrama que ilustra quadros exemplificativos 249S e 249T especificados de acordo com vários aspectos das técnicas descritas nesta invenção. Como mostrado no exemplo da FIG. 7, o quadro 249S inclui campos ChannelSideInfoData (CSID) 154A-154D, campos HOAGainCorrectionData (HOAGCD), campos VVetorData 156A e 156B e campos HOAPredictionInfo. O campo CSID 154A inclui um elemento de sintaxe uintC ("uintC") 267 ajustado a um valor de 10, um elemento de sintaxe bb ("bB") 266 ajustado a um valor de 1 e um elemento de sintaxe bA ("Ba") 265 ajustado a um valor de 0, juntamente com um elemento de sintaxe ChannelType ("ChannelType") 269 ajustado a um valor de 01.

[0190] O elemento de sintaxe uintC 267, o elemento de sintaxe bB 266 e o elemento de sintaxe bA 265 em conjunto formam o elemento de sintaxe NbitsQ 261 com o elemento de sintaxe bA 265 formando os bits mais significativos, o elemento de sintaxe bB 266 formando os segundos bits mais significativos e o elemento de sintaxe uintC 267 formando os bits menos significativos do elemento de sintaxe NbitsQ 261. O elemento de sintaxe NbitsQ 261 pode, como mencionado acima, representar um ou mais bits indicativos de um modo de quantização (por exemplo, um dos modo de quantização vetorial, quantização escalar sem modo

de codificação de Huffman e quantização escalar com modo de codificação de Huffman) utilizados para codificar os dados de áudio ambissônicos de ordem superior.

[0191] O elemento de sintaxe CSID 154A também inclui um elemento de sintaxe PFlag 300 e um elemento de sintaxe CbFlag 302 acima referenciado em várias tabelas de sintaxe. O elemento de sintaxe PFlag 300 pode representar um ou mais bits indicativos de se um elemento codificado de um componente espacial do campo de som representado pelos coeficientes de HOA 11 (onde mais uma vez um componente espacial pode se referir ao vetor V) de um primeiro quadro 249S é predito a partir de um segundo quadro (por exemplo, um quadro anterior neste exemplo). O elemento de sintaxe CbFlag 302 pode representar um ou mais bits indicativos de uma informação do livro de códigos de Huffman, que pode identificar qual dos livros de código de Huffman (ou, em outras palavras, tabelas) utilizou para codificar os elementos do componente espacial (ou, em outras palavras, elementos do vetor V).

[0192] O campo CSID 154B inclui um elemento de sintaxe bB 266 e um elemento de sintaxe bB 265, juntamente com o elemento de sintaxe ChannelType 269, cada um deles ajustado aos valores correspondentes 0 e 0 e 01, no exemplo da FIG. 7. Cada um dos campos CSID 154C e 154D inclui o campo ChannelType 269 com um valor de 3 (11_2). Cada um dos campos CSID 154A-154d corresponde ao respectivo dos canais de transporte 1, 2, 3 e 4. Com efeito, cada campo CSID 154A-154d indica se uma carga útil correspondente são sinais de base direcional (quando o ChannelType correspondente é igual a zero), sinais de base vetorial (quando o ChannelType correspondente é igual a um), um coeficiente de HOA Ambiente adicional (quando o ChannelType

correspondente é igual a dois), ou vazio (quando o ChannelType é igual a três).

[0193] No exemplo da FIG. 7, o quadro 249S inclui dois sinais de base vetorial (considerando que os elementos de sintaxe ChannelType 269 são iguais a 1 nos campos CSID 154A e 154B) e dois vazios (considerando que o ChannelType 269 é igual a 3 nos campos CSID 154C e 154D). Além disso, o dispositivo de codificação de áudio 20 usou predição, tal como indicado pelo elemento de sintaxe PFlag 300 sendo ajustado a um. Mais uma vez, a predição como indicada pelo elemento de sintaxe PFlag 300 se refere a uma indicação do modo de predição indicativa de se a predição foi realizada com respeito ao correspondente dos componentes espaciais comprimidos vl-vn. Quando o elemento de sintaxe PFlag 300 é ajustado a um, o dispositivo de codificação de áudio 20 podem usar predição tomando uma diferença entre, para quantização escalar, um elemento de vetor a partir de um quadro anterior com o elemento de vetor correspondente do quadro corrente ou, para quantização vetorial, uma diferença entre um peso a partir de um quadro anterior com um peso correspondem do quadro corrente.

[0194] O dispositivo de codificação de áudio 20 determinou também que o valor para o elemento de sintaxe NbitsQ 261 para o campo CSID 154B do segundo canal de transporte no quadro 249S é o mesmo que o valor do elemento de sintaxe NbitsQ 261 para o campo CSID 154B do segundo canal de transporte do quadro anterior, por exemplo, quadro 249T no exemplo da FIG. 7. Como resultado, o dispositivo de codificação de áudio 20 especificou um valor de zero para cada um dos elemento de sintaxe bA 265 e o elemento de sintaxe bB 266 para sinalizar que o valor do elemento de sintaxe NbitsQ 261 do segundo canal de transporte no quadro anterior 249T é reutilizado para o elemento de sintaxe

NbitsQ 261 do segundo canal de transporte no quadro 249S. Como resultado, o dispositivo de codificação de áudio 20 pode evitar especificar o elemento de sintaxe uintC 267 para o segundo canal de transporte no quadro 249S, juntamente com o outro elemento de sintaxe acima identificado.

[0195] A FIG. 8 é um diagrama que ilustra quadros exemplificativos para um ou mais canais de pelo menos um fluxo de bits de acordo com as técnicas aqui descritas. O fluxo de bits 450 inclui quadros 810A-810H que podem, cada um, incluir um ou mais canais. O fluxo de bits 450 pode ser um exemplo do fluxo de bits 21 mostrado no exemplo da FIG. 7. No exemplo da FIG. 8, o dispositivo de decodificação de áudio 24 mantém informações de estado, atualizando as informações de estado para determinar como decodificar o quadro corrente k (ver a figura 8). O dispositivo de decodificação de áudio 24 pode utilizar as informações de estado da configuração 814, e quadros 810B-810D.

[0196] Em outras palavras, o dispositivo de codificação de áudio 20 pode incluir, dentro da unidade de geração de corrente de fluxo de bits 42, por exemplo, a máquina de estados 402 que mantém informações de estado para codificar cada um dos quadros 810A-810E, em que a unidade de geração de fluxo de bits 42 pode especificar elementos de sintaxe para cada um dos quadros 810A-810E com base na máquina de estado 402.

[0197] O dispositivo de decodificação de áudio 24 pode também incluir, dentro da unidade de extração de fluxo de bits 72, por exemplo, uma máquina de estado 402 similar que emite elementos de sintaxe (alguns dos quais não são explicitamente especificados no fluxo de bits 21) com base na máquina de estado 402. A máquina de estado 402 do dispositivo de decodificação de áudio 24 pode operar de um

modo semelhante ao da máquina de estado 402 do dispositivo de codificação de áudio 20. Como tal, a máquina de estado 402 do dispositivo de decodificação de áudio 24 pode manter as informações de estado, atualizando as informações de estado com base na configuração 814 e, no exemplo da FIG. 8, a decodificação dos quadros 810B-810D. Com base nas informações de estado, a unidade de extração de fluxo de bits 72 pode extrair o quadro 810E com base nas informações de estado mantidas pela máquina de estado 402. As informações de estado podem proporcionar uma série de elementos de sintaxe implícitos que o dispositivo de codificação de áudio 20 pode utilizar quando decodificando os vários canais de transporte do quadro 810E.

[0198] As técnicas anteriores podem ser realizados com respeito a qualquer número de diferentes contextos e ecossistemas de áudio. Uma série de exemplos de contextos é descrita abaixo, embora as técnicas devam ser limitadas aos contextos exemplificativos. Um ecossistema de áudio exemplificativo pode incluir conteúdo de áudio, estúdios de cinema, estúdios de música, estúdios de áudio de jogos, conteúdo de áudio com base em canal, motores de codificação, hastes de áudio de jogo, motores de renderização / codificação de áudio de jogo e sistemas de distribuição.

