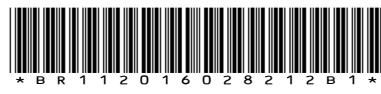




República Federativa do Brasil
Ministério da Economia
Instituto Nacional da Propriedade Industrial

(11) BR 112016028212-4 B1



(22) Data do Depósito: 29/05/2015

(45) Data de Concessão: 23/08/2022

(54) Título: OBTENÇÃO DE INFORMAÇÃO DE SIMETRIA PARA RENDERIZADORES DE ÁUDIO AMBISSÔNICOS DE ORDEM SUPERIOR

(51) Int.Cl.: H04S 7/00; G10L 19/008.

(30) Prioridade Unionista: 11/07/2014 US 62/023,662; 28/05/2015 US 14/724,615; 30/05/2014 US 62/005,829.

(73) Titular(es): QUALCOMM INCORPORATED.

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

(86) Pedido PCT: PCT US2015033273 de 29/05/2015

(87) Publicação PCT: WO 2015/184316 de 03/12/2015

(85) Data do Início da Fase Nacional: 30/11/2016

(57) Resumo: OBTENÇÃO DE INFORMAÇÃO DE SIMETRIA PARA RENDERIZADORES DE ÁUDIO AMBISSÔNICOS DE ORDEM SUPERIOR. Em geral, as técnicas são descritas para obter a informação de renderização de áudio em um bitstream. Um dispositivo configurado para renderizar coeficientes ambissônicos de ordem superior que compreendem um processador e uma memória pode realizar as técnicas. O processador pode ser configurado para obter informação de simetria de sinal indicativa de simetria de sinal de uma matriz utilizada para renderizar os coeficientes ambissônicos de ordem superior para gerar uma pluralidade de alimentações de alto-falantes; e a memória pode ser configurada para armazenar as informações de dispersão.

OBTENÇÃO DE INFORMAÇÃO DE SIMETRIA PARA RENDERIZADORES DE
ÁUDIO AMBISSÔNICOS DE ORDEM SUPERIOR

[0001] Este pedido reivindica o benefício do Pedido provisório de patente US nº 62/023.662, depositado em 11 de julho de 2014, intitulado "SIGNALING AUDIO RENDERING INFORMATION IN A BITSTREAM," e Pedido provisório de patente US nº 62/005.829, depositado em 30 de maio de 2014, intitulado "SIGNALING AUDIO RENDERING INFORMATION IN A BITSTREAM," o conteúdo total de cada um dos Pedidos Provisórios US estão aqui incorporadas, por referência como se aqui apresentado em suas respectivas totalidades.

CAMPO TÉCNICO

[0002] Esta revelação refere-se à renderização de informação e, mais especificamente, renderização da informação para dados de áudio ambissônicos de ordem superior (HOA).

ANTECEDENTES

[0003] Durante a produção de conteúdo de áudio, o engenheiro de som pode renderizar o conteúdo de áudio usando um renderizador específico na tentativa de adaptar o conteúdo de áudio às configurações alvo dos alto-falantes usados para reproduzir o conteúdo de áudio. Em outras palavras, o engenheiro de som pode renderizar o conteúdo de áudio e reproduzir o conteúdo de áudio renderizado usando alto-falantes dispostos na configuração desejada. O engenheiro de som pode então remixar vários aspectos do conteúdo de áudio, renderizar o conteúdo de áudio remixado e reproduzir novamente o conteúdo de áudio renderizado e remixado usando os alto-falantes dispostos na configuração direcionada. O engenheiro de som pode iterar desta forma até que uma certa intenção artística ser fornecida pelo conteúdo de áudio. Desta forma, o engenheiro de som pode produzir conteúdo de áudio que

forneça uma certa intenção artística ou que de outra forma forneça um determinado campo de som durante a reprodução (por exemplo, para acompanhar o conteúdo de vídeo reproduzido juntamente com o conteúdo de áudio).

SUMÁRIO

[0004] Em geral, as técnicas são descritas para especificar a informação de renderização de áudio em um bitstream representativo de dados de áudio. Em outras palavras, as técnicas podem fornecer uma maneira pela qual sinalizar a informação de renderização de áudio utilizada durante a produção de conteúdo de áudio para um dispositivo de reprodução, que pode então utilizar a informação de renderização de áudio para renderizar o conteúdo de áudio. Fornecer a informação de renderização deste modo permite que o dispositivo de reprodução processe o conteúdo de áudio de uma maneira pretendida pelo engenheiro de som e, deste modo, potencialmente assegura a reprodução apropriada do conteúdo de áudio de modo que a intenção artística seja potencialmente compreendida por um ouvinte. Em outras palavras, a informação de renderização utilizada durante a renderização pelo engenheiro de som é fornecida de acordo com as técnicas descritas nesta revelação de modo que o dispositivo de reprodução de áudio possa utilizar a informação de renderização para renderizar o conteúdo de áudio de uma maneira pretendida pelo engenheiro de som, garantindo assim uma experiência mais consistente durante a produção e reprodução do conteúdo de áudio em comparação com sistemas que não fornecem esta informação de renderização de áudio.

[0005] Em um aspecto, um dispositivo configurado para renderizar coeficientes ambissônicos de ordem superior compreende um ou mais processadores configurados para obter informação de esparsidade

indicativa de uma esparsidade de uma matriz utilizada para renderizar os coeficientes ambissônicos de ordem superior para uma pluralidade de feeds de alto-falantes; e uma memória configurada para armazenar as informações de esparsidade.

[0006] Em outro aspecto, um método de renderização de coeficientes ambissônicos de ordem superior compreende obter informação de esparsidade indicativa de uma esparsidade de uma matriz utilizada para renderizar os coeficientes ambissônicos de ordem superior para gerar uma pluralidade de feeds de alto-falantes.

[0007] Em outro aspecto, um dispositivo configurado para produzir um bitstream compreende uma memória configurada para armazenar uma matriz, e um ou mais processadores configurados para obter a informação de esparsidade indicativa de uma esparsidade da matriz usada para renderizar coeficientes ambissônicos de ordem superior para gerar uma pluralidade de feeds de alto-falante.

[0008] Em outro aspecto, um método de produzir um bitstream compreende obter informação de esparsidade indicativa de uma esparsidade de uma matriz utilizada para renderizar os coeficientes ambissônicos de ordem superior para gerar uma pluralidade de feeds de alto-falantes.

[0009] Em outro aspecto, um dispositivo configurado para renderizar coeficientes ambissônicos de ordem superior compreende um ou mais processadores configurados para obter informação de simetria de sinal indicativa da simetria de sinal de uma matriz utilizada para renderizar os coeficientes ambissônicos de ordem superior para gerar uma pluralidade de feeds de alto-falantes; e uma memória configurada para armazenar as informações de esparsidade.

[0010] Em outro aspecto, um método de renderização de coeficientes ambissônicos de ordem superior compreende obter informação de simetria de sinal indicativa da simetria de sinal de uma matriz utilizada para renderizar os coeficientes ambissônicos de ordem superior para gerar uma pluralidade de feeds de alto-falantes.

[0011] Em outro aspecto, um dispositivo configurado para produzir um bitstream compreende uma memória configurada para armazenar uma matriz usada para renderizar o coeficiente ambissônico de ordem superior para gerar uma pluralidade de feeds de alto-falante, e um ou mais processadores configurados para sinalizar a informação de simetria indicativa da simetria de sinal da matriz.

[0012] Em outro aspecto, um método de produzir um bitstream compreende obter informação de esparsidade indicativa de uma esparsidade de uma matriz utilizada para renderizar os coeficientes ambissônicos de ordem superior para gerar uma pluralidade de feeds de alto-falantes.

[0013] Os detalhes de um ou mais aspectos das técnicas são apresentados nos desenhos de acompanhamento e na descrição abaixo. Outras características, objetos e vantagens das técnicas serão evidentes a partir da descrição e desenhos, e a partir das reivindicações.

BREVE DESCRIÇÃO DOS DESENHOS

[0014] A FIG. 1 é um diagrama que ilustra funções de base harmônica esférica de várias ordens e sub-ordens.

[0015] A FIG. 2 é um diagrama ilustrando um sistema que pode realizar vários aspectos das técnicas descritas nesta revelação.

[0016] A FIG. 3 é um diagrama de blocos que ilustra, em mais detalhe, um exemplo do dispositivo de codificação de áudio mostrado no exemplo da Fig. 2 que pode

realizar vários aspectos das técnicas descritas nesta revelação.

[0017] A FIG. 4 é um diagrama de blocos ilustrando o dispositivo de decodificação de áudio da Fig. 2 em maior detalhe.

[0018] A FIG. 5 é um fluxograma ilustrando a operação exemplar de um dispositivo de codificação de áudio na realização de vários aspectos das técnicas de síntese baseada em vetor descritos nesta revelação.

[0019] A FIG. 6 é um fluxograma ilustrando a operação exemplar de um dispositivo de decodificação de áudio na realização de vários aspectos das técnicas descritas nesta revelação.

[0020] A FIG. 7 é um fluxograma que ilustra a operação exemplar de um sistema, como um dos sistemas mostrados nos exemplos das Figs. 2 na realização de vários aspectos das técnicas descritas nesta revelação.

[0021] As FIGs. 8A-8D são diagramas ilustrando bitstreams formados de acordo com as técnicas descritas nesta revelação.

[0022] As FIGs. 8E-8G são diagramas ilustrando porções do bitstream ou informação de canal lateral que podem especificar os componentes espaciais comprimidos em mais detalhe.

[0023] A FIG. 9 é um diagrama ilustrando um exemplo de ganhos min e max dependentes da ordem ambissônica de ordem superior (HOA) dentro de uma matriz de renderização de HOA.

[0024] A FIG. 10 é um diagrama ilustrando uma matriz de renderização HOA de 6^a ordem parcialmente esparsa para 22 alto-falantes.

[0025] A FIG. 11 é um diagrama de fluxo que ilustra a sinalização das propriedades de simetria.

DESCRIÇÃO DETALHADA

[0026] A evolução do som ambiente disponibilizou muitos formatos de saída para o entretenimento hoje em dia. Exemplos de tais formatos de som ambiente para consumidores são em sua maioria "canais", porque eles implicitamente especificam feeds para alto-falantes em certas coordenadas geométricas. Os formatos de som ambiente do consumidor incluem o popular formato 5.1 (que inclui os seguintes seis canais: frente esquerdo (FL), frente direito (FR), centro ou centro frontal, esquerdo traseiro ou esquerdo ambiente, direito traseiro ou direito ambiente e efeitos de baixa frequência (LFE)), o crescente formato 7.1, vários formatos que incluem alto-falantes de altura como o formato 7.1.4 e o formato 22.2 (por exemplo, para uso com o padrão de Televisão de Ultra Alta Definição). Os formatos de não-consumidores podem abranger qualquer número de alto-falantes (em geometrias simétricas e não-simétricas), muitas vezes denominados "arranjos ambiente". Um exemplo desse arranjo inclui 32 alto-falantes posicionados em coordenadas nos cantos de um icosaedro truncado.

[0027] A entrada para um futuro codificador MPEG é opcionalmente um dos três formatos possíveis: (i) áudio baseado em canal tradicional (como discutido acima), que se destina a ser reproduzido através de alto-falantes em posições pré-especificadas; (ii) áudio baseado em objeto, que envolve dados de modulação por código de pulso (PCM) para objetos de áudio individuais com metadados associados contendo as suas coordenadas de localização (entre outras informações); e (iii) áudio baseado em cena, que envolve a representação do campo de som usando coeficientes de funções de base harmônicas esféricas (também chamados de "coeficientes harmônicos esféricos" ou

SHC, "Ambissônicos de Ordem Superior" ou HOA e "coeficientes HOA"). O futuro codificador MPEG pode ser descrito em mais detalhe em um documento intitulado "Call for Proposals for 3D Audio," pela Organização Internacional para Padronização/Comissão Internacional de Eletrotécnica (ISO)/(IEC) JTC1/SC29/WG11/N13411, lançado em Janeiro de 2013 em Genebra, Suíça, e disponível em <http://mpeg.chiariglione.org/sites/default/files/files/standards/parts/docs/w13411.zip>.

[0028] Existem vários formatos de "som ambiente" baseados em canais no mercado. Eles variam, por exemplo, do sistema home theater 5.1 (que tem sido o mais bem-sucedido em termos de fazer incursões em salas de estar além de estéreo) para o 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 de um filme uma vez e não gastar o esforço para remixá-la para cada configuração de alto-falante. Recentemente, as Organizações de Desenvolvimento de Padrões têm considerado modos de fornecer uma codificação para um bitstream padronizado e uma decodificação subsequente que seja adaptável e agnóstica à geometria (e número) do alto-falante e às condições acústicas no local da reprodução (envolvendo um renderizador).

[0029] Para fornecer essa flexibilidade aos 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 nos quais os elementos são ordenados de forma que um conjunto básico de elementos ordenados inferiores fornece uma representação completa do campo de som modelado. À medida que o conjunto é estendido para

incluir elementos de ordem superior, a representação torna-se mais detalhada, aumentando a resolução.

[0030] 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 usando SHC:

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

[0031] A expressão mostra que a pressão p_i em qualquer ponto $\{r_r, \theta_r, \varphi_r\}$ do campo de som, no momento t , pode ser representada unicamente 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 de Bessel esférica da ordem n , e $Y_n^m(\theta_r, \varphi_r)$ são as funções de base harmônica esférica de ordem n e subordem m . Deve-se reconhecer que o termo em colchetes quadrados é uma representação do domínio de frequência do sinal (i.e., $S(\omega, r_r, \theta_r, \varphi_r)$) que pode ser aproximada pelas várias transformadas de tempo-frequência, como a transformada de Fourier discreta (DFT), a transformada de cosseno discreta (DCT) ou uma transformada em ondeleta. Outros exemplos de conjuntos hierárquicos incluem conjuntos de coeficientes de transformada em ondeleta e outros conjuntos de coeficientes de funções de base de multiresolução.

[0032] A FIG. 1 é um diagrama ilustrando funções de base harmônica esférica a partir 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 é mostrada mas não explicitamente observadas no exemplo da FIG. 1 para facilidade dos fins de ilustração.

[0033] O SHC $A_n^m(k)$ pode ou ser fisicamente adquirido (ex., gravado) por várias configurações de

arranjo de microfone ou, alternativamente, eles podem ser derivados de descrições baseada em canal ou baseada em objeto do campo de som. O SHC represente o áudio baseado na cena, onde o SHC pode ser inserido em um codificador de áudio para obter o SHC codificado que pode promover a transmissão ou armazenamento mais eficiente. Por exemplo, uma representação de quarta ordem que envolve coeficientes $(1+4)^2$ (25, e portanto, de quarta ordem) pode ser usada.

[0034] Como observado acima, o SHC pode ser derivado de uma gravação de microfone usando um arranjo de microfone. Vários exemplos de como o SHC pode ser derivado dos arranjos de microfone são descritos em Poletti, M., "Three-Dimensional Surround Sound Systems Based on Spherical Harmonics", J. Audio Eng. Soc., Vol. 53, No. 11, Novembro de 2005, pp. 1004-1025.

[0035] Para ilustrar como os SHCs podem ser derivados de uma descrição baseada em objeto, considere 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:

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

onde 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 fonte de energia do objeto $g(\omega)$ como uma função da frequência (por exemplo, usando técnicas de análise de frequência de tempo, como executar uma transformada de Fourier rápida no fluxo de PCM) nos permite converter cada objeto PCM e o local correspondente no SHC, $A_n^m(k)$. Além disso, pode ser mostrado (uma vez que o acima é uma decomposição linear e ortogonal) que os coeficientes $A_n^m(k)$ para cada objeto são aditivos. Desta maneira, uma multiplicidade de objetos 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 global, na vizinhança do ponto de observação $\{r_r, \theta_r, \varphi_r\}$. As figuras restantes são descritas abaixo no contexto da codificação de áudio baseada em objeto e baseada em SHC.

[0036] A FIG. 2 é um diagrama ilustrando um sistema 10 que pode realizar vários aspectos das técnicas descritas nesta revelação. 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 conteúdos 12 e do dispositivo consumidor de conteúdo 14, as técnicas podem ser implementadas em qualquer contexto em que os SHCs (que também podem ser referidos como coeficientes HOA) ou qualquer outra representação hierárquica de um campo de som são codificados para formar um bitstream 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 nesta revelação, incluindo um telefone móvel (ou telefone celular), um computador tablet, um smartphone ou um computador desktop para fornecer alguns exemplos. De modo similar, o dispositivo consumidor de conteúdo 14 pode representar qualquer forma de dispositivo de computação capaz de implementar as técnicas descritas nesta revelação, incluindo um telefone móvel (ou telefone celular), um computador tablet, um smartphone, um decodificador ou um computador desktop para fornecer alguns exemplos.

[0037] O dispositivo criador de conteúdo 12 pode ser operado por um estúdio de filme ou outra entidade que pode gerar conteúdo de áudio multicanal para consumo pelos operadores de dispositivos de consumo de conteúdo, como o dispositivo consumidor de conteúdo 14. Em alguns exemplos, o dispositivo criador de conteúdo 12 pode ser operado por um usuário individual que gostaria de comprimir os coeficientes 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, que pode se referir a qualquer forma de sistema de reprodução de áudio capaz de renderizar SHC para reprodução como conteúdo de áudio multicanal.

[0038] 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 HOA) e objetos de áudio 9, que o dispositivo criador de conteúdo 12 pode editar utilizando o sistema de edição de áudio 18. Um microfone 5 pode capturar as gravações ao vivo 7. O criador de conteúdo pode, durante o processo de edição, renderizar coeficientes de HOA 11 a partir de objetos de áudio 9, ouvindo os feeds de alto-falante renderizados em uma tentativa de identificar vários aspectos do campo de som que necessitem de mais edição. O dispositivo criador de conteúdo 12 pode então editar os coeficientes HOA 11 (potencialmente indiretamente através da manipulação de diferentes dos objetos de áudio 9 a partir dos quais os coeficientes HOA de fonte podem ser derivados da forma descrita acima). O dispositivo criador de conteúdo 12 pode empregar o sistema de edição de áudio 18 para gerar os

coeficientes HOA 11. O sistema de edição de áudio 18 representa qualquer sistema capaz de editar dados de áudio e produzir os dados de áudio como um ou mais coeficientes de harmônicos esféricos de fonte.

[0039] Quando o processo de edição está completo, o dispositivo criador de conteúdo 12 pode gerar um bitstream 21 com base nos coeficientes HOA 11. Ou seja, 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 os coeficientes HOA 11 de acordo com vários aspectos das técnicas descritas nesta revelação para gerar o bitstream 21. O dispositivo de codificação de áudio 20 pode gerar o bitstream 21 para transmissão, como um exemplo, através de um canal de transmissão, que pode ser um canal com ou sem fios, um dispositivo de armazenamento de dados ou semelhante. O bitstream 21 pode representar uma versão codificada dos coeficientes HOA 11 e pode incluir um bitstream primário e outro bitstream lateral, que podem ser referidos como informação de canal lateral.

[0040] Enquanto mostrado na FIG. 2 como sendo diretamente transmitido para o dispositivo consumidor de conteúdo 14, o dispositivo criador de conteúdo 12 pode emitir o bitstream 21 para um dispositivo intermediário posicionado entre o dispositivo criador de conteúdo 12 e o dispositivo consumidor de conteúdo 14. O dispositivo intermédio pode armazenar o bitstream 21 para distribuição posterior para o dispositivo consumidor de conteúdo 14, o qual pode solicitar o bitstream. O dispositivo intermédio pode compreender um servidor de arquivo, um servidor da web, um computador desktop, um computador portátil, um computador tablet, um telefone móvel, um smartphone ou qualquer outro dispositivo capaz de armazenar o bitstream

21 para posterior recuperação por um decodificador de áudio. O dispositivo intermediário pode residir em uma rede de distribuição de conteúdo capaz de transmitir o fluxo contínuo do bitstream 21 (e possivelmente em conjunto com a transmissão de um fluxo de dados de dados de vídeo correspondente) aos assinantes, como o dispositivo consumidor de conteúdo 14, solicitando o bitstream 21.

[0041] Alternativamente, o dispositivo criador de conteúdo 12 pode armazenar o bitstream 21 em uma mídia de armazenamento, como um disco compacto, um disco de vídeo digital, um disco de vídeo de alta definição ou outra mídia de armazenamento, a maioria das quais são capazes de ser lidas por um computador e portanto, podem ser referidas como mídias de armazenamento legíveis por computador ou mídias de armazenamento legíveis por computador não transitórias. Neste contexto, o canal de transmissão pode se referir aos canais pelos quais o conteúdo armazenado para as mídias é transmitido (e pode incluir lojas de varejo e outro mecanismo de distribuição baseado em loja). Em qualquer caso, as técnicas desta descrição não devem, portanto, ser limitadas a este respeito ao exemplo da FIG. 2.

[0042] Como mostrado adicionalmente no exemplo da FIG. 2, o dispositivo criador de conteúdo 14 inclui um 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 um número de renderizadores diferentes 22. Os renderizadores 22 podem proporcionar, cada um, uma forma diferente de renderização, onde as diferentes formas de renderização podem incluir uma ou mais das várias formas de realizar o posicionamento de amplitude a base de vetor (VBAP) e/ou uma ou mais das

várias formas de realizar a síntese do campo de som. Como aqui utilizado, "A e/ou B" significa "A ou B", ou ambos "A e B".

[0043] 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 os coeficientes HOA 11' a partir do bitstream 21, onde os coeficientes de HOA 11' podem ser semelhantes aos coeficientes de HOA 11 mas diferem devido a operações com perdas (por exemplo, quantização) e/ou através do canal de transmissão. O sistema de reprodução de áudio 16 pode, depois de decodificar o bitstream 21 obter os coeficientes de HOA 11' e renderizar os coeficientes de HOA 11' para fornecer feeds de alto-falante de saída 25. Os feeds de alto-falantes 25 podem acionar um ou mais alto-falantes (que não estão ilustrados no exemplo da FIG. 2 para facilidade dos fins de ilustração).

[0044] Para selecionar o processador apropriado ou, em alguns casos, gerar um renderizador apropriado, o sistema de reprodução de áudio 16 pode obter informação de alto-falante 13 indicativa de um número de alto-falantes e/ou uma geometria espacial dos alto-falantes. Em alguns casos, o sistema de reprodução de áudio 16 pode obter a informação de alto-falante 13 utilizando um microfone de referência e acionando os alto-falantes de modo a determinar dinamicamente a informação do alto-falante 13. Em outros casos ou em conjunto com a determinação dinâmica da informação do alto-falante 13, o sistema de reprodução de áudio 16 pode pedir a um usuário que interaja com o sistema de reprodução de áudio 16 e introduza a informação de alto-falante 13.

[0045] O sistema de reprodução de áudio 16 pode então selecionar um dos renderizadores de áudio 22 com base na informação do alto-falante 13. Em alguns casos, o sistema de reprodução de áudio 16 pode, quando nenhum dos renderizadores de áudio 22 está dentro de uma medida de similaridade de limite (em termos da geometria de alto-falante) para a geometria de alto-falante especificada na informação de alto-falante 13, gerar um dos renderizadores de áudio 22 com base nas informações 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 de alto-falante 13 sem primeiro tentar selecionar um existente dos renderizadores de áudio 22. Um ou mais alto-falantes 3 podem então reproduzir os feeds de alto-falante renderizados 25.

[0046] Em alguns casos, o sistema de reprodução de áudio 16 pode selecionar qualquer um dos renderizadores de áudio 22 e pode ser configurado para selecionar um ou mais dos renderizadores de áudio 22 dependendo da fonte a partir da qual o bitstream 21 é recebido (como um leitor de DVD, um leitor Blu-ray, um smartphone, um computador tablet, um sistema de jogos e uma televisão para fornecer alguns exemplos). Embora qualquer um dos renderizadores de áudio 22 possa ser selecionado, muitas vezes o processador de áudio utilizado ao criar o conteúdo fornece uma forma melhor (e possivelmente a melhor) de renderização devido ao fato de o conteúdo ter sido criado pelo criador de conteúdo 12 utilizando este processador de áudio específico, isto é, o processador de áudio 5 no exemplo da FIG. 3. Selecionar um dos renderizadores de áudio 22 que é o mesmo ou pelo menos próximo (em termos de forma de renderização) pode proporcionar uma melhor representação do campo de som e

pode resultar em uma melhor experiência de som ambiente para o consumidor de conteúdos 14.

[0047] De acordo com as técnicas descritas nesta revelação, o dispositivo de codificação de áudio 20 pode gerar o bitstream 21 para incluir a informação de renderização de áudio 2 ("render info 2"). A informação de renderização de áudio 2 pode incluir um valor de sinal que identifica um renderizador de áudio utilizado quando se gera o conteúdo de áudio multicanal, isto é, o renderizador de áudio 1 no exemplo da FIG. 3. Em alguns casos, o valor de sinal inclui uma matriz usada para renderizar coeficientes harmônicos esféricos para uma pluralidade de feeds de alto-falante.

[0048] Em alguns casos, o valor de sinal inclui dois ou mais bits que definem um índice que indica que o bitstream inclui uma matriz utilizada para renderizar coeficientes harmônicos esféricos para uma pluralidade de feeds de alto-falante. Em alguns casos, quando é utilizado um índice, o valor de sinal inclui ainda dois ou mais bits que definem um número de linhas da matriz incluída no bitstream e dois ou mais bits que definem um número de colunas da matriz incluída no bitstream. Usando estas informações e dado que cada coeficiente da matriz bidimensional é tipicamente definido por um número de ponto flutuante de 32 bits, o tamanho em termos de bits da matriz pode ser calculado como uma função do número de linhas, o número de colunas e o tamanho dos números de ponto flutuante que definem cada coeficiente da matriz, ou seja, 32 bits neste exemplo.

[0049] Em alguns casos, o valor de sinal especifica um algoritmo de renderização usado para renderizar coeficientes harmônicos esféricos para uma pluralidade de feeds de alto-falante. O algoritmo de

renderização pode incluir uma matriz que é conhecida tanto pelo dispositivo de codificação de áudio 20 como pelo dispositivo de decodificação 24. Isto é, o algoritmo de renderização pode incluir a aplicação de uma matriz além de outras etapas de renderização, como posicionamento (por exemplo, VBAP, DBAP ou posicionamento simples) ou filtragem NFC. Em alguns casos, o valor de sinal inclui dois ou mais bits que definem um índice associado com uma de uma pluralidade de matrizes utilizadas para renderizar coeficientes harmônicos esféricos para uma pluralidade de feeds de alto-falante. Mais uma vez, tanto o dispositivo de codificação de áudio 20 como o dispositivo de decodificação 24 podem ser configurados com informação que indica a pluralidade de matrizes e a ordem da pluralidade de matrizes de modo que o índice pode identificar de forma única uma matriz particular da pluralidade de matrizes. Alternativamente, o dispositivo de codificação de áudio 20 pode especificar dados no bitstream 21 que definem a pluralidade de matrizes e/ou a ordem da pluralidade de matrizes de modo que o índice pode identificar de forma única uma matriz particular da pluralidade de matrizes.

[0050] Em alguns casos, o valor de sinal inclui dois ou mais bits que definem um índice associado com uma de uma pluralidade de algoritmos de renderização utilizados para renderizar coeficientes harmônicos esféricos para uma pluralidade de feeds de alto-falante. Mais uma vez, tanto o dispositivo de codificação de áudio 20 como o dispositivo de decodificação 24 podem ser configurados com informação que indica a pluralidade de algoritmos de renderização e a ordem da pluralidade de algoritmos de renderização de modo que o índice pode identificar de forma única uma matriz particular da pluralidade de matrizes. Alternativamente, o dispositivo

de codificação de áudio 20 pode especificar dados no bitstream 21 que definem a pluralidade de matrizes e/ou a ordem da pluralidade de matrizes de modo que o índice pode identificar de forma única uma matriz particular da pluralidade de matrizes.

[0051] Em alguns casos, o dispositivo de codificação de áudio 20 especifica a informação de renderização de áudio 2 em uma base de quadro de áudio no bitstream. Em outros casos, o dispositivo de codificação de áudio 20 especifica a informação de renderização de áudio 2 uma única vez no bitstream.

[0052] O dispositivo de decodificação 24 pode então determinar a informação de renderização de áudio 2 especificada no bitstream. Com base no valor de sinal incluído na informação de renderização de áudio 2, o sistema de reprodução de áudio 16 pode renderizar uma pluralidade de feeds de alto-falante 25 com base na informação de renderização de áudio 2. Como observado acima, o valor de sinal pode, em alguns casos, incluir uma matriz usada para renderizar coeficientes harmônicos esféricos para uma pluralidade de feeds de alto-falante. Neste caso, o sistema de reprodução de áudio 16 pode configurar um dos renderizadores de áudio 22 com a matriz, utilizando este um dos renderizadores de áudio 22 para renderizar os feeds de alto-falante 25 com base na matriz.

[0053] Em alguns casos, o valor de sinal inclui dois ou mais bits que definem um índice que indica que o bitstream inclui uma matriz utilizada para renderizar os coeficientes HOA 11' para os feeds de alto-falante 25. O dispositivo de decodificação 24 pode analisar a matriz a partir do bitstream em resposta ao índice, após o que o sistema de reprodução de áudio 16 pode configurar um dos renderizadores de áudio 22 com a matriz analisada e invocar

este um dos processadores 22 para renderizar os feeds de alto-falante 25. Quando o valor de sinal inclui dois ou mais bits que definem um número de linhas da matriz incluída no bitstream e dois ou mais bits que definem um número de colunas da matriz incluída no bitstream, o dispositivo de decodificação 24 pode analisar a matriz do bitstream em resposta ao índice e com base nos dois ou mais bits que definem um número de linhas e os dois ou mais bits que definem o número de colunas da maneira descrita acima.

[0054] Em alguns casos, o valor de sinal especifica um algoritmo de renderização usado para renderizar os coeficientes HOA 11' para os feeds de alto-falante 25. Nestes casos, alguns ou todos os renderizadores de áudio 22 podem executar estes algoritmos de renderização. O dispositivo de reprodução de áudio 16 pode então utilizar o algoritmo de renderização especificado, por exemplo, um dos renderizadores de áudio 22, para renderizar os feeds de alto-falante 25 a partir dos coeficientes HOA 11'.

[0055] Quando o valor de sinal inclui dois ou mais bits que definem um índice associado a uma de uma pluralidade de matrizes utilizadas para renderizar os coeficientes HOA 11' para os feeds de alto-falantes 25, alguns ou todos os renderizadores de áudio 22 podem representar esta pluralidade de matrizes. Deste modo, o sistema de reprodução de áudio 16 pode renderizar os feeds de alto-falante 25 a partir dos coeficientes de HOA 11' utilizando um dos renderizadores de áudio 22 associados ao índice.

[0056] Quando o valor de sinal inclui dois ou mais bits que definem um índice associado a uma de uma pluralidade de algoritmos de renderização utilizados para renderizar os coeficientes HOA 11' para os feeds de alto-

falantes 25, alguns ou todos os renderizadores de áudio 34 podem representar estes algoritmos de renderização. Deste modo, o sistema de reprodução de áudio 16 pode renderizar os feeds de alto-falante 25 a partir dos coeficientes harmônicos esféricos 11' utilizando um dos renderizadores de áudio 22 associados ao índice.

[0057] Dependendo da frequência com que esta informação de renderização de áudio é especificada no bitstream, o dispositivo de decodificação 24 pode determinar a informação de renderização de áudio 2 em uma base por quadro de áudio ou de uma única vez.

[0058] Especificando a informação de renderização de áudio 3 desta maneira, as técnicas podem potencialmente resultar em uma melhor reprodução do conteúdo de áudio multicanal e de acordo com a maneira pela qual o criador de conteúdo 12 pretendia que o conteúdo de áudio multicanal fosse reproduzido. Como resultado, as técnicas podem fornecer um som ambiente ou experiência de áudio multicanal mais imersiva.

[0059] Em outras palavras e como observado acima, o Ambissônicos de Ordem Superior (HOA) podem representar uma maneira pela qual descrever a informação direcional de um campo de som baseado em uma transformada de Fourier espacial. Tipicamente, quanto maior a ordem de Ambissônicos N, maior a resolução espacial, maior o número de coeficientes harmônicos esféricos (SH) $(N+1)^2$ e maior a largura de banda necessária para transmitir e armazenar os dados.

[0060] Uma vantagem potencial desta descrição é a possibilidade de reproduzir este campo de som na maioria de qualquer das configurações de alto-falante (por exemplo, 5.1, 7.1 22.2, etc.). A conversão da descrição do campo de som em M sinais de alto-falante pode ser feita

através de uma matriz de renderização estática com $(N+1)^2$ entradas e M saídas. Consequentemente, cada configuração de alto-falante pode exigir uma matriz de renderização dedicada. Podem existir vários algoritmos para calcular a matriz de renderização para uma configuração de alto-falante desejada, que pode ser otimizada para determinadas medidas objetivas ou subjetivas, como os critérios de Gerzon. Para configurações de alto-falantes irregulares, os algoritmos podem tornar-se complexos devido aos procedimentos de otimização numéricos iterativos, como a otimização convexa. Para calcular uma matriz de renderização para layouts de alto-falante irregulares sem tempo de espera, pode ser benéfico ter recursos de computação suficientes disponíveis. Configurações de alto-falante irregulares podem ser comuns em ambientes domésticos de sala de estar devido a restrições arquitetônicas e preferências estéticas. Portanto, para a melhor reprodução de campos sonoros, uma matriz de renderização otimizada para tal cenário pode ser preferida na medida em que pode permitir a reprodução do campo de som com mais precisão.

[0061] Como um decodificador de áudio geralmente não requer muitos recursos computacionais, o dispositivo pode não ser capaz de calcular uma matriz de renderização irregular em um tempo amigável ao consumidor. Vários aspectos das técnicas descritas nesta revelação podem proporcionar a utilização de uma abordagem de computação baseada na nuvem como a seguir:

1. O decodificador de áudio pode enviar através de uma conexão de Internet as coordenadas do alto-falante (e, em alguns casos, também as medições de SPL obtidas com um microfone de calibração) para um servidor;

2. O servidor baseado na nuvem pode calcular a matriz de renderização (e possivelmente algumas versões diferentes, para que o cliente possa mais tarde escolher entre essas diferentes versões); e

3. O servidor pode então enviar a matriz de renderização (ou as diferentes versões) de volta para o decodificador de áudio via conexão à Internet.

[0062] Essa abordagem pode permitir que o fabricante mantenha os custos de fabricação de um decodificador de áudio baixo (porque um processador poderoso pode não ser necessário para calcular essas matrizes de renderização irregulares), ao mesmo tempo que facilita uma reprodução de áudio mais ideal em comparação com as matrizes de renderização normalmente projetadas para configurações ou geometrias de alto-falantes. O algoritmo para calcular a matriz de renderização também pode ser otimizado depois que um decodificador de áudio ser enviado, potencialmente reduzindo os custos para revisões de hardware ou até mesmo recalls. As técnicas também podem, em alguns casos, reunir muita informação sobre diferentes configurações de alto-falantes de produtos de consumidor que podem ser benéficas para futuros desenvolvimentos de produto.

[0063] Em alguns casos, o sistema mostrado na FIG. 3 pode não sinalizar a informação de renderização de áudio 2 no bitstream 21 como descrito acima, mas sim sinalizar esta informação de renderização de áudio 2 como metadados separados do bitstream 21. Alternativamente ou em conjunto com o descrito acima, o sistema mostrado na FIG. 3 pode sinalizar uma porção da informação de renderização de áudio 2 no bitstream 21 como descrito acima e sinalizar uma porção desta informação de renderização de áudio 3 como metadados separados do bitstream 21. Em

alguns exemplos, o dispositivo de codificação de áudio 20 pode produzir estes metadados, que podem então ser carregados para um servidor ou outro dispositivo. O dispositivo de decodificação de áudio 24 pode então descarregar ou de outro modo recuperar estes metadados, que são então utilizados para aumentar a informação de renderização de áudio extraída do bitstream 21 pelo dispositivo de decodificação de áudio 24. O bitstream 21 formado de acordo com os aspectos de informação de renderização das técnicas é descrito abaixo em relação aos exemplos das FIGs. 8A-8D.

[0064] A FIG. 3 é um diagrama de blocos que ilustra, em mais detalhe, um exemplo do dispositivo de codificação de áudio 20 mostrado no exemplo da Fig. 2 que pode realizar vários aspectos das técnicas descritas nesta revelação. 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 baseada em direção 28. Embora brevemente descrito a seguir, mais informação sobre o dispositivo de codificação de áudio 20 e os vários aspectos da compressão ou de outro modo codificação de coeficientes de HOA estão disponíveis na Publicação de Pedido de Patente Internacional No. WO 2014/194099, intitulada "INTERPOLATION FOR DECOMPOSED REPRESENTATIONS OF A SOUND FIELD" depositado em 29 de maio de 2014.

[0065] A unidade de análise de conteúdo 26 representa uma unidade configurada para analisar o conteúdo dos coeficientes HOA 11 para identificar se os coeficientes HOA 11 representam o conteúdo gerado a partir de uma gravação ao vivo ou 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 enquadrados 11 foram gerados a partir de uma gravação, a unidade de análise de conteúdo 26 passa os coeficientes HOA 11 para a unidade de decomposição baseada em vetor 27. Em alguns casos, quando os coeficientes de HOA enquadrados 11 foram gerados a partir de um objeto de áudio sintético, a unidade de análise de conteúdo 26 passa os coeficientes HOA 11 para a unidade de decomposição baseada em direção 28. A unidade de síntese com base direcional 28 pode representar uma unidade configurada para executar uma síntese de base direcional dos coeficientes HOA 11 para gerar um bitstream com base direcional 21.

[0066] Como mostrado no exemplo da FIG. 3, a unidade de decomposição baseada em vetor 27 pode incluir uma unidade de transformada inversível linear (LIT) 30, uma unidade de cálculo de parâmetros 32, uma unidade de reordenação 34, uma unidade de seleção de primeiro plano 36, uma unidade de compensação de energia 38, uma unidade de codificador de áudio psicoacústico 40, uma unidade de geração de bitstream 42, uma unidade de análise de campo de som 44, uma unidade de redução de coeficiente 46, uma unidade de seleção de fundo (BG) 48, uma unidade de interpolação espaço-temporal 50 e uma unidade de quantização 52.