[0199] Os estúdios de cinema, os estúdios de música e estúdios de áudio de jogos podem receber conteúdo de áudio. Em alguns exemplos, o conteúdo de áudio pode representar a saída de uma aquisição. Os estúdios de cinema podem emitir conteúdo de áudio com base em canal (por exemplo, em 2.0, 5.1 e 7.1), tal como usando uma estação de trabalho de áudio digital (DAW). Os estúdios de música podem emitir conteúdo de áudio com base em canal (por exemplo, em 2,0 e 5,1), como por meio de um DAW. Em ambos

os casos, os motores de codificação podem receber e codificar o conteúdo de áudio com base em canal com base em um ou mais codecs (por exemplo, AAC, AC3, Dolby True HD, Dolby Digital Plus e DTS Master Audio) para emissão pelos sistemas de distribuição. Os estúdios de áudio de jogos podem emitir de uma ou mais hastes de áudio de jogos, tal como usando um DAW. Os motores de codificação / renderização de áudio de jogos podem codificar e ou renderizar as hastes de áudio em de conteúdo de áudio com base em canal para emissão pelos sistemas de distribuição. Outro contexto exemplificativo em que as técnicas podem ser realizadas compreende um ecossistema de áudio que pode incluir objetos de gravação de áudio de transmissão, sistemas de áudio profissionais, captura em dispositivo de consumidor, formato de áudio HOA, renderização em dispositivo, áudio de consumidor, TV e acessórios, e sistemas de áudio para automóveis.

[0200] Os objetos de gravação de áudio de transmissão, os sistemas de áudio profissionais e a captura em dispositivo de consumidor podem, todos, codificar sua emissão usando formato de áudio HOA. Desta forma, o conteúdo de áudio pode ser codificado usando o formato de áudio HOA em uma única representação que pode ser reproduzida utilizando a renderização em dispositivo, o áudio de consumidor, TV e acessórios, e os sistemas de áudio para automóveis. Em outras palavras, a representação simples do conteúdo de áudio pode ser reproduzida em um sistema de reprodução de áudio genérico (isto é, em oposição à exigência de uma configuração específica, tal como 5.1, 7.1, etc.), tal como um sistema de reprodução de áudio 16.

[0201] Outros exemplos de contexto em que as técnicas podem ser realizadas incluem um ecossistema de

áudio que pode incluir elementos de aquisição e elementos de reprodução. Os elementos de aquisição podem incluir dispositivos de aquisição com fios e / ou sem fios (por exemplo, microfones Eigen), captura de som surround em dispositivo e dispositivos móveis (por exemplo, *smartphones* e *tablets*). Em alguns exemplos, os dispositivos de aquisição com fios e / ou sem fios podem ser acoplados ao dispositivo móvel através de canal(ais) de comunicação com fios e / ou sem fios.

[0202] De acordo com uma ou mais técnicas desta invenção, o dispositivo móvel pode ser usado para adquirir um campo de som. Por exemplo, o dispositivo móvel pode adquirir um campo de som através dos dispositivos de aquisição com fios e / ou sem fios e / ou a captura de som surround em dispositivo (por exemplo, uma pluralidade de microfones integrados ao dispositivo móvel). O dispositivo móvel pode então codificar o campo de som adquirido nos coeficientes de HOA para reprodução por um ou mais dos elementos de reprodução. Por exemplo, um usuário do dispositivo móvel pode gravar (adquirir um campo de som de) um evento ao vivo (por exemplo, uma reunião, uma conferência, uma peça de teatro, um concerto, etc.), e codificar a gravação em coeficientes de HOA.

[0203] O dispositivo móvel pode também utilizar um ou mais dos elementos de reprodução para reproduzir o campo de som codificado por HOA. Por exemplo, o dispositivo móvel pode decodificar o campo de som codificado por HOA e emitir um sinal para um ou mais dos elementos de reprodução que leva o um ou mais dos elementos de reprodução a recriar o campo sonoro. Como um exemplo, o dispositivo móvel pode utilizar os canais de comunicação sem fios e / ou com fios para emitir o sinal para um ou mais alto-falantes (por exemplo, matrizes de alto-falante, barras de som, etc.).

Como outro exemplo, o dispositivo móvel pode utilizar soluções de acoplamento para emitir o sinal para uma ou mais estações de acoplamento e / ou um ou mais alto-falantes acoplados (por exemplo, sistemas de som em carros e / ou domicílios inteligentes). Como outro exemplo, o dispositivo móvel pode utilizar renderização de fone de ouvido para emitir o sinal a um conjunto de fones de ouvido, por exemplo, para criar um som binaural realista.

[0204] Em alguns exemplos, um dispositivo móvel específico pode tanto adquirir um campo de som 3D e reproduzir o mesmo campo de som 3D em um momento posterior. Em alguns exemplos, o dispositivo móvel pode adquirir um campo de som 3D, codificar o campo de som 3D em HOA e transmitir o campo de som 3D codificado para um ou mais outros dispositivos (por exemplo, outros dispositivos móveis e / ou outros dispositivos não móveis) para reprodução.

[0205] Ainda outro contexto em que as técnicas podem ser realizadas inclui um ecossistema de áudio que pode incluir conteúdo de áudio, estúdios de jogos, conteúdo de áudio codificado, mecanismos de renderização e sistemas de distribuição. Em alguns exemplos, os estúdios de jogos podem incluir um ou mais DAWs que podem suportar a edição de sinais HOA. Por exemplo, os um ou mais DAWs podem incluir encaixes e / ou ferramentas HOA que podem ser configurados para operar com (por exemplo, trabalhar com) um ou mais sistemas de áudio de jogos. Em alguns exemplos, os estúdios de jogos podem emitir novos formatos de hastes que suportam HOA. Em qualquer caso, os estúdios de jogos podem emitir conteúdo de áudio codificado para os motores de renderização que podem renderizar um campo de som para reprodução pelos sistemas de distribuição.

[0206] As técnicas podem também ser executadas no que diz respeito a dispositivos de aquisição de áudio exemplificativos. Por exemplo, as técnicas podem ser realizadas no que diz respeito a um microfone Eigen, que pode incluir uma pluralidade de microfones que são coletivamente configurados para registrar um campo de som 3D. Em alguns exemplos, a pluralidade de microfones de microfone Eigen pode ser localizada na superfície de uma esfera substancialmente esférica com um raio de cerca de 4 centímetros. Em alguns exemplos, o dispositivo de codificação de áudio 20 pode ser integrado ao microfone Eigen, de modo a emitir um fluxo de bits 21 diretamente a partir do microfone.

[0207] Outro contexto de aquisição de áudio exemplificativo pode incluir um caminhão de produção que pode ser configurado para receber um sinal a partir de um ou mais microfones, tais como um ou mais microfones Eigen. O caminhão de produção pode também incluir um codificador de áudio, como codificador de áudio 20 da FIG. 3.

[0208] O dispositivo móvel pode também, em certos casos, incluir uma pluralidade de microfones que são coletivamente configurados para registrar um campo de som 3D. Em outras palavras, a pluralidade de microfones pode ter X, Y, Z diversidades. Em alguns exemplos, o dispositivo móvel pode incluir um microfone que pode ser girado para fornecer X, Y, Z diversidades com respeito a um ou mais outros microfones do dispositivo móvel. O dispositivo móvel pode também incluir um codificador de áudio, como o codificador de áudio 20 da FIG. 3.

[0209] Um dispositivo de captura de vídeo robusto pode ainda ser configurado para gravar um campo de som 3D. Em alguns exemplos, o dispositivo de captura de vídeo robusto pode ser ligado a um capacete de um usuário

envolvido em uma atividade. Por exemplo, o dispositivo de captura de vídeo robusto pode estar ligado a um capacete de um usuário de *rafting*. Desta forma, o dispositivo de captura de vídeo robusto pode capturar um campo de som 3D que representa a ação ao redor do usuário (por exemplo, água caindo atrás do usuário, outro praticante de rafting falando em frente ao usuário, etc.).