[0067] A unidade de transformada inversível linear (LIT) 30 recebe os coeficientes HOA 11 na forma dos canais HOA, cada canal representativo de um bloco ou quadro de um coeficiente associado com uma dada ordem, sub-ordem das funções de base esféricas (que podem ser denotadas como HOA[k], onde k pode denotar o quadro atual ou bloco de amostras). A matriz de coeficientes HOA 11 pode ter dimensões D: M x (N+1)².

[0068] A unidade LIT 30 pode representar uma unidade configurada para realizar uma forma de análise referida como uma decomposição de valor singular. Embora descritas em relação à SVD, as técnicas descritas nesta descrição podem ser realizadas com relação a qualquer transformação ou decomposição semelhante que proporcione conjuntos de saída compactada de energia, linearmente não correlacionados. Também, referência aos "conjuntos" nesta revelação é geralmente pretendida para se referir aos conjuntos não nulos, a menos que especificamente indicado em contrário e não se destina a fazer referência à definição matemática clássica de conjuntos que inclui o chamado "conjunto vazio". Uma transformada alternativa pode compreender uma análise de componente principais, que é frequentemente referida como "PCA". Dependendo do contexto, a PCA pode ser referida por vários nomes diferentes, tais como a transformada de Karhunen-Loeve discreta, a transformada de Hotelling, decomposição ortogonal adequada (POD) e decomposição de valores próprios (EVD) para citar alguns exemplos. As propriedades de tais operações que são conducentes à meta subjacente de compressão de dados de áudio são "compactação de energia" e "decodificação" dos dados de áudio multicanal.

[0069] Em qualquer caso, assumindo que a unidade LIT 30 executa uma decomposição de valor singular (que, novamente, pode ser referida como "SVD") para efeitos de exemplo, a unidade LIT 30 pode transformar os coeficientes HOA 11 em dois ou mais conjuntos de coeficientes HOA transformados. Os "conjuntos" de coeficientes HOA transformados pode incluir vetores de coeficientes HOA transformados. No exemplo da FIG. 3, a unidade LIT 30 pode executar a SVD em relação aos coeficientes HOA 11 para gerar uma assim chamada matriz V,

uma matriz S e uma matriz U. SVD, em álgebra linear, pode representar uma fatorização de uma matriz real ou complexa X de y-por-z (onde X pode representar dados de áudio multicanal, como os coeficientes de HOA 11) da seguinte forma:

$$X = USV^*$$

U pode representar uma matriz unitária real ou complexa y-por-y, onde as colunas y de U são conhecidas como os vetores singulares esquerdos dos dados de áudio multicanal. S pode representar uma matriz diagonal y-por-z com números reais não negativos na diagonal, onde os valores diagonais de S são conhecidos como os valores singulares dos dados de áudio multicanal. V^* (que pode denotar uma transposição de V) pode representar uma matriz unitária real ou complexa z-por-z, onde as colunas z de V^* são conhecidas como os vetores singulares direitos dos dados de áudio multicanal.

[0070] Em alguns exemplos, a matriz V^* na expressão matemática SVD referida acima é designada como a transposição conjugada da matriz V para refletir que a SVD pode ser aplicada às matrizes que compreendem números complexos. Quando aplicado às matrizes que compreendem somente números reais, o conjugado complexo da matriz V (ou, em outras palavras, a matriz V^*) pode ser considerado como a transposição da matriz V. Abaixo, assume-se, para facilidade de ilustração, que os coeficientes HOA 11 compreendem números reais, com o resultado de que a matriz V é produzida através de SVD em vez da matriz V^* . Além disso, embora denotada como a matriz V nesta descrição, referência à matriz V deve ser entendida como referindo-se à transposição da matriz V quando apropriado. Embora se presuma que seja a matriz V, as técnicas podem ser aplicadas de uma forma semelhante aos coeficientes HOA 11

com coeficientes complexos, onde a saída da SVD é a matriz V^* . Consequentemente, as técnicas não devem ser limitadas a este respeito apenas para fornecer a aplicação de SVD para gerar uma matriz V , mas podem incluir a aplicação de SVD aos coeficientes HOA 11 tendo componentes complexos para gerar uma matriz V^* .

[0071] Deste modo, a unidade LIT 30 pode realizar SVD com relação aos coeficientes HOA 11 para produzir vetores $US[k]$ 33 (que podem representar uma versão combinada dos vetores S e os vetores U) tendo dimensões $D: M \times (N^2)2$, e vetores $V[k]$ 35 tendo dimensões $D: (N+1)^2 \times (N+1)^2$. Elementos de vetor individuais na matriz $US[k]$ também podem ser chamados $X_{PS}(k)$ enquanto vetores individuais da matriz $V[k]$ também podem ser chamados de $v(k)$.

[0072] Uma análise das matrizes U , 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 N vetores em U (de amostras de comprimento M) pode representar sinais de áudio separados normalizados em função do tempo (para o período de tempo representado por M amostras), que são ortogonais entre si e que foram dissociados de quaisquer características espaciais (o que também pode ser referido como informação direcional). As características espaciais, que representam o formato espacial e posição (r , θ , ϕ) podem ao invés ser representadas pelos vetores individuais $i^{\text{ésimo}}$, $v^{(i)}(k)$, na matriz V (cada do comprimento $(N+1)^2$). Os elementos individuais de cada um dos vetores $v^{(i)}(k)$ podem representar um coeficiente HOA que descreve o formato (incluindo a largura) e a posição do campo de som para um objeto de áudio associado. Ambos os vetores na matriz U e

na matriz V são normalizados de tal forma que suas energias de média quadrática são iguais à unidade. A energia dos sinais de áudio em U é assim representada pelos elementos diagonais em S. Multiplicar U e S para formar US[k] (com os elementos de vetor individual $X_{PS}(k)$), representa assim o sinal de áudio com energias. A capacidade da decomposição SVD para desacoplar os sinais de tempo de áudio (em U), as suas energias (em S) e as suas características espaciais (em V) podem suportar vários aspectos das técnicas descritas nesta revelação. Ainda, o modelo de sintetização dos coeficientes HOA subjacentes [k], X, por uma multiplicação de vetor de US[k] e V[k] dá origem ao termo "decomposição baseada em vetor", o qual é usado por dodo este documento.

[0073] Embora descrito como sendo realizado diretamente em relação aos coeficientes HOA 11, a unidade LIT 30 pode aplicar a transformada linear inversível aos derivados dos coeficientes HOA 11. Por exemplo, a unidade LIT 30 pode aplicar SVD em relação a uma matriz de densidade espectral de potência derivada dos coeficientes 11 de HOA. Ao realizar a SVD em relação à densidade espectral de potência (PSD) dos coeficientes HOA em vez dos próprios coeficientes, a unidade LIT 30 pode potencialmente reduzir a complexidade computacional da realização de SVD em termos de um ou mais ciclos de processador e espaço de armazenamento, enquanto obtendo a mesma eficiência de codificação do áudio da fonte como se a SVD fosse aplicada diretamente aos coeficientes HOA.

[0074] A unidade de cálculo de parâmetro 32 representa uma unidade configurada para calcular vários parâmetros, 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 denotado como $R[k]$, $\theta[k]$, $\varphi[k]$, $r[k]$ e $e[k]$. A unidade de cálculo de parâmetro 32 pode realizar uma análise e/ou correlação de energia (ou a chamada correlação cruzada) com relação aos vetores $US[k]$ 33 para identificar os parâmetros. A unidade de cálculo de parâmetro 32 também pode determinar os parâmetros para o quadro anterior, onde os parâmetros do quadro anterior podem ser denotados $R[k-1]$, $\theta[k-1]$, $\varphi[k-1]$, $r[k-1]$ e $e[k-1]$, com base no quadro anterior do vetor $US[k-1]$ e vetores $V[k-1]$. A unidade de cálculo de parâmetros 32 pode fornecer os parâmetros atuais 37 e os parâmetros anteriores 39 para a unidade de reordenação 34.

[0075] Os parâmetros calculados pela unidade de cálculo de parâmetros 32 podem ser utilizados pela unidade de reordenação 34 para reordenar os objetos de áudio para representarem a sua avaliação ou continuidade natural ao longo do tempo. A unidade de reordenação 34 pode comparar cada um dos parâmetros 37 dos primeiros vetores $US [k]$ 33 por virada contra cada um dos parâmetros 39 para os segundos vetores $US [k-1]$ 33. A unidade de reordenação 34 pode reordenar (usando, como um exemplo, um algoritmo Húngaro) os vários vetores dentro da matriz $US[k]$ 33 e a matriz $V[k]$ 35 com base nos parâmetros atuais 37 e os parâmetros anteriores 39 para produzir uma matriz $US[k]$ reordenada 33' (que pode ser denotada matematicamente como $\overline{US}[k]$) e uma matriz $V[k]$ reordenada 35' (que pode ser denotada matematicamente como $\overline{V}[k]$) para uma unidade de seleção de som de primeiro plano (ou som predominante - PS) 36 ("unidade de seleção de primeiro plano 36") e uma unidade de compensação de energia 38.

[0076] A unidade de análise de campo de som 44 pode representar uma unidade configurada para realizar uma

análise de campo de som com relação aos coeficientes HOA 11 de modo a potencialmente atingir uma taxa de bit alvo 41. A unidade de análise de campo de som 44 pode, com base na análise e/ou em uma taxa de bit alvo recebida 41, determinar o número total de instanciações do codificador psicoacústico (que pode ser uma função do número total de canais ambiente ou de fundo (BG_{TOT}) e o número de canais de primeiro plano ou, em outras palavras, canais predominantes. O número total de instanciações do codificador psicoacusticas pode ser denotado como numHOATransportChannels.

[0077] A unidade de análise de campo de som 44 também pode determinar, novamente atingir potencialmente a taxa de bit alvo 41, o número total de canais de primeiro plano (nFG) 45, a ordem mínima do campo de som de fundo (ou, em outras palavras, ambiente) (N_{BG} ou, alternativamente, MinAmbHOAorder), o número correspondente de canais reais representativos da ordem mínima do campo de som de fundo ($nBGa = (\text{MinAmbHOAorder} + 1)^2$), e índices (i) dos canais BG HOA adicionais para enviar (que podem coletivamente ser denotados como informação de canal de fundo 43 no exemplo da FIG. 3. A informação do canal de fundo 42 também pode ser referida como informação de canal ambiente 43. Cada um dos canais que restam a partir de numHOATransportChannels - nBGa, podem ser ou um "canal de fundo/ambiente adicional", um "canal predominante a base de vetor ativo", um "sinal predominante com base direcional ativo" ou "completamente inativo". Em um aspecto, os tipos de canal podem ser o elemento de sintaxe indicado (como um "ChannelType") por dois bits (por exemplo, 00: sinal com base direcional, 01: sinal predominante baseado em vetor, 10: sinal ambiente adicional, 11: sinal inativo). O número total de sinais de fundo ou 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 bitstream para aquele quadro.

[0078] A unidade de análise de campo de som 44 pode selecionar o número de canais de fundo (ou, em outras palavras, ambiente) e o número de canais de primeiro plano (ou, em outras palavras, predominantes) com base na taxa de bits alvo 41, selecionando mais canais de fundo e/ou de 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, o numHOATransportChannels pode ser definido para 8 enquanto o MinAmbHOAorder pode ser definido para 1 na seção de cabeçalho do bitstream. Neste cenário, em cada quadro, podem ser dedicados quatro canais para representar a porção de fundo ou ambiente do campo de som enquanto os outros 4 canais podem, em uma base quadro a quadro, variar no tipo de canal - por exemplo, utilizado como um canal de fundo/ambiente adicional ou um canal de primeiro de primeiro plano/predominante. Os sinais de primeiro plano/predominantes podem ser um dos sinais baseados em vetor ou direcionais, como descrito acima.

[0079] Em alguns casos, o número total de sinais predominantes baseados em vetor para um quadro, pode ser dado pelo número de vezes que o índice de ChannalType é 01 no bitstream desse quadro. No aspecto acima, para cada canal de fundo/ambiente adicional (por exemplo, correspondente a um ChannelType de 10), a informação correspondente de qual dos possíveis coeficientes de HOA (além dos quatro primeiros) pode ser representado nesse canal. A informação, para o conteúdo HOA de quarta ordem, pode ser um índice para indicar os coeficientes HOA 5-25. Os primeiros quatro coeficientes HOA ambiente 1-4 podem ser

enviados todo o tempo quando minAmbHOAorder é ajustado para 1, portanto, o dispositivo de codificação de áudio pode apenas precisar indicar um dos coeficientes HOA ambiente adicionais com um índice de 5-25. A informação poderia assim ser enviada usando um elemento de sintaxe de 5 bits (para o conteúdo de 4^a ordem), que pode ser denotado como "CodedAmbCoeffIdx". Em qualquer caso, a unidade de análise de campo de som 44 produz a informação de canal de fundo 43 e os coeficientes HOA 11 para a unidade de seleção de fundo (BG) 36, a informação de canal de fundo 43 para a unidade de redução do coeficiente 46 e a unidade de geração de bitstream 42, e o nFG 45 para uma unidade de seleção de primeiro plano 36.

[0080] A unidade de seleção de fundo 48 pode representar uma unidade configurada para determinar coeficientes HOA de fundo ou ambiente 47 com base na informação de canal de fundo (ex., o campo de som de fundo (N_{BG}) e o número (nBGA) e os índices (i) dos canais BG HOA adicionais para enviar). Por exemplo, quando N_{BG} é igual a um, a unidade de seleção de fundo 48 pode selecionar os coeficientes HOA 11 para cada amostrado quadro de áudio tendo uma ordem igual a ou menor que um. A unidade de seleção de fundo 48 pode neste exemplo selecionar então os coeficientes HOA 11 possuindo um índice identificado por um dos índices (i) como coeficientes BG HOA adicionais, em que o nBGA é fornecido à unidade de geração de bitstream 42 para ser especificado no bitstream 21 de modo a permitir que o dispositivo de decodificação de áudio, como o dispositivo de decodificação de áudio 24 mostrado no exemplo das FIGs. 2 e 4, analise os coeficientes HOA de fundo 47 a partir do bitstream 21. A unidade de seleção de fundo 48 pode emitir então os coeficientes de HOA ambiente 47 para a unidade de compensação de energia 38. Os

coeficientes HOA ambiente 47 podem ter dimensões D: $M \times [(N_{BG}+1)^2 + nBGa]$. Os coeficientes de HOA ambiente 47 podem também ser referidos como "coeficientes de HOA ambiente 47", onde cada um dos coeficientes de HOA ambiente 47 corresponde a um canal HOA ambiente separado 47 para ser codificado pela unidade de codificador de áudio psicoacústico 40.

[0081] A unidade de seleção de primeiro plano 36 pode representar uma unidade configurada para selecionar a matriz $US[k]$ reordenada 33' e a matriz $V[k]$ reordenada 35' que representa os componentes de primeiro plano ou distintos do campo de som baseado em nFG 45 (que pode representar um ou mais índices que identificam os vetores de primeiro plano). A unidade de seleção de primeiro plano 36 pode produzir sinais nFG 49 (que podem ser denotados como uma US reordenada $[k]_1, \dots, _{nFG} 49, FG_1, \dots, _{nFG[k]} 49$, ou $X_{PS}^{(1..nFG)}(k) 49$) para a unidade de codificador de áudio psicoacústico 40, onde os sinais nFG 49 podem ter dimensões D: $M \times nFG$ e cada representa objetos de monoaudio. A unidade de seleção de primeiro plano 36 podem também produzir a matriz reordenada $V[k] 35'$ (ou $v^{(1..nFG)}(k) 35'$) que corresponde aos componentes de primeiro plano do campo de som para a unidade de interpolação espaço-temporal 50, onde um subconjunto da matriz reordenada $V[k] 35'$ correspondente aos componentes de primeiro plano pode ser denotado como matriz de primeiro plano $V[k] 51_k$ (que pode ser matematicamente denotada como $\bar{V}_{1,\dots,nFG}[k]$) tendo dimensões D: $(N+1)^2 \times nFG$.

[0082] A unidade de compensação de energia 38 pode representar uma unidade configurada para realizar a compensação de energia com relação aos coeficientes HOA ambiente 47 para compensar a perda de energia devido à

remoção de vários dos canais HOA pela unidade de seleção de fundo 48. A unidade de compensação de energia 38 pode realizar uma análise de energia com relação a uma ou mais da matriz reordenada $US[k]$ 33', a matriz reordenada $V[k]$ 35', os sinais nFG 49, os vetores de primeiro plano $V[k]$ 51_k e os coeficientes HOA ambiente 47 e então realizar a compensação de energia com base na análise de energia para gerar coeficientes HOA ambiente 47' compensados em energia. A unidade de compensação de energia 38 pode produzir os coeficientes HOA ambiente 47' compensados em energia para a unidade de codificador de áudio psicoacústico 40.

[0083] A unidade de interpolação espaço-temporal 50 pode representar uma unidade configurada para receber os vetores $V[k]$ de primeiro plano 51_k para o k^{ésimo} quadro e os vetores $V[k-1]$ de primeiro plano 51_{k-1} para o quadro anterior (aqui a notação k-1) e realizar a interpolação espaço-temporal para gerar os vetores de primeiro plano interpolados $V[k]$. A unidade de interpolação espaço-temporal 50 pode recombinar os sinais nFG 49 com os vetores $V[k]$ de primeiro plano 51_k para recuperar os coeficientes HOA de primeiro plano reordenados. A unidade de interpolação espaço-temporal 50 pode então dividir os coeficientes HOA de primeiro plano reordenados pelos vetores $V[k]$ interpolados para gerar os sinais nFG interpolados 49'. A unidade de interpolação espaço-temporal 50 também pode produzir os vetores $V[k]$ de primeiro plano 51_k que foram usados para gerar os vetores interpolados $V[k]$ de primeiro plano de modo que um dispositivo de decodificação de áudio, como o dispositivo de decodificação de áudio 24, possa gerar os vetores interpolados $V[k]$ de primeiro plano e assim recuperar os vetores $V[k]$ de primeiro plano 51_k. Os vetores $V[k]$ de primeiro plano 51_k usados para gerar os vetores

interpolados $V[k]$ de primeiro plano são denotados como os vetores restantes $V[k]$ de primeiro plano 53. Para garantir que os mesmos $V[k]$ e $V[k-1]$ são usados no codificador e decodificador (para criar os vetores interpolados $V[k]$) versões quantizadas/dequantizadas dos vetores podem ser usadas no codificador e decodificador. A unidade de interpolação espaço-temporal 50 pode produzir os sinais nFG interpolados 49' para a unidade de codificador de áudio psicoacústico 46 e os vetores interpolados $V[k]$ de primeiro plano 51_k para a unidade de redução de coeficiente 46.

[0084] A unidade de redução de coeficiente 46 pode representar uma unidade configurada para realizar a redução de coeficiente com relação aos vetores restantes $V[k]$ de primeiro plano 53 com base na informação do canal de fundo 43 para produzir vetores reduzidos $V[k]$ de primeiro plano 55 para a unidade de quantização 52. Os vetores reduzidos $V[k]$ de primeiro plano 55 podem ter dimensões D: $[(N+1)^2 - (N_{BG}+1)^2 - BG_{TOT}] \times nFG$. A unidade de redução de coeficiente 46 pode, nesse aspecto, representar uma unidade configurada para reduzir o número de coeficientes nos vetores restantes $V[k]$ de primeiro plano 53. Em outras palavras, a unidade de redução de coeficiente 46 pode representar uma unidade configurada para eliminar os coeficientes nos vetores $V[k]$ de primeiro plano (que formam os vetores restantes $V[k]$ de primeiro plano 53) tendo pouca ou nenhuma informação direcional. Em alguns exemplos, os coeficientes do distinto ou, em outras palavras, vetores $V[k]$ de primeiro plano correspondentes a uma função básica de primeira ordem ou ordem zero (que pode ser denotada como N_{BG}) fornece pouca informação direcional e, portanto, pode ser removida dos vetores V do primeiro plano (através de um processo que pode ser referido como "redução de coeficiente"). Neste exemplo, maior

flexibilidade pode ser fornecida não somente para identificar os coeficientes que correspondem a N_{BG} mas para identificar canais HOA adicionais (que podem ser denotados pela variável TotalOfAddAmbHOAChan) a partir do conjunto de $[(N_{BG} + 1)^2 + 1, (N+1)^2]$.

[0085] A unidade de quantização 52 pode representar uma unidade configurada para realizar qualquer forma de quantização para comprimir os vetores reduzidos $V[k]$ de primeiro plano 55 para gerar vetores codificados $V[k]$ de primeiro plano 57, produzindo os vetores codificados $V[k]$ de primeiro plano 57 para a unidade de geração de bitstream 42. Em funcionamento, 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 reduzidos $V[k]$ de primeiro plano 55 neste exemplo. A unidade de quantização 52 pode realizar qualquer um dos 12 modos de quantização a seguir, como indicado por um elemento de sintaxe de modo de quantização denotado "NbtsQ":

Valor de NbtsQ	Tipo de Modo de Quantização		
0-3.	Reservado		
4:	Vetor	Quantização	
5:	Escalar	Quantização	sem
Codificação de Huffman			
6:	6-Bit	Escalar	Quantização
Codificação de Huffman			
7:	7-Bit	Escalar	Quantização
Codificação de Huffman			
8:	8-Bit	Escalar	Quantização
Codificação de Huffman			
...	...		
16:	16-Bit	Escalar	Quantização
Codificação de Huffman			

A unidade de quantização 52 também pode executar versões previstas de qualquer um dos tipos anteriores de modos de quantização, onde é determinada uma diferença entre um elemento de (ou um peso quando a quantização do vetor é realizada) do vetor V de um quadro anterior e o elemento (ou peso quando a quantização do vetor é realizada) do vetor V de um quadro atual é determinado. A unidade de quantização 52 pode então quantizar a diferença entre os elementos ou pesos do quadro atual e do quadro anterior em vez do valor do elemento do vetor V do próprio quadro atual.

[0086] A unidade de quantização 52 pode executar múltiplas formas de quantização em relação a cada um dos vetores [V] [k] de primeiro plano reduzidos para obter múltiplas versões codificadas dos vetores [V] [k] de primeiro plano reduzidos. A unidade de quantização 52 pode selecionar uma das versões codificadas dos vetores [V] [k] de primeiro plano reduzidos 55 como o vetor V[k] codificado de primeiro plano 57. Em outras palavras, a unidade de quantização 52 pode selecionar um vetor V quantizado por vetor não previsto, vetor V quantizado por vetor previsto, vetor V quantizado escalar codificado não-Huffman e o vetor V quantizado escalar codificado d Huffman para usar como o vetor V quantizado comutado de saída com base em qualquer combinação dos critérios discutidos nesta revelação. 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 de vetor e um ou mais modos de quantização escalares e quantizar um vetor de entrada V baseado em (ou de acordo com) o modo selecionado. A unidade de quantização 52 pode então fornecer o vetor V quantizado do vetor não previsto (por exemplo, em termos de valores de peso ou bits

indicativos do mesmo), vetor V quantizado por vetor previsto (por exemplo, em termos de valores de erro ou bits indicativos do mesmo), o vetor V quantizado escalar codificado não-Huffman e o vetor V quantizado escalar codificado de Huffman para a unidade de geração de bitstream 52 como os vetores de primeiro plano codificados V[k] 57. A unidade de quantização 52 também pode fornecer os elementos de sintaxe indicativos do modo de quantização (por exemplo, o elemento de sintaxe de NbitsQ) e quaisquer outros elementos de sintaxe usados para desquantizar ou de outra forma reconstruir o vetor V.

[0087] A unidade de codificador de áudio psicoacústica 40 incluída no dispositivo de codificação de áudio 20 pode representar múltiplas instâncias de um codificador de áudio psicoacústico, sendo cada uma delas utilizada para codificar um objeto de áudio diferente ou canal de HOA de cada um dos coeficientes de HOA ambiente compensados por energia 47' e os sinais de nFG interpolados 49' para gerar coeficientes de HOA ambiente codificados 59 e sinais de nFG codificados 61. A unidade de codificador de áudio psicoacústico 40 pode emitir os coeficientes de HOA ambiente codificados 59 e os sinais de nFG codificados 61 para a unidade de geração de bitstream 42.

[0088] A unidade de geração de bitstream 42 incluída no dispositivo de codificação de áudio 20 representa uma unidade que formata os dados de acordo com um formato conhecido (que pode referir-se a um formato conhecido por um dispositivo de decodificação), gerando deste modo o bitstream baseado em vetor 21. O bitstream 21 pode, em outras palavras, representar dados de áudio codificados, tendo sido codificados da maneira descrita acima. A unidade de geração de bitstream 42 pode representar um multiplexador em alguns exemplos, os quais

podem receber os vetores de primeiro plano $V[k]$ codificados 57, os coeficientes de HOA ambiente codificados 59, os sinais de nFG codificados 61 e as informações de canal de fundo 43. A unidade de geração de bitstream 42 pode então gerar um bitstream 21 baseado nos vetores de primeiro plano $V[k]$ codificados 57, os coeficientes de HOA ambiente codificados 59, os sinais de nFG codificados 61 e as informações de canal de fundo 43. Deste modo, a unidade de geração de bitstream 42 pode, neste modo, especificar os vetores 57 no bitstream 21 para obter o bitstream 21. O bitstream 21 pode incluir um bitstream primário ou principal e um ou mais bitstreams de canal lateral.

[0089] Vários aspectos das técnicas também podem permitir que a unidade de geração de bitstream 46, como descrito acima, especifique a informação de renderização de áudio 2 no bitstream 21. Embora a versão atual do próximo rascunho de trabalho de compressão de áudio 3D forneça sinalização de matrizes de mistura específicas dentro do bitstream 21, o rascunho de trabalho não fornece a especificação de renderizadores usados na renderização de coeficientes HOA no bitstream 21. Para o conteúdo de HOA, o equivalente dessa matriz de mistura é a matriz de renderização que converte a representação de HOA nos feeds dos alto-falantes desejados. Vários aspectos das técnicas descritas nesta revelação propõem harmonizar ainda mais os conjuntos de características de conteúdo de canal e HOA permitindo que a unidade de geração de bitstream 46 sinalize as matrizes de renderização de HOA dentro do bitstream (como, por exemplo, informação de renderização de áudio 2).

[0090] Uma solução de sinalização exemplar baseada no esquema de codificação das matrizes de mistura e otimizada para HOA é apresentada a seguir. Semelhante à

transmissão de matrizes de mistura, as matrizes de renderização de HOA podem ser sinalizadas dentro de *mpegh3daConfigExtension()*. As técnicas podem fornecer um novo tipo de extensão ID_CONFIG_EXT_HOA_MATRIX conforme estabelecido nas tabelas a seguir (com itálico e negrito indicando alterações na tabela existente).

Tabela – Sintaxe de mpegh3daConfigExtension()
 (Tabela 13 no CD)

Sintaxe	Nº de bits.	Mnemonic
<pre>mpegh3daConfigExtension() { numConfigExtensions = escapedValue(2,4,8) + 1; para (confExtIdx=0; confExtIdx<numConfigExtensions; confExtIdx++) { usacConfigExtType[confExtIdx] = escapedValue(4,8,16); usacConfigExtLength[confExtIdx] = escapedValue(4,8,16); switch (usacConfigExtType[confExtIdx]) { caso ID_CONFIG_EXT_FILL: enquanto (usacConfigExtLength[confExtIdx]--) { fill_byte[i]; /* deve ser '10100101' */ } quebra; caso ID_CONFIG_EXT_DMX_MATRIX: DownmixMatrixSet() quebra; <u>caso ID_CONFIG_EXT_HOA_MATRIX:</u> <u>HOARenderingMatrixSet()</u> <u>quebra;</u> caso ID_CONFIG_EXTLOUDNESS_INFO: loudnessInfoSet (); quebra; padrão: enquanto (usacConfigExtLength[confExtIdx]--) { tmp; } quebra; } } } }</pre>	8 8	uimsbf uimsbf

Tabela – Valor de usacConfigExtType (Tabela 1 no CD)

usacConfigExtType	Valor
ID_CONFIG_EXT_FILL	0
ID_CONFIG_EXT_DMX_MATRIX	1
ID_CONFIG_EXT_LOUDNESS_INFO	2
<u>ID_CONFIG_EXT_HOA_MATRIX</u>	<u>3</u>
/* reservado para uso de ISO */	<u>4-127</u>
/* reservado para uso fora do escopo de ISO */	128 e superior

[0091] O campo de bits *HOARenderingMatrixSet()* pode ser igual em estrutura e funcionalidade comparado ao *DownmixMatrixSet()*. Ao invés de *inputCount(audioChannelLayout)*, o *HOARenderingMatrixSet()* pode usar o valor “equivalente” *NumOfHoaCoeffs*, calculado em *HOAConfig*. Ainda, porque o ordenação dos coeficientes HOA pode ser fixa dentro do decodificador HOA (ver, ex., Anexo G no CD), o *HOARenderingMatrixSet* não precisa de qualquer equivalente à *inputConfig(audioChannelLayout)*.

Tabela 2 – Sintaxe de HOARenderingMatrixSet()
 (adotado a partir da Tabela 15 no CD)

Sintaxe	Nº de bits.	Mnemônico
<pre>HOARenderingMatrixSet() { numHOARenderingMatrices para (k=0; k< numHOARenderingMatrices; ++k) { downmixId; CICPspeakerLayoutIdx; DmxMatrixLenBits = escapedValue(8,8,12); HOARenderingMatrix (NumOfHoaCoeffs, outputConfig(CICPspeakerLayoutIdx), outputCount(CICPspeakerLayoutIdx)); } }</pre>	5 6 6 8..28	Uimsbf Uimsbf Uimsbf DmxMatr ixLenBi ts

[0092] Vários aspectos das técnicas também podem permitir que a unidade de geração de bitstream 46, ao comprimir os dados de áudio HOA (por exemplo, os coeficientes de HOA 11 no exemplo da FIG. 4), utilizando um primeiro esquema de compressão (como o esquema de compressão de decomposição representado pela unidade de decomposição baseada em vetor 27), especifique o bitstream 21 de modo que os bits correspondentes a um segundo esquema de compressão (por exemplo, o sistema de compressão baseado e direção ou esquema de compressão baseado em direcionalidade representado pela unidade de decomposição baseada na direcção 28) não estão incluídos no bitstream 21. Por exemplo, a unidade de geração de bitstream 42 pode gerar o bitstream 21 de modo a não incluir elementos de sintaxe HOAPredictionInfo ou campo que pode ser reservado para utilização para especificar a informação de predição entre sinais direcionais do esquema de compressão com base na direção. Exemplos do bitstream 21 gerado de acordo com

vários aspectos das técnicas descritas nesta descrição são mostrados nos exemplos das FIGs. 8E e 8F.

[0093] Em outras palavras, a previsão de sinais direcionais pode ser parte da Síntese de Som Predominante empregada pela unidade de decomposição com base direcional 28 e depende da existência de ChannelType 0 (o que pode indicar um sinal baseado em direção). Quando nenhum sinal com base na direção está presente dentro de um quadro, nenhuma previsão de sinais direcionais pode ser efetuada. No entanto, a informação de banda lateral associada *HOAPredictionInfo()* pode, embora não utilizada, ser escrita em cada quadro independentemente da existência de sinais com base na direção. Quando não existe sinal direcional dentro de um quadro, as técnicas descritas nesta descrição podem permitir que a unidade de geração de bitstream 42 reduza o tamanho da banda lateral não sinalizando HOAPredictionInfo na banda lateral conforme estabelecido na seguinte Tabela (em que os itálicos com sublinhado denotam adições):

Tabela: Sintaxe de HOAFrame

Sintaxe	Nº de bits.	Mnemoni c
<pre> HOAFrame(usacIndependencyFlag) { NumOfDirSigs = 0; NumOfVecSigs = 0; NumOfContAddHoaChans = 0; Se (usacIndependencyFlag){ hoaIndependencyFlag = usacIndependencyFlag; } ainda{ hoaIndependencyFlag; } para(i=0; i< NumOfAdditionalCoders; ++i){ ChannelSideInfoData(i); se (<i>MaxGainCorrAmpExp>0</i>) { HOAGainCorrectionData(i); } switch ChannelType[i] { caso 0: DirSigChannelIds[NumOfDirSigs] = i + 1; NumOfDirSigs++; quebra; caso 1: VecSigChannelIds[NumOfVecSigs] = i + 1; NumOfVecSigs++; quebra; caso 2: se (AmbCoeffTransitionState[i] == 0) { ContAddHoaCoeff [NumOfContAddHoaChans] = AmbCoeffIdx[i]; NumOfContAddHoaChans++; } quebra; } } se (<i>MaxGainCorrAmpExp>0</i>) { </pre>	1	bslbf

```

        para ( i=
NumOfAdditionalCoders;
        i<
NumHOATransportChannels; ++i){
    HOAGainCorrectionData(i);
}
}
Para (i=0; i< NumOfVecSigs;
++i){
    VVectorData (
VecSigChannelIds(i) );
}
se (NumOfDirSigs>0)
{
    HOAPredictionInfo(
DirSigChannelIds, NumOfDirSigs );
    byte_alignment();
}

```

[0094] Nesse aspecto, as técnicas podem permitir que um dispositivo, como o dispositivo de codificação de áudio 20 seja configurado para, ao comprimir dados de áudio ambissônicos de ordem superior utilizando um primeiro esquema de compressão, especificar um bitstream representativo de uma versão comprimida dos dados de áudio de ordem ambissônica superior que não incluem bits correspondentes a um segundo esquema de compressão também utilizado para comprimir os dados de áudio ambissônicos de ordem superior.

[0095] Em alguns casos, o primeiro esquema de compressão compreende um esquema de compressão de decomposição baseado em vetor. Nestes e em outros casos, o esquema de compressão de decomposição baseado em vetor compreende um esquema de compressão que envolve a aplicação de uma decomposição de valor singular (ou equivalentes deste descrito com mais detalhe nesta descrição) aos dados de áudio ambissônicos de ordem superior.

[0096] Nestes e em outros casos, o dispositivo de codificação de áudio 20 pode ser configurado para especificar o bitstream que não inclui os bits que

correspondem a pelo menos um elemento de sintaxe utilizado para executar o segundo tipo de esquema de compressão. O segundo esquema de compressão pode, como referido acima, compreender um esquema de compressão baseado em direcionalidade.

[0097] O dispositivo de codificação de áudio 20 também pode ser configurado para especificar o bitstream 21 de modo que o bitstream 21 não inclua os bits correspondentes a um elemento de sintaxe HOAPredictionInfo do segundo esquema de compressão.

[0098] Quando o segundo esquema de compressão compreende um esquema de compressão baseado em direcionalidade, o dispositivo de codificação de áudio 20 pode ser configurado para especificar o bitstream 21 de modo que o bitstream 21 não inclua os bits correspondentes a um elemento de sintaxe HOAPredictionInfo do esquema de compressão baseado na direcionalidade. Em outras palavras, o dispositivo de codificação de áudio 20 pode ser configurado para especificar o bitstream 21 de modo que o bitstream 21 não inclua os bits que correspondem a pelo menos um elemento de sintaxe utilizado para executar o segundo tipo de esquemas de compressão, pelo menos um elemento de sintaxe indicativo de uma previsão entre dois ou mais sinais baseados em direção. Reiniciado novamente, quando o segundo esquema de compressão compreende um esquema de compressão baseado em direcionalidade, o dispositivo de codificação de áudio 20 pode ser configurado para especificar o bitstream 21 de modo que o bitstream 21 não inclua os bits correspondentes a um elemento de sintaxe HOAPredictionInfo do esquema de compressão baseado na direcionalidade, onde o elemento de sintaxe HOAPredictionInfo é indicativo de uma previsão entre dois ou mais sinais baseados em direção.

[0099] Vários aspectos das técnicas podem, ainda, permitir que a unidade de geração de bitstream 46 especifique o bitstream 21 em determinados casos de modo que o bitstream 21 não inclui dados de correção de ganho. A unidade de geração de bitstream 46 pode, quando a correção de ganho é suprimida, especificar o bitstream 21 de modo que o bitstream 21 não inclui os dados de correção de ganho. Exemplos do bitstream 21 gerado de acordo com vários aspectos das técnicas como mostrado, conforme observado acima, nos exemplos das FIGs. 8E e 8F.

[0100] Em alguns casos, a correção de ganho é aplicada quando certos tipos de codificação psicoacústica são executados dado o intervalo dinâmico relativamente menor destes tipos específicos de codificação psicoacústica em comparação com outros tipos de codificação psicoacústica. Por exemplo, AAC tem um intervalo dinâmico relativamente menor do que a codificação unificada de voz e áudio (USAC). Quando o esquema de compressão (como um esquema de compressão de síntese baseado em vetor ou um esquema de compressão baseado em direção) envolve USAC, a unidade de geração de bitstream 46 pode sinalizar no bitstream 21 que a correção de ganho foi suprimida (por exemplo, especificando um elemento d sintaxe *MaxGainCorrAmpExp* no *HOAConfig* com um valor de zero no bitstream 21) e, em seguida, especificar o bitstream 21 para não incluir os dados de correção de ganho (em um campo *HOAGainCorrectionData*()).

[0101] Em outras palavras, o campo de bit *MaxGainCorrAmpExp* como parte do *HOAConfig* (consulte a Tabela 71 no CD) pode controlar a extensão em que o módulo de controle de ganho automático afeta os sinais do canal de transporte antes da codificação do núcleo USAC. Em alguns casos, este módulo foi desenvolvido para RM0 para melhorar

a faixa dinâmica não ideal da implementação do codificador AAC disponível. Com a mudança de AAC para o codificador de núcleo USAC durante a fase de integração, o intervalo dinâmico do codificador de núcleo pode melhorar e, portanto, a necessidade deste módulo de controle de ganho não pode ser tão crítica como antes.