[0210] As técnicas podem também ser executadas no que diz respeito a um dispositivo móvel melhorado por acessório, que pode ser configurado para registrar um campo de som 3D. Em alguns exemplos, o dispositivo móvel pode ser semelhante aos dispositivos móveis discutidos acima, com a adição de um ou mais acessórios. Por exemplo, um microfone Eigen pode ser ligado ao dispositivo móvel acima notado, para formar um dispositivo móvel melhorado por acessório. Desta forma, o dispositivo móvel melhorado por acessório pode capturar uma versão de alta qualidade do campo de som 3D do que apenas usando componentes de captura de som integrantes ao dispositivo móvel melhorado por acessório.

[0211] Exemplos de dispositivos de reprodução de áudio que podem executar vários aspectos das técnicas descritas nesta invenção são discutidos abaixo. Em conformidade com uma ou mais técnicas desta invenção, alto-falantes e / ou barras de som podem ser dispostas em qualquer configuração arbitrária, ao mesmo tempo ainda reproduzindo um campo de som 3D. Além disso, em alguns exemplos, os dispositivos de reprodução de fones de ouvido podem ser acoplados a um decodificador 24, através ou de uma ligação com fios ou sem fios. Em conformidade com uma ou mais técnicas desta invenção, uma única representação genérica de um campo de som pode ser utilizada para renderizar o campo de som em qualquer combinação dos alto-

falantes, as barras de som, e os dispositivos de reprodução de fones de ouvido.

[0212] Uma série de diferentes exemplos de ambientes de reprodução de áudio também pode ser adequada para a realização de vários aspectos das técnicas descritas nesta invenção. Por exemplo, um ambiente de reprodução de alto-falantes 5.1, um ambiente de reprodução alto-falantes 2.0 (por exemplo, estéreo), um ambiente de reprodução de alto-falantes 9.1 com alto-falantes frontais de altura total, um ambiente de reprodução de alto-falantes 22.2, um ambiente de reprodução de alto-falantes 16.0, um ambiente de reprodução de alto-falantes automotivo e um dispositivo móvel com o ambiente de reprodução de botão auricular podem ser ambientes adequados para a realização de vários aspectos das técnicas descritas nesta invenção.

[0213] Em conformidade com uma ou mais técnicas desta invenção, uma única representação genérica de um campo de som pode ser utilizada para renderizar o campo de som em qualquer dos ambientes de reprodução anteriores. Adicionalmente, as técnicas desta invenção habilitam uma renderização para renderizar um campo de som a partir de uma representação genérica para reprodução nos ambientes de reprodução diferentes daqueles descritos acima. Por exemplo, se as considerações de design proibem a colocação adequada de alto-falantes de acordo com um ambiente de reprodução de alto-falantes 7.1 (por exemplo, se não é possível colocar um alto-falante surround direito), as técnicas desta invenção permitem uma renderização para compensar com os outros 6 alto-falantes, de modo que a reprodução pode ser alcançada em um ambiente de reprodução de alto-falantes 6.1.

[0214] Além disso, um usuário pode assistir a um jogo de esportes enquanto usa fones de ouvido. Em

conformidade com uma ou mais das técnicas da presente invenção, o campo sonoro 3D do jogo de esportes pode ser adquirido (por exemplo, um ou mais microfones Eigen podem ser colocados no interior e / ou em torno do estádio de beisebol), coeficientes de HOA correspondentes ao campo de som 3D podem ser obtidos e transmitidos para um decodificador, o decodificador pode reconstruir o campo de som 3D com base nos coeficientes de HOA e emitir o campo de som 3D reconstruído a um renderizador, o renderizador pode obter uma indicação quanto ao tipo de ambiente de reprodução (por exemplo, fones de ouvido), e renderizar o campo de som 3D reconstruído em sinais que levam os fones de ouvido a emitir uma representação do campo de som 3D do jogo de esportes.

[0215] Em cada um dos vários exemplos descritos acima, deve ser entendido que o dispositivo de codificação de áudio 20 pode executar um método ou de outro modo compreender meios para realizar cada etapa do método para o qual o dispositivo de codificação de áudio 20 está configurado para executar. Em alguns casos, os meios podem compreender um ou mais processadores. Em alguns casos, os um ou mais processadores podem representar um processador de fim especial configurado por meio de instruções armazenadas para um meio de armazenamento de leitura por computador não transitório. Em outras palavras, os vários aspectos das técnicas em cada um dos conjuntos de exemplos de codificação podem prever um meio de armazenamento legível por computador não transitório tendo neles armazenadas instruções que, quando executadas, levam o um ou mais processadores a executar o método para o qual o dispositivo de codificação de áudio 20 foi configurado para executar.

[0216] Em um ou mais exemplos, as funções descritas podem ser implementadas em hardware, software, firmware ou qualquer combinação dos mesmos. Se implementadas em software, as funções podem ser armazenadas ou transmitidas como uma ou mais instruções ou código em um meio de leitura por computador e executadas por uma unidade de processamento com base em hardware. Meios legíveis por computador podem incluir meios de armazenamento legíveis por computador, o que corresponde a um meio tangível, tal como meios de armazenamento de dados. Meios de armazenamento de dados podem ser qualquer material disponível que possa ser acessado por um ou mais computadores ou um ou mais processadores para recuperar instruções, código, e / ou estruturas de dados para a implementação das técnicas descritas nesta invenção. Um produto de programa de computador pode incluir um meio de leitura por computador.

[0217] Do mesmo modo, em cada um dos vários exemplos descritos acima, deve ser entendido que o dispositivo de decodificação de áudio 24 pode executar um método ou de outro modo compreender meios para realizar cada etapa do método para o qual o dispositivo de decodificação de áudio 24 está configurado para executar. Em alguns casos, os meios podem compreender um ou mais processadores. Em alguns casos, os um ou mais processadores podem representar um processador de finalidade especial configurado por meio de instruções armazenadas para um meio de armazenamento não transitório de leitura por computador. Em outras palavras, vários aspectos das técnicas em cada um dos conjuntos de exemplos de codificação podem prever um meio de armazenamento não transitório legível por computador tendo nele armazenadas instruções que, quando executadas, levam o um ou mais processadores a executar o

método para o qual o dispositivo de decodificação de áudio 24 foi configurado para executar.

[0218] A título de exemplo, e não como limitação, esses meios de armazenamento legíveis por computador podem compreender RAM, ROM, EEPROM, CD-ROM ou outro armazenamento em disco óptico, armazenamento em disco magnético, ou outros dispositivos de armazenamento magnéticos, memória flash, ou qualquer outro meio que possa ser usado para armazenar o código de programa desejado sob a forma de instruções ou estruturas de dados, e que possa ser acessado por um computador. Deve ser entendido, no entanto, que os meios de armazenamento legíveis por computador e meios de armazenamento de dados não incluem conexões, ondas portadoras, sinais ou outros meios transitórios, mas são direcionados a meios de armazenamento tangíveis não transitórios. Disco (disk) e disco (disc), como aqui utilizados, incluem disco compacto (CD), disco laser, disco óptico, disco versátil digital (DVD), disquete e disco Blu-ray, em que discos (discs) geralmente reproduzem dados magneticamente, ao passo que discos (discs) reproduzem dados opticamente com lasers. Combinações dos anteriores também devem ser incluídas dentro do âmbito dos meios de leitura por computador.

[0219] As instruções podem ser executadas por um ou mais processadores, tal como um ou mais processadores de sinal digital (DSPs), microprocessadores de uso geral, circuitos integrados de aplicação específica (ASICs), arranjos de portas programáveis em campo (FPGAs), ou outros circuitos lógicos discretos ou integrados equivalentes. Por conseguinte, o termo "processador" tal como aqui utilizado pode referir-se a qualquer uma das estruturas precedentes ou qualquer outra estrutura adequada para aplicação das técnicas aqui descritas. Além disso, em alguns aspectos, a

funcionalidade aqui descrita pode ser fornecida dentro de módulos de hardware e / ou software dedicados configurados para codificação e decodificação, ou incorporados em um codec combinado. Além disso, as técnicas podem ser totalmente implementadas em um ou mais circuitos ou elementos lógicos.

[0220] As técnicas da presente invenção podem ser implementadas em uma vasta variedade de dispositivos ou aparelhos, incluindo um dispositivo portátil sem fios, um circuito integrado (IC) ou um conjunto de ICs (por exemplo, um conjunto de chips). Vários componentes, módulos ou unidades são descritas na presente descrição para enfatizar aspectos funcionais de dispositivos configurados para executar as técnicas divulgadas, mas não precisam necessariamente de realização por diferentes unidades de hardware. Em vez disso, tal como descrito acima, várias unidades podem ser combinadas em uma unidade de hardware codec ou fornecidas por um conjunto de unidades de hardware interoperativas, incluindo um ou mais processadores, conforme descrito acima, em conjunto com o software e / ou firmware adequado.

[0221] Vários aspectos das técnicas têm sido descritos. Estes e outros aspectos das técnicas estão dentro do âmbito das reivindicações a seguir.

REIVINDICAÇÕES

1. Método de uso de bits eficiente, o método **caracterizado** pelo fato de que compreende:

obter um fluxo de bits que compreende uma versão comprimida de um componente especial de um campo de som, o componente especial do campo de som sendo representado por um vetor representativo de um eixo espacial ortogonal em um domínio de harmônicos esféricos, em que o fluxo de bits compreende adicionalmente um indicador quanto à reutilização, a partir de um quadro anterior, um elemento de sintaxe indicativo de um modo de predição indicativo de se a predição foi realizada em relação ao vetor.

2. Método, de acordo com a reivindicação 1, **caracterizado** pelo fato de que o indicador compreende um ou mais bits de um elemento de sintaxe que indicam um modo de quantização usado quando se comprime o vetor.

3. Método, de acordo com a reivindicação 2, **caracterizado** pelo fato de que o um ou mais bits do elemento de sintaxe, quando definido para um valor de zero, indica reutilizar o pelo menos um elemento de sintaxe do quadro anterior.

4. Método, de acordo com a reivindicação 2, **caracterizado** pelo fato de que a porção do elemento de sintaxe inclui um bit mais significativo do elemento de sintaxe e um segundo bit mais significativo do elemento de sintaxe.

5. Dispositivo configurado para executar o uso de bits eficiente, o dispositivo **caracterizado** pelo fato de que compreende:

meios para obter um fluxo de bits que compreende uma versão comprimida de um componente especial de um campo de som, o componente especial do campo de som sendo representado por um vetor representativo de um eixo espacial

ortogonal em um domínio de harmônicos esféricos, em que o fluxo de bits compreende adicionalmente um indicador quanto à reutilização, a partir de um quadro anterior, um elemento de sintaxe indicativo de um modo de predição indicativo de se a predição foi realizada em relação ao vetor ; e

meios para armazenar o indicador.

6. Dispositivo, de acordo com a reivindicação 5, **caracterizado** pelo fato de que o elemento de sintaxe do quadro anterior compreende um elemento de sintaxe indicativo de um modo de previsão usado quando se comprime o vetor.

7. Dispositivo, de acordo com a reivindicação 5, **caracterizado** pelo fato de que o elemento de sintaxe do quadro anterior compreende um elemento de sintaxe indicativo de uma tabela de Huffman usada quando se comprime o vetor, um identificador de categoria que identifica uma categoria de compressão à qual o vetor corresponde, se um elemento do vetor for um valor positivo ou um valor negativo, diversos vetores de código usados quando se comprime o vetor ou um livro de código usado quando se comprime o vetor.

8. Dispositivo, de acordo com a reivindicação 5, **caracterizado** pelo fato de que a versão comprimida do vetor é representada no fluxo de bits com o uso, pelo menos em parte, de um código de Huffman para representar um valor residual de um elemento do vetor.

9. Dispositivo, de acordo com a reivindicação 5, **caracterizado** pelo fato de que compreende adicionalmente:

meios para decompor os dados de áudio ambissônicos de ordem mais elevada para obter o vetor; e

meios para especificar o vetor no fluxo de bits para obter o fluxo de bits.

10. Dispositivo, de acordo com a reivindicação 5, **caracterizado** pelo fato de que compreende adicionalmente:

meios para obter, a partir do fluxo de bits, um objeto de áudio que corresponde ao vetor; e

meios para combinar o objeto de áudio com o vetor para reconstruir os dados de áudio ambissônicos de ordem superior.

11. Dispositivo, de acordo com a reivindicação 5, **caracterizado** pelo fato de que a compressão do vetor inclui a quantização do vetor.

12. Dispositivo, de acordo com a reivindicação 11, **caracterizado** pelo fato de que o indicador compreende um ou mais bits de um elemento de sintaxe que indicam um modo de quantização usado quando se comprime o vetor.

13. Dispositivo, de acordo com a reivindicação 11, **caracterizado** pelo fato de que compreende adicionalmente:

meios para decompor dados de áudio ambissônicos de ordem mais elevada para obter o vetor; e

meios para especificar o vetor no fluxo de bits para obter o fluxo de bits.

14. Dispositivo, de acordo com a reivindicação 11, **caracterizado** pelo fato de que compreende adicionalmente:

meios para obter, a partir do fluxo de bits, um objeto de áudio que corresponde ao vetor; e

meios para combinar o objeto de áudio com o vetor para reconstruir os dados de áudio ambissônicos de ordem superior.

15. Dispositivo, de acordo com a reivindicação 5, **caracterizado** pelo fato de que o um ou mais processadores são configurados para decodificar o fluxo de bits, com base no indicador, e emitir uma pluralidade de alimentações de alto-falante.

16. Dispositivo, de acordo com a reivindicação 15, **caracterizado** pelo fato de que compreende adicionalmente um

ou mais alto-falantes configurados para renderizar a pluralidade de alimentações de alto-falante.

17. Memória **caracterizada** pelo fato de que compreende instruções armazenadas na mesma, as instruções armazenadas por um computador para realizar o método conforme definido em qualquer uma das reivindicações 1 a 4.