[0102] Em alguns casos, a funcionalidade de controle de ganho pode ser suprimida se `MaxGainCorrAmpExp` for definido como 0. Nesses casos, a informação de banda lateral associada `HOAGainCorrectionData()` pode não ser gravada em cada quadro de HOA pela tabela acima ilustrando a "Sintaxe de HOAFrame". Para a configuração onde `MaxGainCorrAmpExp` é definido como 0, as técnicas descritas nesta revelação podem não sinalizar o `HOAGainCorrectionData`. Além disso, em tal cenário, o módulo de controle de ganho inverso pode ainda ser anulado, reduzindo a complexidade do decodificador em cerca de 0,05 MOPS por canal de transporte sem qualquer efeito secundário negativo.

[0103] Nesse aspecto, as técnicas podem configurar o dispositivo de codificação de áudio 20 para, quando a correção de ganho é suprimida durante a compressão de dados de áudio ambissônicos de ordem superior, especificar o bitstream 21 representativo de uma versão comprimida dos dados de áudio ambissônicos de ordem superior de modo que o bitstream 21 não inclui dados de correção de ganho.

[0104] Nestes e em outros casos, o dispositivo de codificação de áudio 20 pode ser configurado para comprimir os dados de áudio ambissônicos de ordem superior de acordo com um esquema de compressão de decomposição baseado em vetor para gerar a versão comprimida dos dados de áudio ambissônicos de ordem superior. Exemplos do

esquema de compressão de decomposição podem envolver a aplicação de uma decomposição de valor singular (ou seus equivalentes descritos em maior detalhe acima) aos dados de áudio ambissônicos de ordem superior para gerar a versão comprimida dos dados de áudio ambissônicos de ordem superior.

[0105] Nestes e em outros casos, o dispositivo de codificação de áudio 20 pode ser configurado para especificar um elemento de sintaxe MaxGainCorrAmbExp no bitstream 21 como zero para indicar que a correção de ganho é suprimida. Em alguns casos, o dispositivo de codificação de áudio 20 pode ser configurado para especificar, quando a correção de ganho é suprimida, o bitstream 21 de modo que o bitstream 21 não inclua um campo de dados de HOAGainCorrection que armazene os dados de correção de ganho. Em outras palavras, o dispositivo de codificação de áudio 20 pode ser configurado para especificar um elemento de sintaxe MaxGainCorrAmbExp no bitstream 21 como zero para indicar que a correção de ganho é suprimida e não incluir no bitstream um campo de dados de HOAGainCorrection que armazena os dados de correção de ganho.

[0106] Nestes e em outros casos, o dispositivo de codificação de áudio 20 pode ser configurado para suprimir a correção de ganho quando a compressão dos dados de áudio ambissônicos de ordem superior inclui a aplicação de uma codificação de uma fala de áudio unificada e codificação de áudio de fala (USAC) aos dados de áudio ambissônicos de ordem superior.

[0107] As otimizações potenciais anteriores para a sinalização de várias informações no bitstream 21 podem ser adaptadas ou de outro modo atualizadas da maneira descrita em maior detalhe abaixo. As atualizações podem ser aplicadas em conjunto com outras atualizações

discutidas abaixo ou usadas para atualizar apenas vários aspectos das otimizações discutidas acima. Como tal, cada combinação potencial de atualizações para as otimizações descritas acima são consideradas, incluindo a aplicação de uma única atualização descrita abaixo para as otimizações descritas acima ou quaisquer combinações particulares das atualizações descritas abaixo para as otimizações descritas acima.

[0108] Para especificar uma matriz no bitstream, a unidade de geração de bitstream 42 pode, por exemplo, especificar um ID_CONFIG_EXT_HOA_MATRIX em um mpegh3daConfigExtension() do bitstream 21, como ilustrado em negrito e realçado na Tabela seguinte. A Tabela a seguir é representativa da sintaxe para especificar a porção mpegh3daConfigExtension() do bitstream 21:

Tabela - Sintaxe de mpegh3daConfigExtension()

Sintaxe	Nº de bits.	Mnemonico
<pre>mpegh3daConfigExtension() { numConfigExtensions = escapedValue(2,4,8) + 1; para (confExtIdx=0; confExtIdx<numConfigExtensions; confExtIdx++) { usacConfigExtType[confExtIdx] = escapedValue(4,8,16); usacConfigExtLength[confExtIdx] = escapedValue(4,8,16); switch (usacConfigExtType[confExtIdx]) { caso ID_CONFIG_EXT_FILL: enquanto (usacConfigExtLength[confExtIdx]--) { fill_byte[i]; /* deve ser 8 uimsbf '10100101' */ } quebra; caso ID_CONFIG_EXT_DOWNMIX: downmixConfig(); quebra; caso ID_CONFIG_EXTLOUDNESS_INFO: loudnessInfoSet(); quebra; caso ID_CONFIG_EXT_AUDIOSCENE_INFO: mae_AudioSceneInfo(); quebra; caso ID_CONFIG_EXT_HOA_MATRIX : HoaRenderingMatrixSet(); quebra; padrão: enquanto (usacConfigExtLength[confExtIdx]--) { tmp; 8 uimsbf } quebra; } } } }</pre>		

O ID_CONFIG_EXT_HOA_MATRIX na tabela anterior fornece um recipiente no qual especificar a matriz de

renderização, o recipiente denotado como "HoaRenderingMatrixSet()".

[0109] O conteúdo do recipiente HoaRenderingMatrixSet() pode ser definido de acordo com a sintaxe estabelecida na Tabela a seguir:

Tabela - Sintaxe de HoaRenderingMatrixSet()

Sintaxe	Nº de bits.	Mnemoni c
<pre>HoaRenderingMatrixSet() { numHoaRenderingMatrices; para (k=0; k< numHoaRenderingMatrices; ++k) { HoaRenderingMatrixId; CICPspeakerLayoutIdx; HoaMatrixLenBits = escapedValue(8,8,12); HoaRenderingMatrix(NumOfHoaCoeffs outputConfig(CICPspeakerLayoutIdx), outputCount(CICPspeakerLayoutIdx)); } }</pre>	5 7 6 8..28	uimsbf uimsbf uimsbf HoaMatrixLenBits

Como mostrado na Tabela acima, o HoaRenderingMatrixSet() inclui vários elementos de sintaxe diferentes, incluindo um numHoaRenderingMatrices, um HoaRendereringMatrixId, um CICPspeakerLayoutIdx, um HoaMatrixLenBits e um HoARenderingMatrix.

[0110] O elemento de sintaxe numHoaRenderingMatrices pode especificar um número de definições HoaRendereringMatrixId presentes no elemento de bitstream. O elemento de sintaxe HoaRenderingMatrixId pode representar um campo que define unicamente um ID para uma matriz de renderização HOA padrão disponível no lado do decodificador ou uma matriz de renderização HOA transmitida. Nesse aspecto, o HoaRenderingMatrixId pode representar um exemplo do valor de sinal que inclui dois ou mais bits que definem um índice que indica que o bitstream

inclui uma matriz usada para renderizar coeficientes harmônicos esféricos para uma pluralidade de feeds de alto-falante ou o valor de sinal que inclui dois ou mais bits que definem um índice associado a uma de uma pluralidade de matrizes utilizadas para renderizar coeficientes harmônicos esféricos para uma pluralidade de feeds de alto-falante. O elemento de sintaxe CICPspeakerLayoutIdx pode representar um valor que descreve o layout do alto-falante de saída para a dada matriz de renderização HOA e pode corresponder a um elemento ChannelConfiguration definido na ISO/IEC 23000 1-8. O elemento de sintaxe HoaMatrixLenBits (que também pode ser denominado como "HoaRenderingMatrixLenBits") pode especificar um comprimento do seguinte elemento de bitstream (por exemplo, o recipiente HoaRenderingMatrix()) em bits.

[0111] O recipiente HoaRenderingMatrix() inclui um NumOfHoaCoeffs seguido de um recipiente outputConfig() e um recipiente outputCount(). O recipiente outputConfig() pode incluir vetores de configuração de canal especificando as informações sobre cada alto-falante. A unidade de geração de bitstream 42 pode assumir que esta informação de alto-falante é conhecida a partir das configurações de canal do layout de saída. Cada entrada, outputConfig[i], pode representar uma estrutura de dados com os seguintes membros:

AzimuthAngle (que pode indicar o valor absoluto do ângulo de azimute do alto-falante);

AzimuthDirection (que pode indicar a direção azimutal usando, como um exemplo, 0 para esquerda e 1 para direita);

ElevationAngle (que pode indicar o valor absoluto dos ângulos de elevação do alto-falante);

`ElevationDirection` (que pode indicar a direção de elevação usando, por exemplo, 0 para cima e 1 para baixo); e

`isLFE` (que pode indicar se o alto-falante é um alto-falante de efeito de baixa frequência (LFE)).

A unidade de geração de bitstream 42 pode invocar uma função auxiliar, em alguns casos, designada por “`findSymmetricSpeakers`”, que pode especificar ainda o seguinte:

`pairType` (que pode armazenar um valor de `SYMMETRIC` (que significa um par simétrico de dois alto-falantes em algum exemplo), `CENTER`, ou `ASYMMETRIC`); e

`symmetricPair-> originalPosition` (que pode indicar a posição na configuração do canal original do segundo (por exemplo, direito) alto-falante no grupo, apenas para grupos `SYMMETRIC`).

O recipiente `outputCount()` pode especificar um número de alto-falantes para os quais a matriz de renderização HOA é definida.

[0112] A unidade de geração de bitstream 42 pode especificar o recipiente `HoaRenderingMatrix()` de acordo com a sintaxe estabelecida na Tabela seguinte:

Tabela - Sintaxe de `HoaRenderingMatrix()`

Sintaxe	Nº de bits.	Mnemoni c
HoaRenderingMatrix ()		
{		
lfeExist = 0;		
hasLfeRendering = 0;		
para (i=0; i< inputCount; ++i)		
isHoaCoefSparse[i] = 0;		
precisionLevel	1	uimsbf
Se (gainLimitPerHoaOrder) {	1	uimsbf
para (i = 0; i<(maxHoaOrder+1); ++i		
) {		
maxGain[i] = - escapedValue(3,		
5, 5);		
minGain[i] = -(escapedValue(4,		
5, 6) + 1 - maxGain[i]);		
}		
} ainda {		
maxGain[0] = - escapedValue (3, 5,		
5);		
minGain[0] = -(escapedValue (4, 5,		
6) + 1- maxGain[0]);		
for (i = 1; i<(maxHoaOrder+1); ++i) {		
maxGain[i] = maxGain[0];		
minGain[i] = minGain[0];		
}		
}		
Se (isFullMatrix) {	1	uimsbf
firstSparseOrder	1	uimsbf
para (i =		
(firstSparseOrder*firstSparseOrder);		
i<inputCount; ++i)		
isHoaCoefSparse[i] = 1;		
}		
para (i=0; i< outputCount; ++i){		
se(outputConfig[i].isLFE)		
lfeExist = 1;		
}		
se(lfeExist)		
hasLfeRendering;	1	uimsbf
numPairs =		
findSymmetricSpeakers(outputCount,		
outputConfig,		
hasLfeRendering);		
para (i=0; i<numPairs; ++i) {		
valueSymmetricPairs[i] = 0;		
signSymmetricPairs[i] = 0;		
}		
zerothOrderAlwaysPositive;	1	uimsbf

```

    se (isAllValueSymmetric) {           1      uimsbf
        para (i=0; i<numPairs; ++i) {
            valueSymmetricPairs[i] = 1;
        } ainda {
            se (isAnyValueSymmetric) {           1      uimsbf
                para (i=0; i<numPairs; ++i)
                    valueSymmetricPairs[i] =
                boolVal;
            Se { isAnySignSymmetric) {           1      uimsbf
                para (i=0; i<numPairs; ++i) {
                    se
                    (0==valueSymmetricPairs[i])
                        signSymmetricPairs[i] = 1
                    boolVal;
                }
            }
        } ainda {
            se (isAllSignSymmetric) {           1      uimsbf
                para (i=0; i<numPairs; ++i)
                    signSymmetricPairs[i] = 1;
            } ainda { /* isAnyValueSymmetric==0
        */
            Se { isAnySignSymmetric) {           1      uimsbf
                para (i=0; i<numPairs; ++i)
                    signSymmetricPairs[i] =
                boolVal;
            }
        }
        hasVerticalCoef;                 1      uimsbf
        DecodeHoaMatrixData()
    }
}

```

Como mostrado na Tabela diretamente acima, o elemento de sintaxe numPairs é definido para o valor de saída a partir da invocação da função auxiliar findSymmetricSpeakers usando o outputCount e outputConfig e hasLfeRendering como entradas. O numPairs pode, portanto, indicar o número de pares de alto-falantes simétricos identificados na configuração de alto-falantes de saída que podem ser considerados para a codificação de simetria eficiente. O elemento de sintaxe precisionLevel na Tabela acima pode indicar uma precisão usada para quantização uniforme dos ganhos de acordo com a seguinte Tabela:

Tabela - Tamanho da etapa de quantização uniforme de hoaGain em função de precisionLevel

precisionLevel	menor tamanho da etapa de quantização [dB]
0	1,0
1	0,5
2	0,25
3	0,125

[0113] O elemento de sintaxe de gainLimitPerHoaOrder mostrado na tabela acima que estabelece a sintaxe de HoaRenderingMatrix() pode representar um sinalizador indicando se o maxGain e minGain são especificados individualmente para cada ordem ou para toda a matriz de renderização de HOA. Os elementos de sintaxe maxGain[i] podem especificar um ganho real máximo na matriz para coeficientes para a ordem HOA *i* expressa, como um exemplo, em decibéis (dB). Os elementos de sintaxe mixGain[i] podem especificar um ganho real mínimo na matriz para coeficientes da ordem HOA *i* expressa, como um exemplo, em dB. O elemento de sintaxe isFullMatrix pode representar um sinalizador que indica se a matriz de renderização HOA é esparsa ou cheia. O elemento de sintaxe firstSparseOrder pode especificar, no caso de a matriz de renderização HOA ter sido especificada como esparsa pelo elemento de sintaxe isFullMatrix, a primeira ordem de HOA que está codificada de forma esparsa. O elemento de sintaxe isHoaCoefSparse pode representar um vetor de máscara de bits derivado do elemento de sintaxe firstSparseOrder. O elemento de sintaxe lfeExists pode representar um sinalizador indicativo de se existe um ou mais LFEs em outputConfig. O elemento de sintaxe hasLfeRendering indica se a matriz de renderização contém elementos diferentes de zero para um ou mais canais LFE. O elemento de sintaxe

`zerothOrderAlwaysPositive` pode representar um sinalizador indicativo de se a 0^a ordem HOA tem apenas valores positivos.

[0114] O elemento de sintaxe `isAllValueSymmetric` pode representar um sinalizador indicativo de se todos os pares de alto-falantes simétricos têm valores absolutos iguais na matriz de renderização HOA. O elemento de sintaxe `isAnyValueSymmetric` representa um sinalizador que indica, quando falso, por exemplo, se alguns dos pares de alto-falantes simétricos têm valores absolutos iguais na matriz de renderização HOA. O elemento de sintaxe `valueSymmetricPairs` pode representar uma máscara de bit de comprimento `numPairs` indicando os pares de alto-falantes com simetria de valor. O elemento de sintaxe `isValueSymmetric` pode representar uma máscara de bits derivada da maneira mostrada na Tabela 3 a partir do elemento de sintaxe `valueSymmetricPairs`. O elemento de sintaxe `isAllSignSymmetric` pode indicar, quando não há simetrias de valor na matriz, se todos os pares de alto-falantes simétricos têm pelo menos simetrias de sinal de número. O elemento de sintaxe `isAnySignSymmetric` pode representar um sinalizador indicativo de se existem pelo menos alguns pares de alto-falantes simétricos com simetrias de sinal numérico. O elemento de sintaxe `signSymmetricPairs` pode representar uma máscara de bit de comprimento `numPairs` indicando os pares de alto-falantes com simetria de sinal. A variável `isSignSymmetric` pode representar uma máscara de bits derivada do elemento de sintaxe `signSymmetricPairs` da maneira mostrada acima na Tabela estabelecendo a sintaxe de `HoaRenderingMatrix()`. O elemento de sintaxe `hasVerticalCoef` pode representar um sinalizador indicativo de se a matriz é uma matriz de renderização HOA somente horizontal. O elemento de sintaxe

bootVal pode representar uma variável usada no circuito de decodificação.

[0115] Em outras palavras, a unidade de geração de bitstream 42 pode analisar o renderizador de áudio 1 para gerar uma ou mais das informações de simetria de valor acima (por exemplo, qualquer combinação de um ou mais elemento de sintaxe isAllValueSymmetric, elemento de sintaxe isAnyValueSymmetric, elemento de sintaxe valueSymmetricPairs, elemento de sintaxe isValueSymmetric e elemento de sintaxe valueSymmetricPairs) ou obter a informação de simetria de valor. A unidade de geração de bitstream 42 pode especificar a informação de renderização de áudio 2 no bitstream 21 da maneira mostrada acima de modo que a informação de renderização de áudio 2 inclui a informação de simetria de sinal de valor.

[0116] Além disso, em outras palavras, a unidade de geração de bitstream 42 também pode analisar o renderizador de áudio 1 para gerar uma ou mais das informações de simetria de sinal acima (por exemplo, qualquer combinação de um ou mais elemento de sintaxe isAllSignSymmetric, elemento de sintaxe isAnySignSymmetric, elemento de sintaxe signSymmetricPairs, elemento de sintaxe isSignSymmetric e elemento de sintaxe signSymmetricPairs) ou obter a informação de simetria de sinal. A unidade de geração de bitstream 42 pode especificar a informação de renderização de áudio 2 no bitstream 21 da maneira mostrada acima de modo que a informação de renderização de áudio 2 inclui a informação de simetria de sinal de áudio.

[0117] Quando se determina a informação de simetria de valor e a informação de simetria de sinal, a unidade de geração de bitstream 42 pode analisar os vários valores do processador de áudio 1, que podem ser especificados como uma matriz. Uma matriz de renderização

pode ser formulada como uma pseudo-inversa de uma matriz R. Em outras palavras, para renderizar $(N+1)^2$ canais HOA (indicados como Z abaixo) para L sinais de alto-falante (indicados pelo vetor coluna, p, dos sinais do alto-falante L), pode ser dada a seguinte equação:

$$Z = R * p.$$

Para chegar à matriz de renderização que produz os sinais de alto-falante L, o inverso da matriz R é multiplicado pelos Z canais HOA como mostrado na seguinte equação:

$$p = R^{-1} * Z.$$

A menos que o número de canais de alto-falantes, L, seja o mesmo que o número de canais HOA Z, $(N+1)^2$, a matriz R não será quadrada e um inverso perfeito não pode ser determinado. Como resultado, o pseudo-inverso pode ser usado em vez disso, que é definido da seguinte forma:

$$\text{pinv}(R) = R^T (R * R^T)^{-1},$$

onde R^T denota a transposição da matriz R. Substituindo R^{-1} na equação acima, a resolução para os sinais de alto-falante L denotados pelo vetor de coluna p pode ser denotada matematicamente como se segue:

$$p = \text{pinv}(R) * Z = R^T (R * R^T)^{-1} * Z.$$

[0118] As entradas da matriz R são os valores dos harmônicos esféricos para as posições de alto-falante com $(N+1)^2$ fileiras para os diferentes harmônicos esféricos e L colunas para os alto-falantes. A unidade de geração de bitstream 42 pode determinar pares de alto-falantes com base nos valores para os alto-falantes. Analisando os valores dos harmônicos esféricos para as posições de alto-falante, a unidade de geração de bitstream 42 pode determinar com base nos valores cujas posições de alto-falantes são pares (por exemplo, como pares podem ter

valores semelhantes, quase iguais ou iguais, mas com sinais opostos).

[0119] Depois de identificar os pares, a unidade de geração de bitstream 42 pode determinar para cada par, se os pares têm o mesmo valor ou quase o mesmo valor. Quando todos os pares têm o mesmo valor, a unidade de geração de bitstream 42 pode definir o elemento de sintaxe `isAllValueSymmetric` para um. Quando todos os pares não têm o mesmo valor, a unidade de geração de bitstream 42 pode definir o elemento de sintaxe `isAllValueSymmetric` para zero. Quando um ou mais mas não todos os pares têm o mesmo valor, a unidade de geração de bitstream 42 pode definir o elemento de sintaxe `isAnyValueSymmetric` para um. Quando nenhum os pares têm o mesmo valor, a unidade de geração de bitstream 42 pode definir o elemento de sintaxe `isAnyValueSymmetric` para zero. Para pares com valores simétricos, a unidade de geração de bitstream 42 pode especificar apenas um valor em vez de dois valores separados para o par de alto-falantes, reduzindo assim o número de bits utilizados para representar a informação de renderização de áudio 2 (por exemplo, a matriz neste exemplo) no bitstream 21.

[0120] Quando não existem simetrias de valor entre os pares, a unidade de geração de bitstream 42 também pode determinar para cada par, se os pares de alto-falantes têm simetria de sinal (significando que um alto-falante tem um valor negativo enquanto o outro alto-falante tem um valor positivo). Quando todos os pares têm simetria de sinal, a unidade de geração de bitstream 42 pode definir o elemento de sintaxe `isAllSignSymmetric` para um. Quando todos os pares não têm simetria de sinal, a unidade de geração de bitstream 42 pode definir o elemento de sintaxe `isAllSignSymmetric` para zero. Quando um ou mais mas não

todos os pares têm simetria de sinal, a unidade de geração de bitstream 42 pode definir o elemento de sintaxe `isAnySignSymmetric` para um. Quando nenhum dos pares têm simetria de sinal, a unidade de geração de bitstream 42 pode definir o elemento de sintaxe `isAnySignSymmetric` para zero. Para pares com sinais simétricos, a unidade de geração de bitstream 42 pode especificar apenas um ou nenhum sinal em vez de dois sinais separados para o par de alto-falantes, reduzindo assim o número de bits utilizados para representar a informação de renderização de áudio 2 (por exemplo, a matriz neste exemplo) no bitstream 21.

[0121] A unidade de geração de bitstream 42 pode especificar o recipiente `DecodeHoaMatrixData()` mostrado na Tabela que estabelece a sintaxe de `HoaRenderingMatrix()` de acordo com a sintaxe mostrada na Tabela seguinte:

Tabela - Sintaxe de `DecodeHoaMatrixData`

Sintaxe	Nº de bits.	Mnemonic
<pre> DecodeHoaMatrixData () { j = 0; para (i=0; i<outputCount; ++i) { isValueSymmetric[i] = 0; isSignSymmetric[i] = 0; se ((outputConfig[i].pairType == SP_PAIR_SYMMETRIC) && (outputConfig[i].symmetricPair != NULL)) { se (0==(outputConfig[i].isLFE && (0==hasLfeRendering))) { isValueSymmetric[i] = valueSymmetricPairs[j]; isSignSymmetric[i] = signSymmetricPairs[j++]; } } para (i = 0; i < inputCount; ++i) { currentHoaOrder = ceil(sqrt(i+1)- 1); para (j = outputCount-1; j >= 0; --j) { signMatrix[i * outputCount + j] = 1; hoaMatrix [i * outputCount + j] = 0.0; se ((vertBitmask[i] && hasVerticalCoef) !vertBitmask[i])) { hasValue = 1; se (0 == isValueSymmetric[j]) { se ((hasLfeRendering && outputConfig[j].isLFE) (!outputConfig[j].isLFE)) { se (isHoaCoefSparse[i]){ hasValue 1 } se (hasValue) { hoaMatrix [i * outputCount + j] = } } } } } } DecodeHoaGainValue(currentHoaOrder); </pre>	1	Uimsbf

```

        se
(0==isSignSymmetric[j]) {
    se (hoaMatrix [i * outputCount + j]
!= 0.0) {
        se (currentHoaOrder ||
!zerothOrderAlwaysPositive) {
            signMatrix[i * outputCount +      1
j] = boolVal*2-1;
        }
    }
} ainda { // isSignSymmetric[i] == 1
    pairIdx =
outputConfig[j].symmetricPair-
>originalPosition;
    signMatrix[i * outputCount + j]
= symSigns[i] * signMatrix[i *
outputCount + pairIdx];
}
}
}
}
ainda { // isAllValueSymmetric
    pairIdx =
outputConfig[j].symmetricPair-
>originalPosition;
    hoaMatrix [i*outputCount+j]
=
    hoaMatrix
[i*outputCount+pairIdx];
    SignMatrix[i
*outputCount+j] = symSigns[i] *

signMatrix[i*outputCount+pairIdx];
}
}
}
}
para (i = 0; i < inputCount; ++i) {
    para (j = 0; j < outputCount; ++j)
{
    hoaMatrix[i *outputCount+j] *=
signMatrix[i*outputCount+j];
}
}
}
}

```

[0122] O elemento de sintaxe hasValue na Tabela anterior que estabelece a sintaxe de

DecodeHoaMatrixData pode representar um sinalizador indicativo de se o elemento de matriz é codificado de modo esparso. O elemento de sintaxe signMatrix pode representar uma matriz com os valores de sinal da matriz de renderização HOA em, como um exemplo, forma vetorial linearizada. O elemento de sintaxe hoaMatrix pode representar os valores da matriz de renderização HOA em, como um exemplo, forma vetorial linearizada. A unidade de geração de bitstream 42 pode especificar o recipiente DecodeHoaGainValeu() mostrado na Tabela que estabelece a sintaxe de DecodeHoaMatrixData de acordo com a sintaxe mostrada na Tabela seguinte:

Tabela - Sintaxe de DecodeHoaGainValue

Sintaxe	Nº de bits.	Mnemoni c
<pre>DecodeHoaGainValue(order) { nAlphabet = (maxGain[order] - minGain[order]) * 2 ^ precisionLevel + 2; gainValueIndex = ReadRange(nAlphabet); gainValue = maxGain[order] - gainValueIndex / 2 ^ precisionLevel; se (gainValue < minGain) { gainValue = 0.0; } ainda { gainValue = 10.0 ^ (gainValue / 20.0); } return gainValue; }</pre>		

[0123] A unidade de geração de bitstream 42 pode especificar o recipiente readRange() mostrado na Tabela que estabelece a sintaxe de DecodeHoaGainValeu de acordo com a sintaxe especificada na Tabela seguinte:

Tabela 7 - Sintaxe de ReadRange

Sintaxe	Nº de bits.	Mnemomic
<pre>ReadRange(alphabetSize) { nBits = floor(log2(alphabetSize)); nUnused = 2 ^ (nBits + 1) - alphabetSize; faixa; nBits uimsbf se (faixa >= nUnused) { rangeExtra; 1 uimsbf faixa = faixa * 2 - nUnused + rangeExtra; } faixa de retorno; }</pre>		

[0124] Embora não ilustrado no exemplo da FIG. 3, o dispositivo de codificação de áudio 20 também pode incluir uma unidade de saída de bitstream que comuta a saída de bitstream do dispositivo de codificação de áudio 20 (por exemplo, entre o bitstream baseado na direção 21 e o bitstream baseado no vetor 21) com base em se um quadro atual deve ser codificado utilizando a síntese de base direcional ou a síntese baseada em vetor. A unidade de saída de bitstream pode executar a comutação com base no elemento de sintaxe emitido pela unidade de análise de conteúdo 26 indicando se foi efetuada uma síntese baseada em direção (como resultado da detecção de que os coeficientes HOA foram gerados a partir de um objeto de áudio sintético) ou uma síntese baseada em vetor foi realizada (como resultado da detecção de que os coeficientes de HOA foram registrados). A unidade de saída de bitstream pode especificar a sintaxe de cabeçalho correta para indicar a codificação de comutação ou corrente usada para a quadro atual juntamente com um respectivo dos bitstreams 21.

[0125] Além disso, como observado acima, a unidade de análise de campo de som 44 pode identificar os coeficientes HOA ambiente BG_{TOT} 47, os quais podem mudar em uma base quadro a quadro (embora às vezes a BG_{TOT} possa permanecer constante ou igual através de dois ou mais quadros adjacentes (no tempo)). A alteração em BG_{TOT} pode resultar em alterações nos coeficientes expressos nos vetores $V[k]$ de primeiro plano reduzidos. A alteração em BG_{TOT} pode resultar em coeficientes HOA de fundo (que também podem ser referidos como "coeficientes HOA ambiente") que mudam em uma base quadro a quadro (embora, novamente, às vezes BG_{TOT} possa permanecer constante ou igual em dois ou mais quadros adjacentes (no tempo)). As alterações resultam frequentemente em uma alteração de energia para os aspectos do campo de som representados pela adição ou remoção dos coeficientes HOA ambiente adicionais e a correspondente remoção de coeficientes de ou adição de coeficientes para aos vetores $V[k]$ de primeiro plano reduzidos.

[0126] Como resultado, a unidade de análise de campo de som 44 pode determinar adicionalmente quando os coeficientes HOA ambiente mudam de quadro para quadro e gerar um sinalizador ou outro elemento sintático indicativo da alteração para o coeficiente HOA ambiente em termos de serem utilizados para representar os componentes ambiente do campo de som (onde a alteração também pode ser referida como uma "transição" do coeficiente HOA ambiente ou como uma "transição" do coeficiente HOA ambiente). Em particular, a unidade de redução de coeficientes 46 pode gerar o sinalizador (que pode ser denotado como um sinalizador `AmbCoeffTransition` ou um sinalizador `AmbCoeffIdxTransition`), fornecendo o sinalizador para a unidade de geração de bitstream 42 de modo que o

sinalizador pode ser incluído no bitstream 21 (possivelmente como parte da informação do canal lateral).

[0127] A unidade de redução de coeficientes 46 pode, além de especificar o indicador de transição de coeficiente ambiente, também modificar como são produzidos os vetores de primeiro plano $V[k]$ reduzidos 55. Em um exemplo, ao determinar que um dos coeficientes ambiente de HOA ambiente está em transição durante a quadro atual, a unidade de redução de coeficientes 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 de primeiro plano $V[k]$ reduzidos 55 que corresponde ao coeficiente de HOA ambiente em transição. Novamente, o coeficiente HOA ambiente em transição pode adicionar ou remover do número total de BG_{TOT} dos coeficientes de fundo. Por conseguinte, a alteração resultante no número total de coeficientes de fundo afeta se o coeficiente HOA ambiente está incluído ou não no bitstream e se o elemento correspondente dos vetores V está incluído para os vetores V especificados no bitstream nos segundo e terceiro modos de configuração descritos acima. Mais informações sobre como a unidade de redução de coeficientes 46 pode especificar os vetores $V[k]$ de primeiro plano reduzidos para superar as mudanças na energia são fornecidas no Pedido N° de Série 14/594.533, intitulada "TRANSITIONING OF AMBIENT HIGHER_ORDER AMBISONIC COEFFICIENTS", depositado em 12 de janeiro de 2015.

[0128] A FIG. 4 é um diagrama de blocos ilustrando o dispositivo de decodificação de áudio 24 da Fig. 2 em maior detalhe. Como mostrado no exemplo da FIG. 4 dispositivo de codificação de áudio 24 pode incluir uma unidade de extração 72, uma unidade de reconstrução de renderizador 81, uma unidade de reconstrução de base

direcional 90 e uma unidade de reconstrução baseada em vetor 92. Embora descrito a seguir, mais informação sobre o dispositivo de decodificação de áudio 24 e os vários aspectos da descompressão ou de outro modo decodificação de coeficientes de HOA estão disponíveis na Publicação de Pedido de Patente Internacional No. WO 2014/194099, intitulada "INTERPOLATION FOR DECOMPOSED REPRESENTATIONS OF A SOUND FIELD" depositado em 29 de maio de 2014.

[0129] A unidade de extração 72 pode representar uma unidade configurada para receber o bitstream 21 e extrair a informação de renderização de áudio 2 e as várias versões codificadas (por exemplo, uma versão codificada com base direcional ou uma versão codificada baseada no vetor) dos coeficientes HOA 11. Em outras palavras, as matrizes de renderização de Amnissônicos de Ordem Superior (HOA) podem ser transmitidas pelo dispositivo de codificação de áudio 20 para permitir o controle sobre o processo de renderização de HOA no sistema de reprodução de áudio 16. A transmissão pode ser facilitada por meio do mpegh3daConfigExtension do Tipo ID_CONFIG_EXT_HOA_MATRIX mostrado acima. O mpegh3daConfigExtension pode conter várias matrizes de renderização HOA para diferentes configurações de reprodução de alto-falante. Quando as matrizes de renderização de HOA são transmitidas, o dispositivo de codificação de áudio 20 sinaliza, para cada sinal de matriz de renderização HOA, o layout de alto-falante alvo associado que determina juntamente com o HoaOrder as dimensões da matriz de renderização.

[0130] A transmissão de um único HoaRenderingMatrixId permite referenciar uma matriz de renderização de HOA predefinida disponível no sistema de reprodução de áudio 16, ou a uma matriz de renderização de

HOA transmitida a partir do exterior do bitstream de áudio 21. Em alguns casos, cada matriz de renderização de HOA é assumida como normalizada em N3D e segue a ordenação dos coeficientes de HOA como definido no bitstream 21.

[0131] A função findSymmetricSpeakers pode, como referido acima, indicar um número e uma posição de todos os pares de alto-falantes dentro da configuração de alto-falante fornecida que são simétricos em relação, como um exemplo, ao plano médio de um ouvinte no chamado "sweet spot". Esta função auxiliar pode ser definida da seguinte forma:

```
int findSymmetricSpeakers(int outputCount,
SpeakerInformation* outputConfig, int hasLfeRendering);
```

A unidade de extração 72 pode invocar a função createSymSigns para calcular um vetor de valores 1,0 e -1,0 que podem então ser utilizados para gerar os elementos de matriz associados com alto-falantes simétricos. Esta função createSymSigns pode ser definida da seguinte forma:

```
void createSymSigns(int* symSigns, int hoaOrder)
{
    int n, m, k = 0;
    para (n = 0; n<=hoaOrder; ++n) {
        para (m = -n; m<=n; ++m)
            symSigns[k++] = ((m>=0)*2)-1;

    }
}
```

[0132] A unidade de extração 72 pode invocar a função create2dBitmask para gerar uma máscara de bit para identificar os coeficientes HOA que somente são usados no plano horizontal. A função create2dBitmask pode ser definida da seguinte forma:

```
void create2dBitmask(int* bitmask, int hoaOrder)
```

```

{
    int n, m, k = 0;
    bitmask[k++] = 0;
    para (n = 1; n<=hoaOrder; ++n) {
        para (m = -n; m<=n; ++m)
            bitmask[k++] = abs(m)!=n;
    }
}

```

[0133] Para decodificar os Coeficientes da Matriz de Renderização de HOA, a unidade de extração 72 pode primeiro extrair o elemento de sintaxe HoaRenderingMatrixSet(), o qual, como referido acima, pode conter uma ou mais matrizes de renderização de HOA que podem ser aplicadas para obter uma renderização de HOA para um layout de alto-falante desejado. Em alguns casos, um determinado bitstream pode não conter mais de uma instância de HoaRenderingMatrixSet(). O elemento de sintaxe HoaRenderingMatrix() contém a informação de matriz de renderização HOA (que pode ser indicada como informação de processador 2 no exemplo da FIG. 4). A unidade de extração 72 pode primeiro ler as informações de configuração, que podem guiar o processo de decodificação. Depois, a unidade de extração 72 lê os elementos da matriz em conformidade.

[0134] Em alguns casos, a unidade de extração 72, no início, lê os campos precisionLevel e gainLimitPerOrder. Quando o sinalizador gainLimitPerOrder é definida, a unidade de extração 72 lê e decodifica os campos maxGain e minGain para cada ordem HOA separadamente. Quando o sinalizador gainLimitPerOrder não está definido, a unidade de extração 72 lê e decodifica os campos maxGain e minGain uma vez e aplica estes campos a todas as ordens de HOA durante o processo de decodificação. Em alguns casos, o valor minGain deve estar entre 0db e -69dB. Em alguns

casos, o valor maxGain deve estar entre 1dB e 111dB menor que o valor minGain. A FIG. 9 é um diagrama ilustrando um exemplo de ganhos min e max dependentes de HOA dentro de uma matriz de renderização de HOA.

[0135] A unidade de extração 72 pode depois ler o sinalizador isFullMatrix, o qual pode sinalizar se uma matriz é definida como cheia ou como parcialmente esparsa. Quando a matriz é definida como parcialmente esparsa, a unidade de extração 72 lê o campo seguinte (por exemplo, o elemento de sintaxe firstSparseOrder), que especifica a ordem HOA a partir da qual a matriz de renderização HOA é codificada de forma esparsa. As matrizes de renderização de HOA podem muitas vezes ser esparsas para baixa ordem e tornar-se esparsas nas ordens mais altas, dependendo da configuração da reprodução do alto-falante. A FIG. 10 é um diagrama ilustrando uma matriz de renderização HOA de 6^a ordem parcialmente esparsa para 22 alto-falantes. A esparsidade da matriz ilustrada na FIG. 10 começa com o 26º coeficiente de HOA (HOA ordem 5).

[0136] Dependendo se existe um ou mais canais de efeitos de baixa frequência (LFE) dentro da configuração de reprodução do alto-falante (indicada pelo elemento de sintaxe lfeExists), a unidade de extração 72 pode ler o campo hasLfeRendering. Quando hasLfeRendering não é definido, a unidade de extração 72 é configurada para assumir que os elementos de matriz relacionados com os canais LFE são zeros digitais. O campo seguinte lido pela unidade de extração 72 é o indicador zerothOrderAlwaysPositive, que sinaliza se os elementos de matriz associados com o coeficiente da ordem 0 são positivos. Neste caso em que zerothOrderAlwaysPositive indica que os coeficientes de HOA de ordem zero são positivos, a unidade de extração 72 determina que os sinais

numéricos não são codificados para os coeficientes de matriz de renderização correspondentes aos coeficientes HOA de ordem zero.