 Extensões positivas
 Extensões negativas

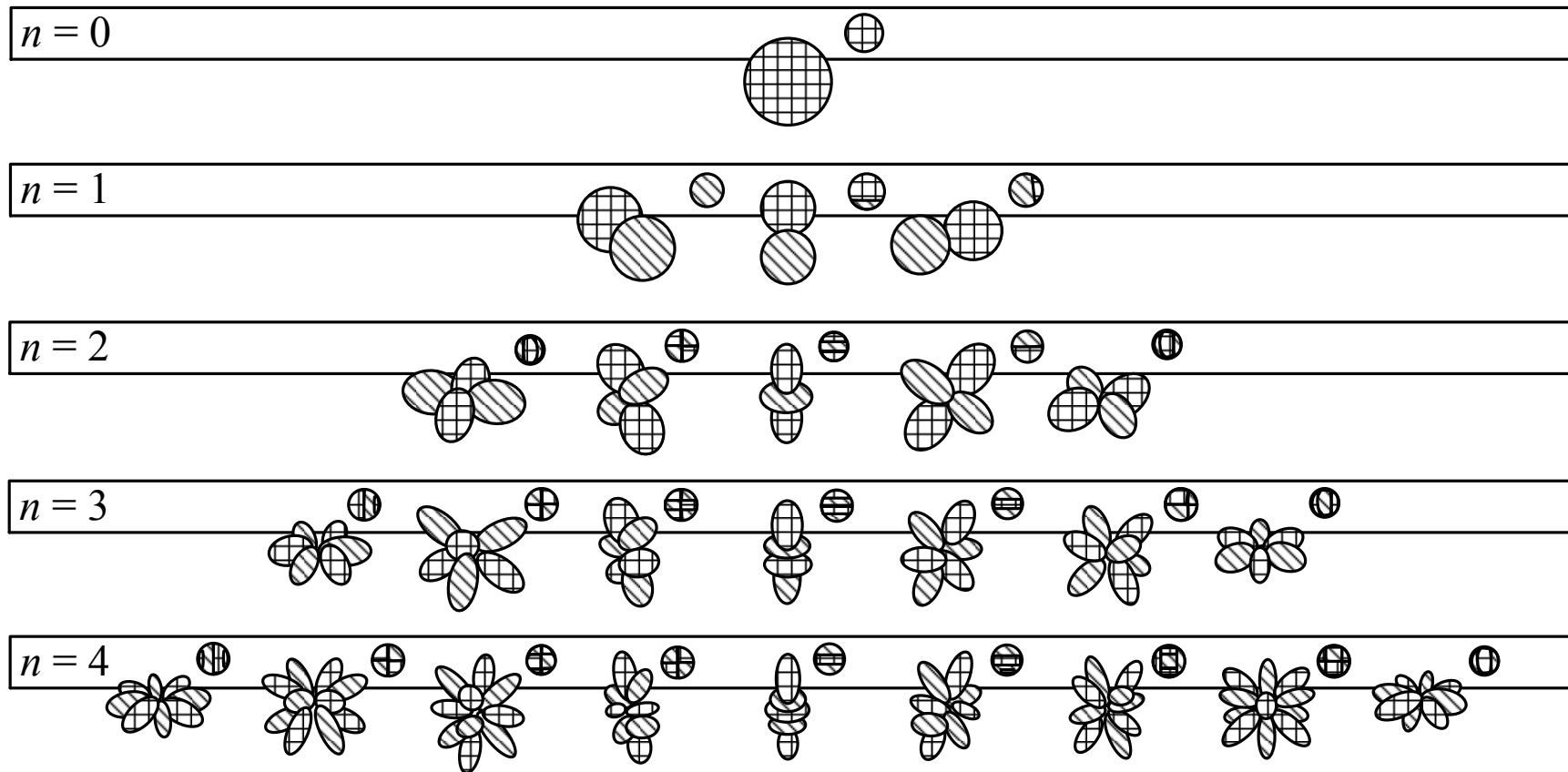


FIG. 1

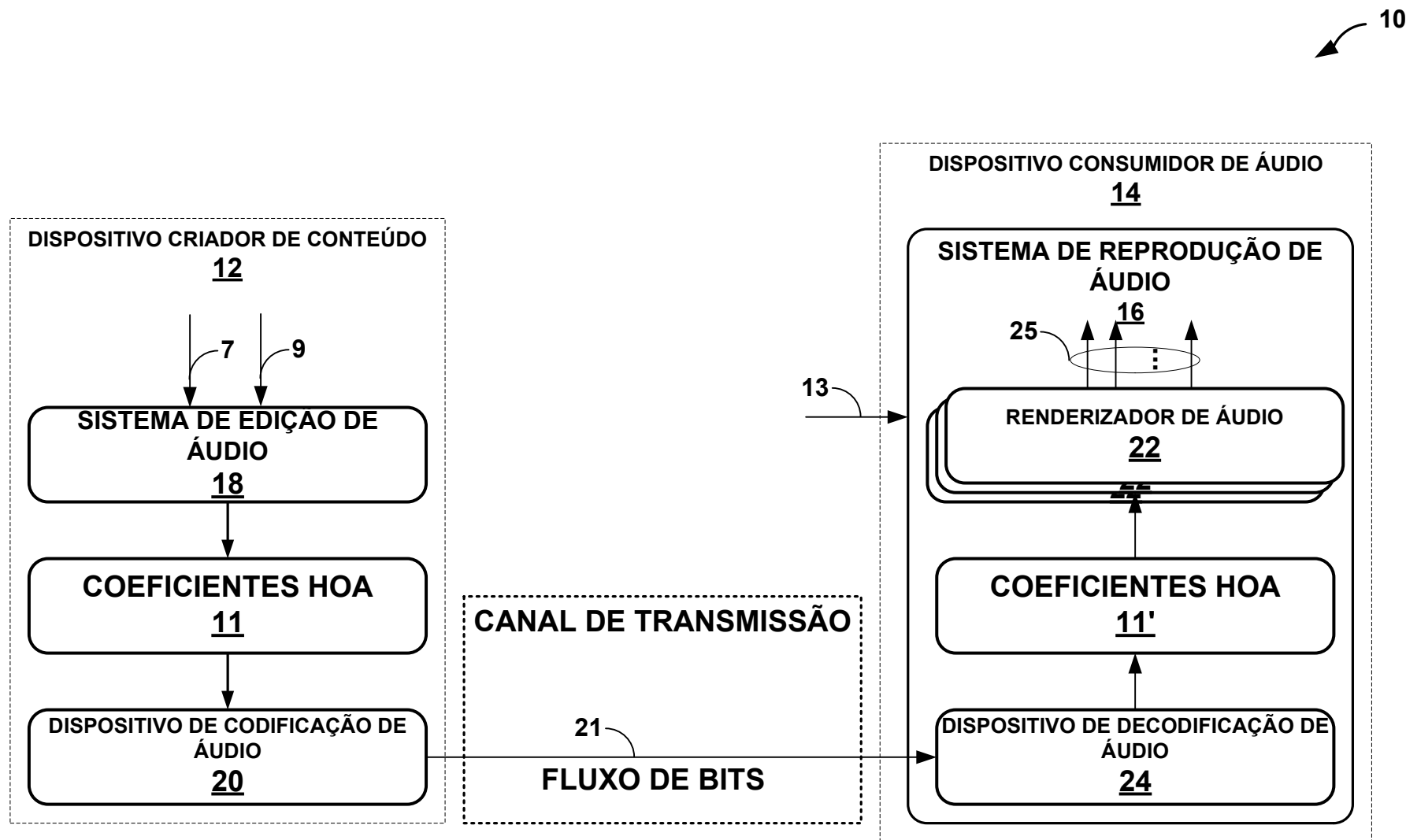


FIG. 2