[0137] No que se segue, as propriedades da matriz de renderização HOA podem ser sinalizadas para pares de alto-falantes simétricos em relação ao plano médio. Em alguns casos, existem duas propriedades de simetria relativas a) simetria de valores e b) simetria de sinal. No caso da simetria de valor, os elementos de matriz do alto-falante esquerdo do par de alto-falantes simétricos não são codificados, mas sim a unidade de extração 72 deriva aqueles elementos dos elementos de matriz decodificados do alto-falante direito utilizando a função auxiliar `createSymSigns` que executa o seguinte:

```
pairIdx      =     outputConfig[j].symmetricPair-
>originalPosition;

    hoaMatrix[i * outputCount + j] = hoaMatrix[i *
outputCount + pairIdx]; and

    signMatrix[i * outputCount + j] = symSigns[i] *
signMatrix[i * outputCount + pairIdx].
```

[0138] Quando um par de alto-falantes não é simétrico em valor, então os elementos de matriz podem ser simétricos em relação aos seus sinais numéricos. Quando um par de alto-falantes é simétrico em sinal, os sinais numéricos dos elementos de matriz do alto-falante esquerdo do par de alto-falantes simétricos não são codificados e a unidade de extração 72 deriva estes sinais numéricos dos sinais numéricos dos elementos de matriz associados com o alto-falante direito usando a função auxiliar `createSymSigns`, que executa o seguinte:

```
pairIdx      =     outputConfig[j].symmetricPair-
>originalPosition;
```

```

    signMatrix[i * outputCount + j] = symSigns[i] *
signMatrix[i * outputCount + pairIdx];

```

[0139] A FIG. 11 é um diagrama que ilustra a sinalização das propriedades de simetria. Um par de alto-falantes não pode ser definido como simétrico em valor e simétrico em sinal ao mesmo tempo. O sinalizador de decodificação final hasVerticalCoef especificado se apenas os elementos de matriz associados com coeficientes de HOA circulares (isto é, 2D) são codificados. Se hasVerticalCoef não for definido, os elementos de matriz associados aos coeficientes HOA definidos com a função auxiliar create2dBitmask são definidos como zero digital.

[0140] Isto é, a unidade de extração 72 pode extrair a informação de renderização de áudio 2 de acordo com o processo apresentado na FIG. 11. A unidade de extração 72 pode primeiro ler o elemento de sintaxe isAllValueSymmetric a partir do bitstream 21 (300). Quando o elemento de sintaxe isAllValueSymmetric é definido como um (ou, em outras palavras, um booleano verdadeiro), a unidade de extração 72 pode iterar através do valor do elemento de sintaxe numPairs, definindo o elemento de sintaxe do arranjo valueSymmetricPairs como um valor de um (efetivamente indicando que todos os pares de alto-falantes são valores simétricos) (302).

[0141] Quando o elemento de sintaxe isAllValueSymmetric é definido como zero (ou, em outras palavras, um booleano falso), a unidade de extração 72 pode ler em seguida o elemento de sintaxe isAnyValueSymmetric (304). Quando o elemento de sintaxe isAnyValueSymmetric é definido como um (ou, em outras palavras, um booleano verdadeiro), a unidade de extração 72 pode iterar através do valor do elemento de sintaxe numPairs, definindo o elemento de sintaxe do arranjo valueSymmetricPairs para um

bit lido sequencialmente a partir do bitstream (306). A unidade de extração 72 também pode obter o elemento de sintaxe `isAnySignSymmetric` para qualquer um dos pares tendo um elemento de sintaxe `valueSymmetricPairs` definido como zero (308). A unidade de extração 72 pode então iterar através do número de pares novamente e, quando o `valueSymmetricPairs` for igual a zero, definir um bit de `signSymmetricPairs` para um valor lido a partir do bitstream 21 (310).

[0142] Quando o elemento de sintaxe `isAnyValueSymmetric` é definido como zero (ou, em outras palavras, um booleano falso), a unidade de extração 72 pode ler em seguida o elemento de sintaxe `isAllValueSymmetric` a partir do bitstream 21 (312). Quando o elemento de sintaxe `isAllSignSymmetric` é definido como um valor de um (ou, em outras palavras, um booleano verdadeiro), a unidade de extração 72 pode iterar através do valor do elemento de sintaxe `numPairs`, definindo o elemento de sintaxe do arranjo `signSymmetricPairs` como um valor de um (efetivamente indicando que todos os pares de alto-falantes são simétricos em sinal) (316).

[0143] Quando o elemento de sintaxe `isAllSignSymmetric` é definido como zero (ou, em outras palavras, um booleano falso), a unidade de extração 72 pode ler em seguida o elemento de sintaxe `isAnySignSymmetric` a partir do bitstream 21 (316). A unidade de extração 72 pode iterar através do valor do elemento de sintaxe `numPairs`, definindo o elemento de sintaxe do arranjo `signSymmetricPairs` para um bit lido sequencialmente a partir do bitstream (318). A unidade de geração de bitstream 42 pode executar um processo recíproco em relação ao descrito acima em relação à unidade de extração 72 para especificar a informação de simetria de valor, a informação

de simetria de sinal ou uma combinação de ambas as informações de simetria de valor e de sinal.

[0144] A unidade de reconstrução de renderizador 81 pode representar uma unidade de configuração d para reconstruir um processador com base na informação de renderização de áudio 2. Isto é, usando as propriedades acima mencionadas, a unidade de reconstrução de renderizador 81 pode ler uma série de valores de ganho de elemento de matriz. Para ler o valor de ganho absoluto, a unidade de reconstrução de renderizador 81 pode invocar a função DecodeGainValue(). A unidade de reconstrução de renderizador 81 pode invocar a função ReadRange() do índice de alfabeto para decodificar uniformemente os valores de ganho. Quando o valor de ganho decodificado não é um zero digital, a unidade de reconstrução de renderizador 81 pode ler o valor do sinal numérico adicionalmente (de acordo com a Tabela a abaixo). Quando o elemento de matriz está associado a um coeficiente de HOA que foi sinalizado como esparsa (via isHoaCoefSparse) o sinalizador hasValue precede o valor de gainValueIndex (ver Tabela b). Quando o sinalizador hasValue é zero, este elemento é definido como zero digital e nenhum gainValueIndex e sinal são sinalizados.

Tabelas a e b - Exemplos de sintaxe de bitstream para decodificar um elemento de matriz

bitfield size	gainValueIndex	sign
	alphabetSize	1

a)

bitfield size	hasValue	gainValueIndex	sign
	1	alphabetSize	1

b)

[0145] Dependendo das propriedades de simetria especificadas para pares de alto-falantes, a unidade de

reconstrução de renderização 81 pode derivar os elementos de matriz associados com o alto-falante esquerdo a partir do alto-falante direito. Neste caso, a informação de renderização de áudio 2 no bitstream 21 para decodificar um elemento de matriz para o alto-falante esquerdo é reduzida ou potencialmente completamente omitida em conformidade.

[0146] Deste modo, o dispositivo de decodificação de áudio 24 pode determinar a informação de simetria para reduzir um tamanho da informação de renderização de áudio a ser especificada. Em alguns casos, o dispositivo de decodificação de áudio 24 pode determinar informações de simetria para reduzir um tamanho da informação de renderização de áudio a ser especificada e derivar pelo menos uma porção do processador de áudio com base na informação de simetria.

[0147] Nestes e em outros casos, o dispositivo de decodificação de áudio 24 pode determinar a informação de simetria de valor para reduzir um tamanho da informação de renderização de áudio a ser especificada. Nestes e em outros casos, o dispositivo de decodificação de áudio 24 pode derivar pelo menos uma porção do processador de áudio com base na informação de simetria de valor.

[0148] Nestes e em outros casos, o dispositivo de decodificação de áudio 24 pode determinar a informação de simetria de sinal para reduzir um tamanho da informação de renderização de áudio a ser especificada. Nestes e em outros casos, o dispositivo de decodificação de áudio 24 pode derivar pelo menos uma porção do processador de áudio com base na informação de simetria de sinal.

[0149] Nestes e em outros casos, o dispositivo de decodificação de áudio 24 pode determinar a informação de esparsidade indicativa de uma esparsidade de uma matriz

utilizada para renderizar coeficientes harmônicos esféricos para uma pluralidade de feeds de alto-falante.

[0150] Nestes e em outros casos, o dispositivo de decodificação de áudio 24 pode determinar um layout de alto-falante para o qual uma matriz deve ser usada para renderizar coeficientes harmônicos esféricos para uma pluralidade de feeds de alto-falante.

[0151] O dispositivo de decodificação de áudio 24, nesse aspecto, pode então determinar a informação de renderização de áudio 2 especificada no bitstream. Com base no valor de sinal incluído na informação de renderização de áudio 2, o sistema de reprodução de áudio 16 pode renderizar uma pluralidade de feeds de alto-falante 25 usando um dos renderizadores de áudio 22. Os feeds de alto-falante podem acionar os alto-falantes 3. Como observado acima, o valor de sinal pode, em alguns casos, incluir uma matriz (que é decodificada e fornecida como um dos renderizadores de áudio 22) usada para renderizar coeficientes harmônicos esféricos para uma pluralidade de feeds de alto-falante. Neste caso, o sistema de reprodução de áudio 16 pode configurar um dos renderizadores de áudio 22 com a matriz, utilizando este um dos renderizadores de áudio 22 para renderizar os feeds de alto-falante 25 com base na matriz.

[0152] Para extrair e depois decodificar as várias versões codificadas dos coeficientes HOA 11 para que os coeficientes HOA 11 estejam disponíveis para serem processados utilizando o processador de áudio 22 obtido, a unidade de extração 72 pode determinar a partir do elemento de sintaxe observado acima indicativo de se os coeficientes HOA 11 foram codificados através das várias versões baseadas em direção ou baseadas em vetor. Quando uma codificação baseada em direção foi realizada, a unidade de

extração 72 pode extrair a versão com base direcional dos coeficientes HOA 11 e os elementos de sintaxe associados à versão codificada (que é denotado como informação de base direcional 91 no exemplo da FIG. 4), passando a informação de base direcional 91 para a unidade de reconstrução com base direcional 90. A unidade de reconstrução com base direcional 90 pode representar uma unidade configurada para reconstruir os coeficientes HOA na forma de coeficientes HOA 11' com base na informação baseada em direção 91.

[0153] Quando o elemento de sintaxe indica que os coeficientes HOA 11 foram codificados utilizando uma decomposição baseada em vetor, a unidade de extração 72 pode extrair os vetores $V[k]$ codificados em primeiro plano 57 (que podem incluir pesos codificados 57 e/ou índices 63 ou vetores V quantizados escalares), os coeficientes HOA ambiente codificados 59 e os objetos de áudio correspondentes 61 (que também podem ser referidos como os sinais de 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 de primeiro plano $V[k]$ codificados para a unidade de reconstrução de vetor V 74 e os coeficientes HOA de ambiente codificados 59 juntamente com os sinais de nFG codificados 61 para a unidade de decodificação psicoacústica 80.

[0154] A unidade de reconstrução de vetor V 74 pode representar uma unidade configurada para reconstruir os vetores V a partir dos vetores de primeiro plano codificados $V[k]$ 57. A unidade de reconstrução do vetor V 74 pode funcionar de um modo recíproco ao da unidade de quantização 52.

[0155] A unidade de decodificação psicoacústica 80 pode funcionar de um modo recíproco à unidade de codificador de áudio psicoacústico 40 mostrada

no exemplo da FIG. 3 de modo a decodificar os coeficientes HOA codificados ambiente 59 e os sinais de nFG codificados 61 e assim gerar coeficientes HOA ambiente compensados em energia 47' e os sinais de 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 HOA ambiente compensados em energia 47' para a unidade de desvanecimento 770 e os sinais de nFG 49' para a unidade de formulação de primeiro plano 78.

[0156] A unidade de interpolação espaço-temporal 76 pode funcionar de uma maneira semelhante à descrita 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 relação aos vetores $V[k]$ de primeiro plano 55_k e aos vetores de primeiro plano $V[k-1]$ reduzidos 55_{k-1} para gerar vetores $V[k]$ interpolados de primeiro plano $55_k''$. A unidade de interpolação espaço-temporal 76 pode encaminhar os vetores de primeiro plano interpolados $V[k] 55_k''$ para a unidade de desvanecimento 770.

[0157] A unidade de extração 72 também pode emitir um sinal 757 indicativo de quando um dos coeficientes HOA ambiente está em transição para a unidade de desvanecimento 770, que pode então determinar qual dos SHC_{BG} 47' (onde o SHC_{BG} 47' também pode ser designado como "coeficientes HOA ambiente 47'" ou "coeficientes HOA ambiente 47'") e os elementos dos vetores $[V][k]$ de primeiro plano interpolados $55_k''$ devem aparecer ou desvanecer. Em alguns exemplos, a unidade de desvanecimento 770 pode operar oposta em relação a cada um dos coeficientes HOA ambiente 47' e os elementos dos vetores de primeiro plano $V[k]$ interpolados $55_k''$. Ou seja,

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

[0158] A unidade de formulação de primeiro plano 78 pode representar uma unidade configurada para realizar a multiplicação de matriz em relação aos vetores de primeiro plano $V[k]$ ajustados $55_k'''$ e os sinais de nFG interpolados 49' para gerar os coeficientes HOA de primeiro plano 65. Nesse aspecto, a unidade de formulação de primeiro plano 78 pode combinar os objetos de áudio 49' (que é outra forma de designar os sinais de nFG interpolados 49') com os vetores $55_k''$ para reconstruir o primeiro plano ou, em outras palavras, aspectos predominantes dos coeficientes HOA 11'. A unidade de formulação de primeiro plano 78 pode realizar uma multiplicação de matriz dos sinais de nFG interpolados 49' pelos vetores de primeiro plano ajustados $V[k] 55_k''$.

[0159] A unidade de formulação de coeficientes HOA 82 pode representar uma unidade configurada para combinar os coeficientes HOA de primeiro plano 65 com os coeficientes HOA ajustados 47" de modo a obter os coeficientes HOA 11'. A notação primária reflete que os coeficientes HOA 11' podem ser semelhantes, mas não iguais aos coeficientes HOA 11. As diferenças entre os coeficientes HOA 11 e 11' podem resultar da perda devido a transmissão através de uma mídia de transmissão com perdas, quantização ou outras operações com perdas.

[0160] Além disso, a unidade de extração 72 e o dispositivo de decodificação de áudio 24 mais genericamente também podem ser configurados para operarem de acordo com vários aspectos das técnicas descritas nesta revelação para obter os bitstreams 21 que são potencialmente otimizados nas formas descritas acima em relação a não incluir vários elementos de sintaxe ou campos de dados em determinadas instâncias.

[0161] Em alguns casos, o dispositivo de decodificação de áudio 24 pode ser configurado para, ao descomprimir dados de áudio ambissônicos de ordem superior comprimidos utilizando um primeiro esquema de compressão, obter um bitstream 21 representativo de uma versão comprimida dos dados de áudio ambissônicos de ordem superior que não inclui bits correspondentes para um segundo esquema de compressão também usado para comprimir os dados de áudio ambissônicos de ordem superior. O primeiro esquema de compressão pode compreender um esquema de compressão baseado em vetor, o vetor resultante definido no domínio harmônico esférico e enviado através do bitstream 21. O esquema de compressão de decomposição baseado em vetor pode, em alguns exemplos, compreender um esquema de compressão que envolve a aplicação de uma

decomposição de valor singular (ou equivalentes do mesmo como descrito em maior detalhe em relação ao exemplo da FIG. 3) para os dados de áudio ambissônicos de ordem superior.

[0162] O dispositivo de decodificação de áudio 24 pode ser configurado para obter o bitstream 21 que não inclui os bits que correspondem a pelo menos um elemento de sintaxe utilizado para executar o segundo tipo de esquema de compressão. Como observado acima, o segundo esquema de compressão compreende um esquema de compressão baseado em direcionalidade. Mais especificamente, o dispositivo de decodificação de áudio 24 pode ser configurado para obter o bitstream 21 que não inclua os bits correspondentes a um elemento de sintaxe HOAPredictionInfo do segundo esquema de compressão. Em outras palavras, quando o segundo esquema de compressão compreende um esquema de compressão baseado em direcionalidade, o dispositivo de decodificação de áudio 24 pode ser configurado para obter o bitstream 21 que não inclua os bits correspondentes a um elemento de sintaxe HOAPredictionInfo do esquema de compressão baseado na direcionalidade. Conforme mencionado acima, o elemento de sintaxe HOAPredictionInfo pode ser indicativo de uma previsão entre dois ou mais sinais baseados em direcionalidade.

[0163] Em alguns casos, ou como uma alternativa ou em conjunto com os exemplos anteriores, o dispositivo de decodificação de áudio 24 pode ser configurado para, quando a correção de ganho é suprimida durante a compressão de dados de áudio ambissônicos de ordem superior, obter o bitstream 21 representativo de uma versão comprimida dos dados de áudio ambissônicos de ordem superior que não incluem dados de correção de ganho. O dispositivo de decodificação de áudio 24 pode, nestas

instâncias, ser configurado para descomprimir os dados de áudio ambissônicos de ordem superior de acordo com um esquema de descompressão de síntese baseado em vetor. A versão comprimida dos dados ambissônicos de ordem superior é gerada através da aplicação de uma decomposição de valor singular (ou seus equivalentes descritos em maior detalhe em relação ao exemplo da FIG. 3 acima) para os dados de áudio ambissônicos de ordem superior. Quando SVD é aplicado ou equivalentes aos dados de áudio de HOA, o dispositivo de codificação de áudio 20 especifica pelo menos um dos vetores ou bits resultantes indicativos no bitstream 21, onde os vetores descrevem características espaciais de objetos de áudio de primeiro plano correspondentes (tais como um largura, localização e volume dos objetos de áudio de primeiro plano correspondentes).

[0164] Mais especificamente, o dispositivo de decodificação de áudio 24 pode ser configurado para obter um elemento de sintaxe MaxGainCorrAmbExp a partir do bitstream 21 com um valor definido como zero para indicar que a correção de ganho é suprimida. Ou seja, o dispositivo de codificação de áudio 24 pode ser configurado para obter, quando a correção de ganho é suprimida, o bitstream de modo que o bitstream não inclua um campo de dados de HOAGainCorrection que armazene os dados de correção de ganho. O bitstream 21 pode compreender um elemento de sintaxe MaxGainCorrAmbExp tendo um valor de zero para indicar que a correção de ganho é suprimida e não incluir um campo de dados de HOAGainCorrection que armazena os dados de correção de ganho. A supressão da correção de ganho pode ocorrer quando a compressão dos dados de áudio ambissônicos de ordem superior inclui a aplicação de uma codificação de fala e áudio e fala unificada (USAC) aos dados de áudio ambissônicos de ordem superior.

[0165] A FIG. 5 é um fluxograma que ilustra a operação exemplar de um dispositivo de codificação de áudio, como o dispositivo de codificação de áudio 20 mostrado no exemplo da Fig. 3 na realização de vários aspectos das técnicas de síntese baseada em vetor descritas nesta revelação. Inicialmente, o dispositivo de codificação de áudio 20 recebe os coeficientes HOA 11 (106). O dispositivo de codificação de áudio 20 pode invocar a unidade LIT 30, que pode aplicar um LIT em relação aos coeficientes HOA para produzir os coeficientes HOA transformados (por exemplo, no caso de SVD, os coeficientes 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 realizar a análise acima descrita com relação 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 do modo descrito acima. Ou seja, a unidade de cálculo de parâmetros 32 pode determinar pelo menos um parâmetro com base em uma análise dos coeficientes HOA transformados 33/35 (108).