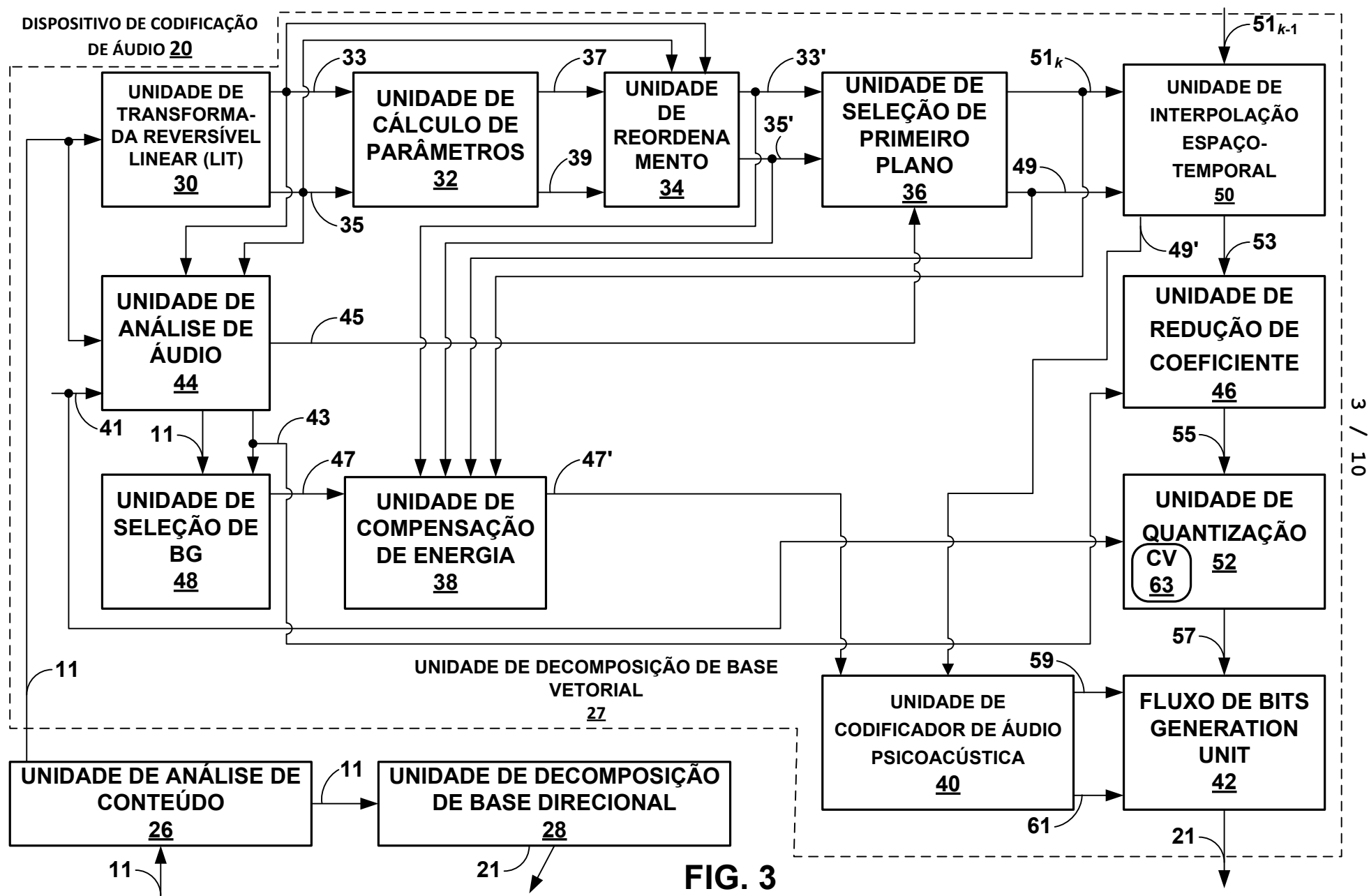


FIG. 3

3 / 10

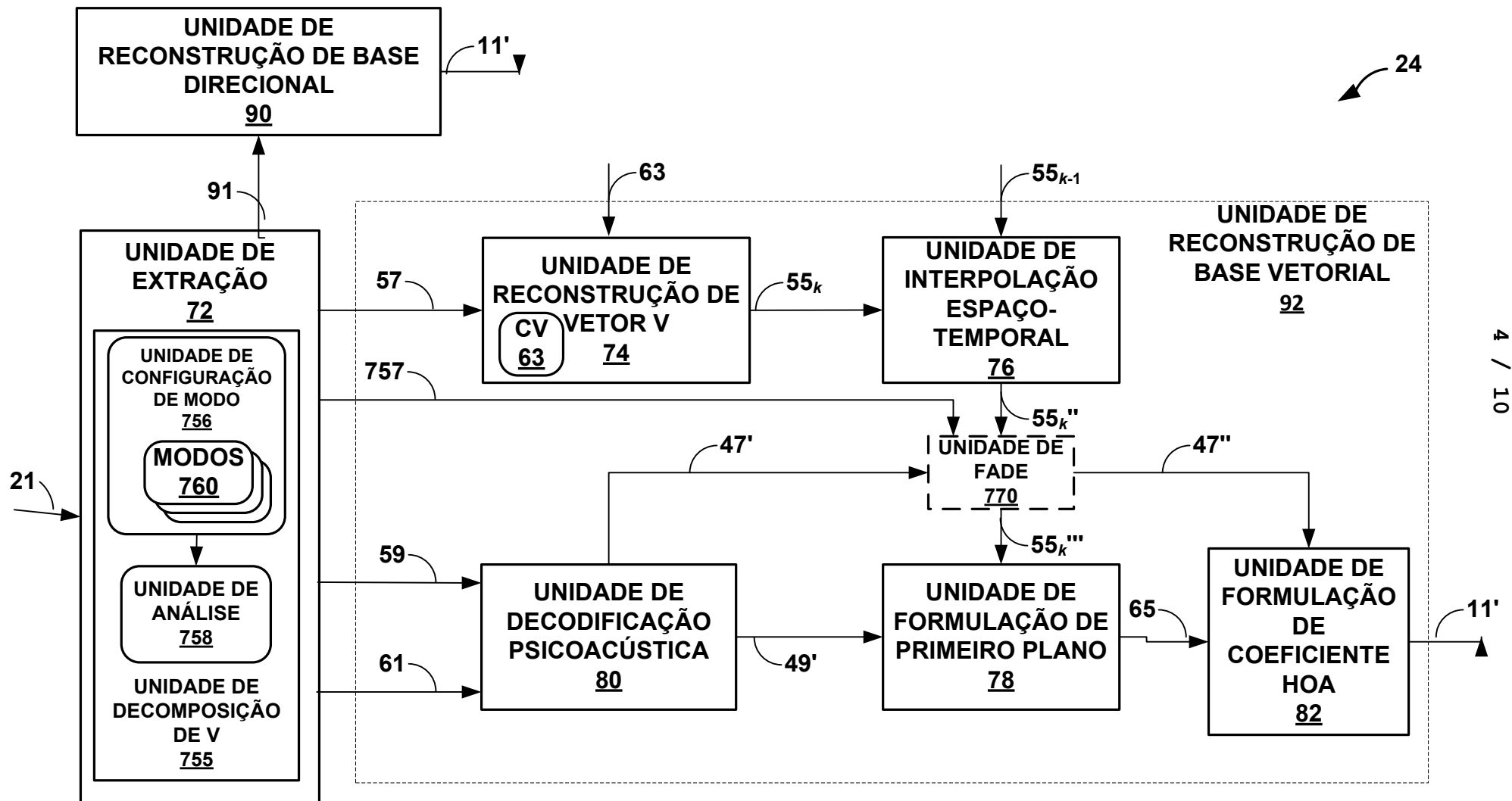


FIG. 4

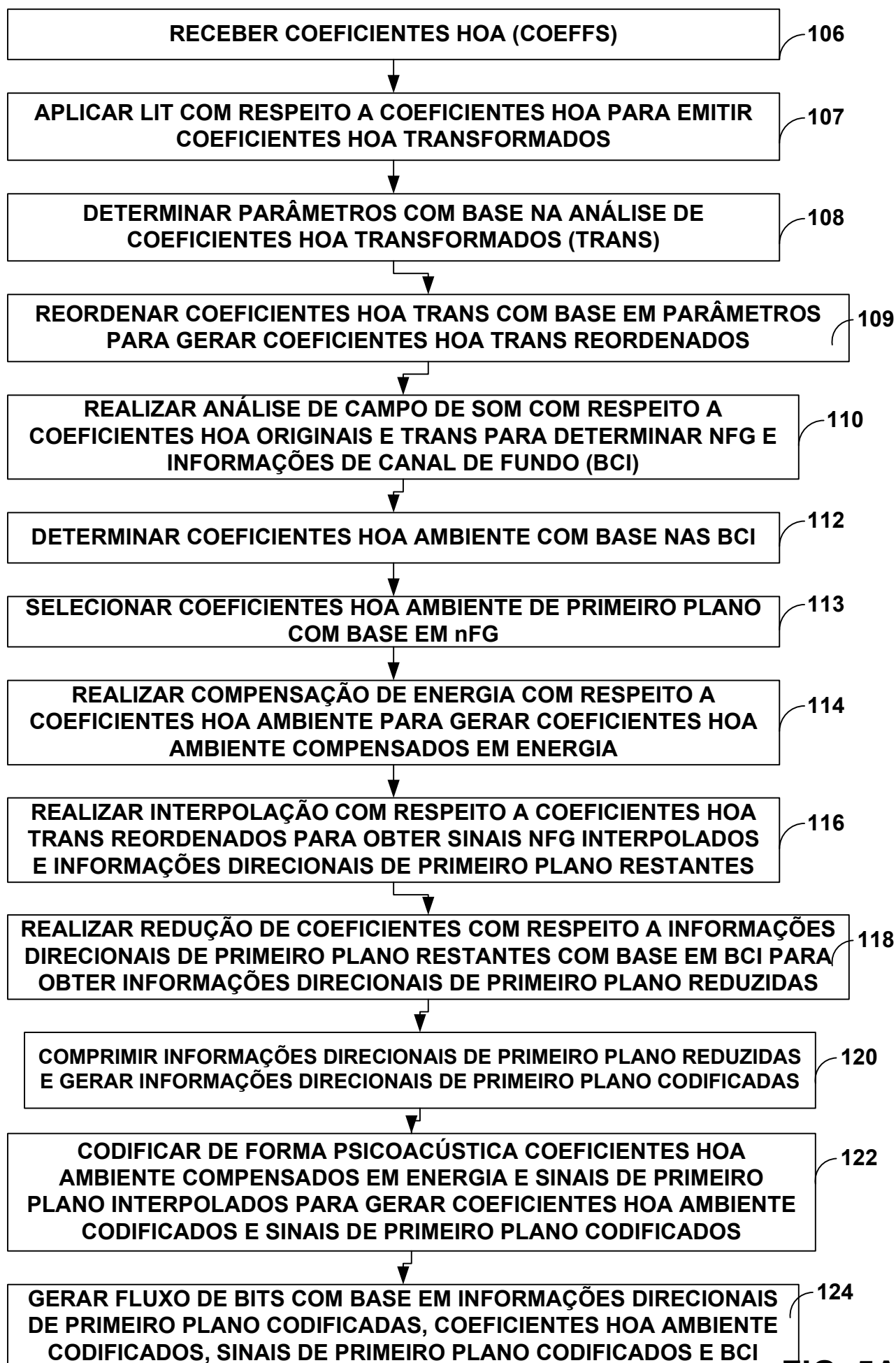


FIG. 5A

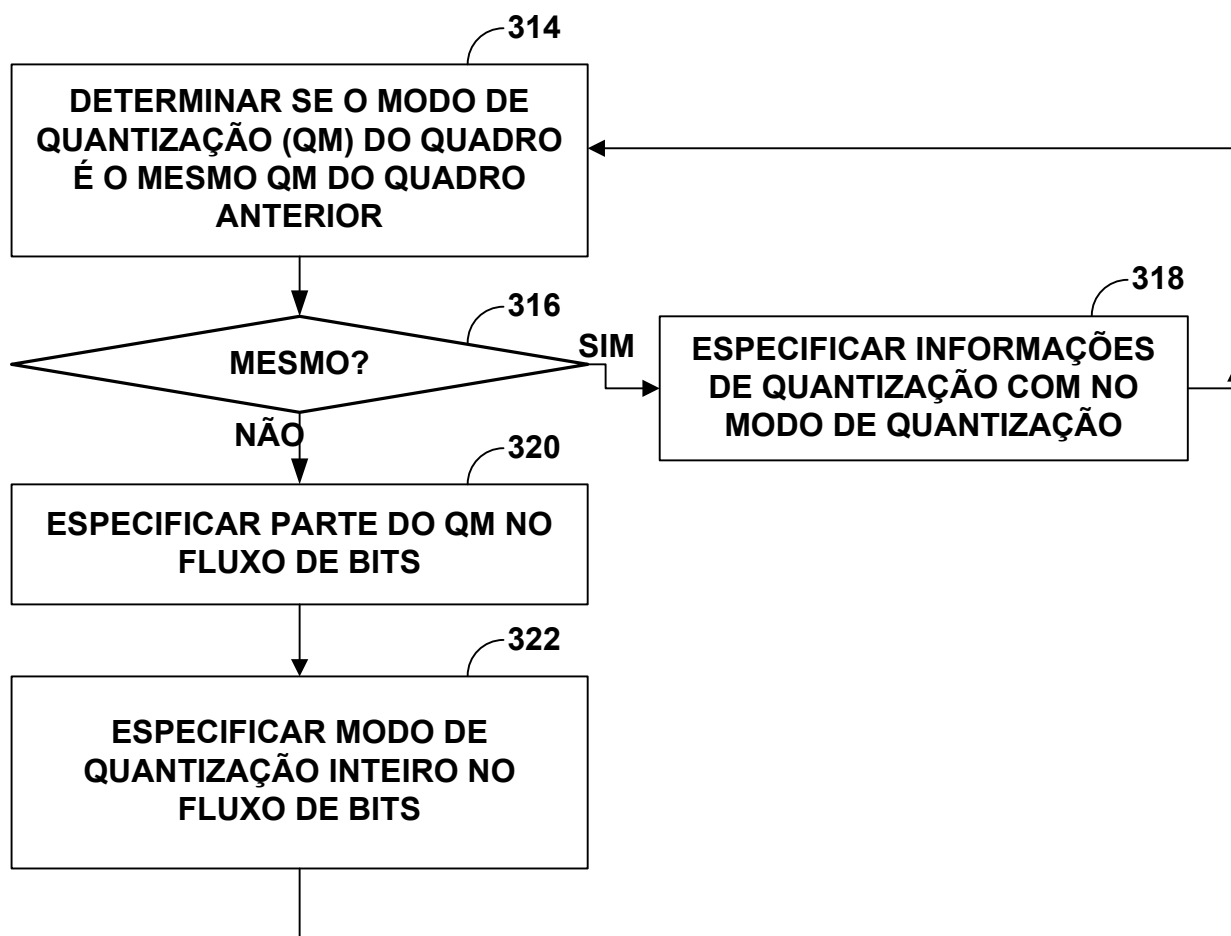


FIG. 5B

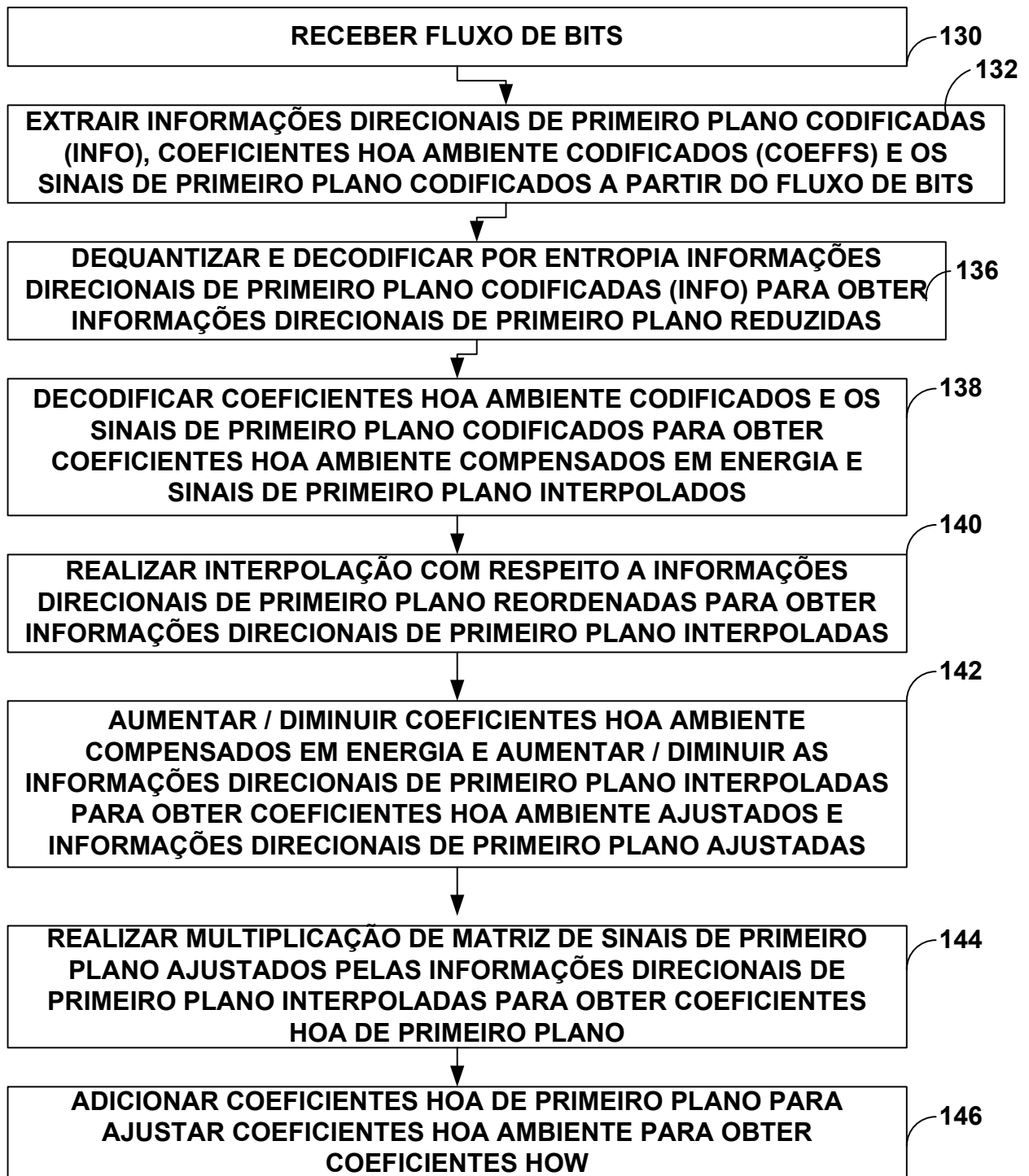


FIG. 6A

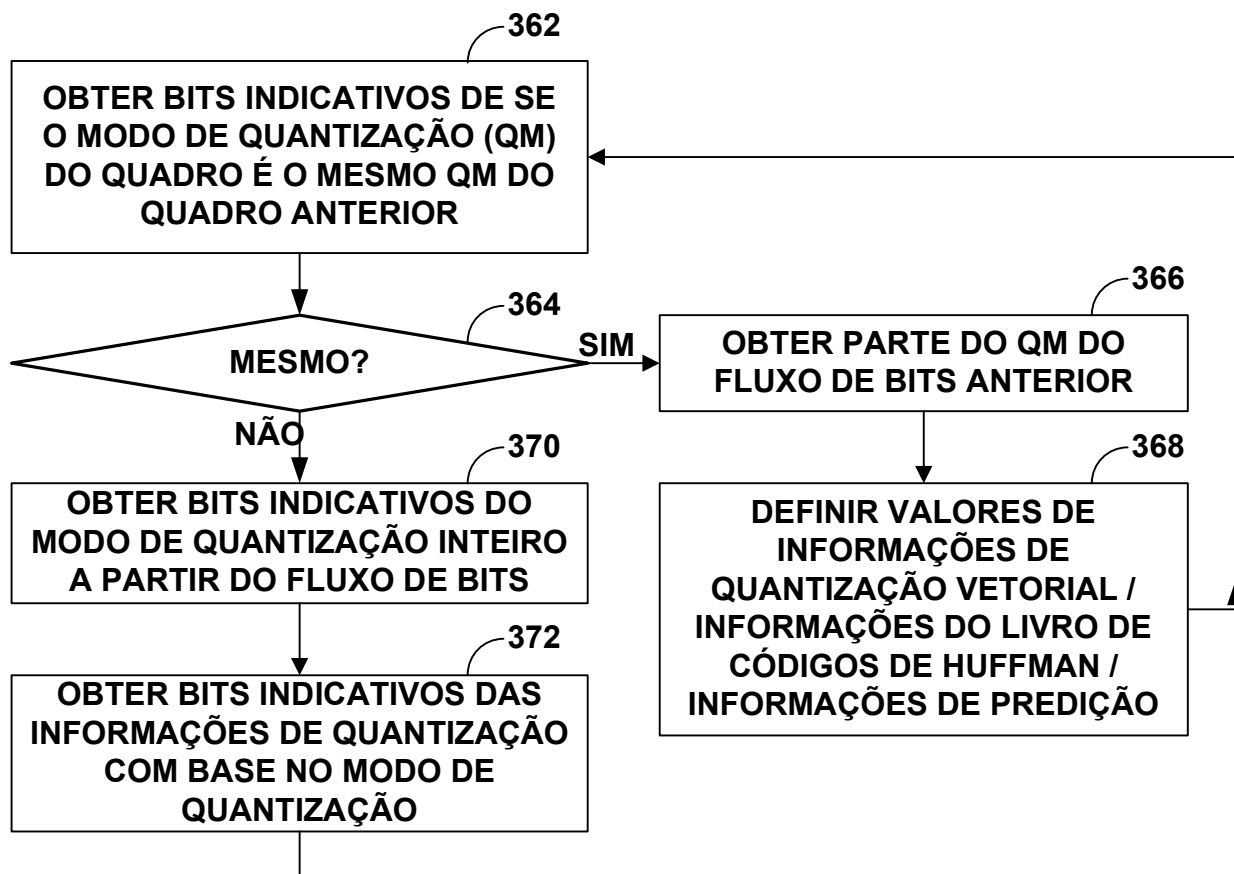
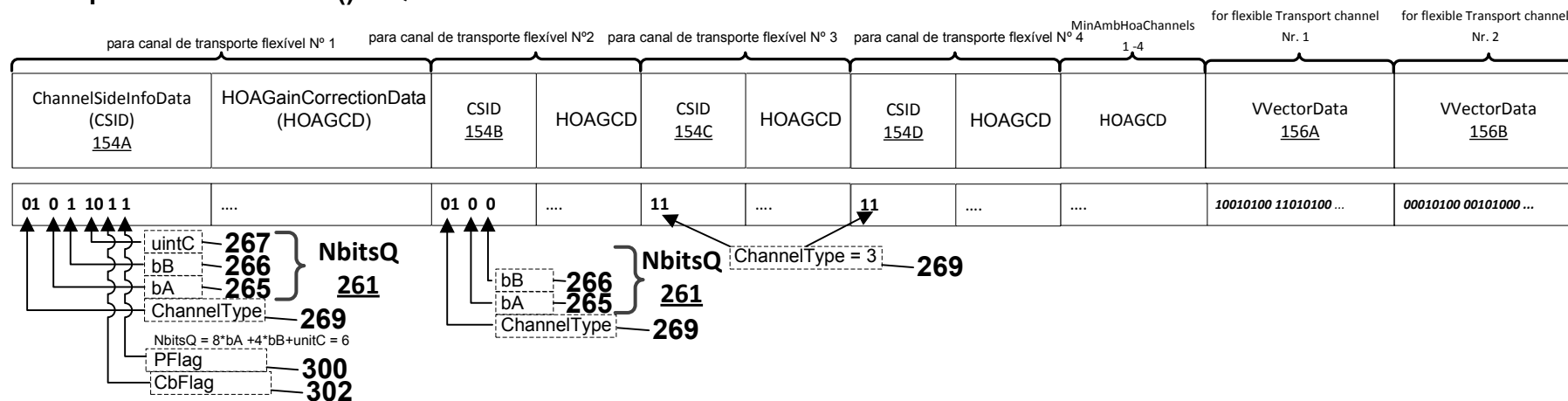


FIG. 6B

249T



249S



9 / 10

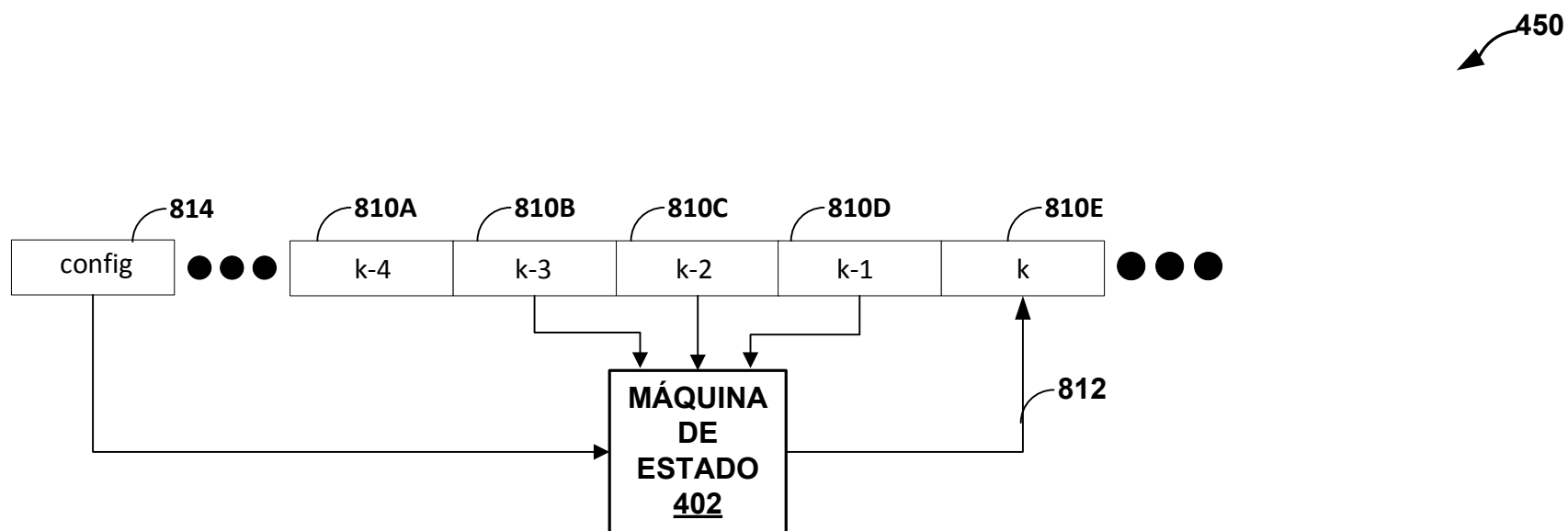


FIG. 8