[0167] O dispositivo de codificação de áudio 20 pode então invocar a unidade de reordenação 34, a qual pode reordenar os coeficientes HOA transformados (que, novamente no contexto de SVD, podem referir-se aos vetores US[k] 33 e V[k] 35 com base nos parâmetros para gerar os coeficientes HOA transformados reordenados 33'/35' (ou, em outras palavras, os vetores US[k] 33' e os vetores V[k] 35'), como descrito acima (109). O dispositivo de codificação de áudio 20 pode, durante qualquer das operações precedentes ou operações subsequentes, invocar também a unidade de análise de campo de som 44. A unidade

de análise de campo de som 44 pode, como descrito acima, realizar uma análise de campo de som com relação aos coeficientes HOA 11 e/ou aos coeficientes HOA transformados 33/35 para determinar o número total de canais de primeiro plano (nFG) 45, a ordem do campo de som de fundo (N_{BG}) e o número (nBGA) e os índices (i) de canais BG HOA adicionais para enviar (que podem ser coletivamente denotados como informação de canal de fundo 43 no exemplo da FIG. 3) (109).

[0168] O dispositivo de codificação de áudio 20 também pode invocar a unidade de seleção de fundo 48. A unidade de seleção de fundo 48 pode determinar coeficientes HOA de fundo ou ambiente 47 com base na informação de canal de fundo 43 (110). A unidade 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 baseado em nFG 45 (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 realizar a compensação de energia com relação aos coeficientes HOA ambiente 47 para compensar a perda de energia devido à remoção de vários dos coeficientes HOA pela unidade de seleção de fundo 48 (114) e assim gerar os coeficientes HOA 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 relação aos coeficientes 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 coeficientes 46 pode executar uma redução de coeficientes em relação aos vetores V[k] de primeiro plano restantes com base na informação de canal de fundo 43 para obter informação de direção de primeiro plano reduzida 55 (que também pode ser referida como os vetores de primeiro plano reduzido V[k] 55) (118).

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

[0172] O dispositivo de codificação de áudio 20 também pode invocar a unidade de codificador de áudio psicoacústica 40. A unidade de codificador de áudio psicoacústica 40 pode codificar psicoacusticamente cada vetor dos coeficientes HOA ambiente com compensação de energia 47' e os sinais de nFG interpolados 49' para gerar coeficientes HOA ambiente codificados 59 e sinais de nFG codificados 61. O dispositivo de codificação de áudio pode então invocar a unidade de geração de bitstream 42. A unidade de geração de bitstream 42 pode então gerar o bitstream 21 baseado na informação direcional de primeiro plano codificada 57, nos coeficientes HOA ambiente codificados 59, nos sinais de nFG codificados 61 e nas informações de canal de fundo 43.

[0173] A FIG. 6 é um fluxograma que ilustra a operação exemplar de um dispositivo de decodificação de

áudio, como o dispositivo de decodificação de áudio 24 mostrado na Fig. 4 na realização de vários aspectos das técnicas descritas nesta revelação. Inicialmente, o dispositivo de decodificação de áudio 24 pode receber o bitstream 21 (130). Ao receber o bitstream, o dispositivo de decodificação de áudio 24 pode invocar a unidade de extração 72. Assumindo, para fins de discussão, que o bitstream 21 indica que a reconstrução baseada em vetor deve ser executada, a unidade de extração 72 pode analisar o bitstream para recuperar a informação acima mencionada, passando a informação para a unidade de reconstrução baseada em vetor 92.

[0174] Em outras palavras, a unidade de extração 72 pode extrair a informação direcional de primeiro plano codificada 57 (que, de novo, também pode ser referida como os vetores de primeiro plano codificados $V[k]$ 57), os coeficientes HOA ambiente codificados 59 e os sinais de primeiro plano codificados que também podem ser referidos como os sinais de nFG de primeiro plano codificados 59 ou os objetos de áudio de primeiro plano codificados 59) a partir do bitstream 21 da maneira descrita acima (132).

[0175] O dispositivo de decodificação de áudio 24 pode ainda invocar a unidade de desquantização 74. A unidade de desquantização 74 pode decodificar e desquantizar por entropia a informação direcional de primeiro plano codificada 57 para obter informação direcional de primeiro plano reduzida 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 HOA ambiente codificados 59 e os sinais de primeiro plano codificados 61 para obter

coeficientes HOA ambiente compensados em energia 47' e os sinais de primeiro plano interpolados 49' (138). A unidade de decodificação psicoacústica 80 pode passar os coeficientes HOA ambiente compensados em energia 47' para a unidade de desvanecimento 770 e os sinais de nFG 49' para a unidade de formulação de primeiro plano 78.

[0176] 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 a informação direcional reordenada de primeiro plano 55_k' e realizar a interpolação espaço-temporal com relação à informação direcional de primeiro plano 55_k/55_{k-1} reduzida para gerar a informação interpolada 55_k" (140) de direção de primeiro plano. A unidade de interpolação espaço-temporal 76 pode encaminhar os vetores de primeiro plano interpolados V[k] 55_k'' para a unidade de desvanecimento 770.

[0177] O dispositivo de decodificação de áudio 24 pode invocar a unidade de desvanecimento 770. A unidade de desvanecimento 770 pode receber ou de outro modo obter elementos de sintaxe (por exemplo, a partir da unidade de extração 72) indicativa de quando os coeficientes HOA ambiente compensados em energia 47' estão em transição (por exemplo, o elemento de sintaxe AmbCoeffTransition). A unidade de desvanecimento 770 pode, com base nos elementos de sintaxe de transição e na informação de estado de transição mantida, fazer surgir ou desvanecer os coeficientes HOA ambiente com compensação de energia 47', fornecendo coeficientes HOA 47" ajustados para a unidade de formulação de coeficiente HOA 82. A unidade de desvanecimento 770 pode também, com base nos elementos de sintaxe e na informação de estado de transição mantida, e fazer surgir ou desvanecer os um ou mais elementos dos

vetores de primeiro plano interpolados $V[k] 55_k''$ correspondentes emitindo os vetores de primeiro plano ajustados $V[k] 55_k'''$ para a unidade de formulação de primeiro plano 78 (142).

[0178] 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 da matriz dos sinais de nFG 49' pela informação direcional de primeiro plano ajustada $55_k'''$ para obter os coeficientes HOA de primeiro plano 65 (144). O dispositivo de decodificação de áudio 24 também pode invocar a unidade de formulação de coeficiente HOA 82. A unidade de formulação de coeficientes HOA 82 pode adicionar os coeficientes HOA de primeiro plano 65 aos coeficientes HOA ajustados 47" de modo a obter os coeficientes HOA 11' (146).

[0179] A FIG. 7 é um fluxograma que ilustra a operação exemplar de um sistema, como o sistema 10 mostrado no exemplo da Fig. 2 na realização de vários aspectos das técnicas descritas nesta revelação. Conforme discutido acima, o dispositivo criador de conteúdo 12 pode utilizar o sistema de edição de som 18 para criar ou editar conteúdos de áudio capturados ou gerados (o que é mostrado como os coeficientes HOA 11 no exemplo da FIG. 2). O dispositivo criador de conteúdo 12 pode então renderizar os coeficientes HOA 11 utilizando o processador de áudio 1 para os feeds de alto-falante de multicanal gerados, como discutido em maior detalhe acima (200). O dispositivo criador de conteúdo 12 pode então reproduzir estes feeds de alto-falantes utilizando um sistema de reprodução de áudio e determinar se são necessários mais ajustes ou edição para capturar, por exemplo, a intenção artística desejada (202). Quando são desejados mais ajustes ("SIM" 202), o

dispositivo criador de conteúdo 12 pode remixar os coeficientes HOA (204), renderizar os coeficientes HOA (11) (200) e determinar se são necessários outros ajustes (202). Quando não são desejados mais ajustes ("NÃO" 202), o dispositivo de codificação de áudio 20 pode codificar o conteúdo de áudio para gerar o bitstream 21 da maneira descrita acima em relação ao exemplo da FIG. 5) (206). O dispositivo de codificação de áudio 20 também pode gerar e especificar a informação de renderização de áudio 2 no bitstream 21, como descrito em maior detalhe acima (208).

[0180] O dispositivo consumidor de conteúdo 14 pode então obter a informação de renderização de áudio 2 a partir do bitstream 21 (210). O dispositivo de decodificação 24 pode então decodificar o bitstream 21 para obter o conteúdo de áudio (que é mostrado como os coeficientes de HOA 11' no exemplo da FIG. 2 da maneira acima descrita com relação ao exemplo da FIG. 6) (211). O sistema de reprodução de áudio 16 pode então renderizar os coeficientes HOA 11' com base na informação de renderização de áudio 2 da maneira descrita acima (212) e reproduzir o conteúdo de áudio renderizado através dos alto-falantes 3 (214).

[0181] As técnicas descritas nesta revelação podem, por conseguinte, permitir, como um primeiro exemplo, um dispositivo que gera um bitstream representativo de conteúdo de áudio de multi-canal para especificar informação de renderização de áudio. O dispositivo pode, neste primeiro exemplo, incluir meios para especificar a informação de renderização de áudio que inclui um valor de sinal que identifica um processador de áudio utilizado quando se gera o conteúdo de áudio multicanal.

[0182] O dispositivo do primeiro exemplo, sendo que o valor de sinal inclui uma matriz usada para

renderizar coeficientes harmônicos esféricos para uma pluralidade de feeds de alto-falante.

[0183] Em um segundo exemplo, o dispositivo do primeiro exemplo, sendo que o valor de sinal inclui dois ou mais bits que definem um índice que indica que o bitstream inclui uma matriz utilizada para renderizar coeficientes harmônicos esféricos para uma pluralidade de feeds de alto-falante.

[0184] O dispositivo do segundo exemplo, sendo que a informação de renderização de áudio inclui ainda dois ou mais bits que definem um número de linhas da matriz incluída no bitstream e dois ou mais bits que definem um número de colunas da matriz incluída no bitstream.

[0185] O dispositivo do primeiro exemplo, sendo que o valor de sinal especifica um algoritmo de renderização usado para renderizar objetos de áudio para uma pluralidade de feeds de alto-falante.

[0186] O dispositivo do primeiro exemplo, sendo que o valor de sinal especifica um algoritmo de renderização usado para renderizar coeficientes harmônicos esféricos para uma pluralidade de feeds de alto-falante.

[0187] O dispositivo do primeiro exemplo, sendo que o valor de sinal inclui dois ou mais bits que definem um índice associado com uma de uma pluralidade de matrizes utilizadas para renderizar coeficientes harmônicos esféricos para uma pluralidade de feeds de alto-falante.

[0188] O dispositivo do primeiro exemplo, sendo que o valor de sinal inclui dois ou mais bits que definem um índice associado com uma de uma pluralidade de algoritmos de renderização usados para renderizar objetos de áudio para uma pluralidade de feeds de alto-falante.

[0189] O dispositivo do primeiro exemplo, sendo que o valor de sinal inclui dois ou mais bits que

definem um índice associado com uma de uma pluralidade de algoritmos de renderização utilizadas para renderizar coeficientes harmônicos esféricos para uma pluralidade de feeds de alto-falante.

[0190] O dispositivo do primeiro exemplo, sendo que os meios para especificar a informação de renderização de áudio compreendem meios para especificar a informação de renderização de áudio em uma base por quadro de áudio no bitstream.

[0191] O dispositivo do primeiro exemplo, sendo que os meios para especificar a informação de renderização de áudio compreendem meios para especificar a informação de renderização de áudio em um único momento no bitstream.

[0192] Em um terceiro exemplo, uma mídia de armazenamento legível por computador não transitória que tem instruções armazenadas na mesma que, quando executadas, fazem com que um ou mais processadores especifiquem a informação de renderização de áudio no bitstream, sendo que a informação de renderização de áudio identifica um processador de áudio utilizado ao gerar o conteúdo de áudio multicanal.

[0193] Em um quarto exemplo, um dispositivo para renderizar conteúdo de áudio multicanal a partir de um bitstream, o dispositivo compreendendo meios para determinar informação de renderização de áudio que inclui um valor de sinal que identifica um renderizador de áudio usado quando se gera o conteúdo de áudio multicanal e meios para renderizar uma pluralidade de feeds de alto-falante com base na informação de renderização de áudio especificada no bitstream.

[0194] O dispositivo do quarto exemplo, sendo que o valor de sinal inclui uma matriz utilizada para

renderizar coeficientes harmônicos esféricos para uma pluralidade de feeds de alto-falante, e sendo que os meios para renderizar a pluralidade de feeds de alto-falante compreendem meios para renderizar a pluralidade de feeds de alto-falantes com base na matriz.

[0195] Em um quinto exemplo, o dispositivo do quarto exemplo, sendo que o valor de sinal inclui dois ou mais bits que definem um índice que indica que o bitstream inclui uma matriz utilizada para renderizar coeficientes harmônicos esféricos para uma pluralidade de feeds de alto-falante, em que o dispositivo adicionalmente compreendendo meios para analisar a matriz a partir do bitstream em resposta ao índice e sendo que os meios para renderizar a pluralidade de feeds de alto-falante compreendem meios para renderizar a pluralidade de feeds de alto-falante com base na matriz analisada.

[0196] O dispositivo do quinto exemplo, sendo que o valor de sinal inclui ainda dois ou mais bits que definem um número de linhas da matriz incluída no bitstream e dois ou mais bits que definem um número de colunas da matriz incluída no bitstream, e sendo que os meios para analisar a matriz a partir do bitstream compreende meios para analisar a matriz do bitstream em resposta ao índice e com base nos dois ou mais bits que definem um número de linhas e os dois ou mais bits que definem o número de colunas.

[0197] O dispositivo do quarto exemplo, sendo que o valor de sinal especifica o algoritmo de renderização usado para renderizar objetos de áudio para a pluralidade de feeds de alto-falante, e sendo que os meios para renderizar a pluralidade de feeds de alto-falante compreendem meios para renderizar a pluralidade de feeds de

alto-falantes a partir dos objetos de áudio usando o algoritmo de renderização especificado.

[0198] O dispositivo do quarto exemplo, sendo que o valor de sinal especifica o algoritmo de renderização usado para renderizar coeficientes harmônicos esféricos para a pluralidade de feeds de alto-falante, e sendo que os meios para renderizar a pluralidade de feeds de alto-falante compreendem meios para renderizar a pluralidade de feeds de alto-falantes a partir dos coeficientes harmônicos esféricos usando o algoritmo de renderização especificado

[0199] O dispositivo do quarto exemplo, sendo que o valor de sinal inclui dois ou mais bits que definem um índice associado com uma da pluralidade de matrizes usada para renderizar coeficientes harmônicos esféricos para a pluralidade de feeds de alto-falante, e sendo que os meios para renderizar a pluralidade de feeds de alto-falante compreendem meios para renderizar a pluralidade de feeds de alto-falantes a partir dos coeficientes harmônicos esféricos usando uma da pluralidade de matrizes associadas com o índice.

[0200] O dispositivo do quarto exemplo, sendo que o valor de sinal inclui dois ou mais bits que definem um índice associado com uma da pluralidade de algoritmos de renderização usados para renderizar os objetos de áudio para a pluralidade de feeds de alto-falante, e sendo que os meios para renderizar a pluralidade de feeds de alto-falante compreendem meios para renderizar a pluralidade de feeds de alto-falantes a partir dos objetos de áudio usando uma da pluralidade de algoritmos de renderização associadas com o índice.

[0201] O dispositivo do quarto exemplo, sendo que o valor de sinal inclui dois ou mais bits que definem um índice associado com uma da pluralidade de algoritmos de

renderização usados para renderizar coeficientes harmônicos esféricos para uma pluralidade de feeds de alto-falante, e sendo que os meios para renderizar a pluralidade de feeds de alto-falante compreendem meios para renderizar a pluralidade de feeds de alto-falantes a partir dos coeficientes harmônicos esféricos usando uma da pluralidade de algoritmos de renderização associadas com o índice.

[0202] O dispositivo do quarto exemplo, sendo que os meios para determinar a informação de renderização de áudio incluem meios para determinar a informação de renderização de áudio em uma base por quadro de áudio a partir do bitstream.

[0203] O dispositivo do quarto exemplo, sendo que os meios para determinar os meios de informação de renderização de áudio incluem determinar a informação de renderização de áudio em um momento único a partir do bitstream.

[0204] Em um sexto exemplo, uma mídia de armazenamento legível por computador não transitória tendo armazenadas nela informações que, quando executadas, fazem com que um ou mais processadores determinem a informação de renderização de áudio que inclui um valor de sinal que identifica um renderizador de áudio usado quando se gera o conteúdo de áudio multicanal; e renderizem uma pluralidade de feeds de alto-falante com base na informação de renderização de áudio especificada no bitstream.

[0205] As FIGS. 8A-8D são diagramas ilustrando bitstreams 21A-21D formados de acordo com as técnicas descritas nesta revelação. No exemplo da FIG. 8A, o bitstream 21A pode representar um exemplo do bitstream 21 mostrado nas FIGS. 2-4 acima. O bitstream 21A inclui a informação de renderização de áudio 2A que inclui um ou mais bits que definem um valor de sinal 554. O valor de

sinal 554 pode representar qualquer combinação dos tipos de informação descritos abaixo. O bitstream 21A inclui também conteúdo de áudio 558, o qual pode representar um exemplo do conteúdo de áudio 7/9.

[0206] No exemplo da FIG. 8B, o bitstream 21B pode ser semelhante ao bitstream 21A onde o valor de sinal 554 da informação de renderização de áudio 2B compreende um índice 554A, um ou mais bits que definem um tamanho de fileira 554B da matriz sinalizada, um ou mais bits que definem um tamanho de coluna 554C da matriz sinalizada, e coeficientes de matriz 554D. O índice 554A pode ser definido utilizando dois a cinco bits, enquanto que cada tamanho de fileira 554B e tamanho de coluna 554C pode ser definido utilizando dois a dezasseis bits.

[0207] A unidade de extração 72 pode extrair o índice 554A e determinar se o índice indica que a matriz está incluída no bitstream 21B (onde certos valores de índice, como 0000 ou 1111, podem sinalizar que a matriz é explicitamente especificada no bitstream 21B). No exemplo da FIG. 8B, o bitstream 21B inclui um índice 554A sinalizando que a matriz é explicitamente especificada no bitstream 21B. Como resultado, a unidade de extração 72 pode extrair o tamanho da fileira 554B e o tamanho da coluna 554C. A unidade de extração 72 pode ser configurada para calcular o número de bits para analisar que representam coeficientes de matriz em função do tamanho de fileira 554B, do tamanho da coluna 554C e de um tamanho de bit sinalizado (não ilustrado na FIG. 8A) ou implícito de cada coeficiente de matriz. Usando o número determinado de bits, a unidade de extração 72 pode extrair os coeficientes de matriz 554D, que o sistema de reprodução de áudio 16 pode utilizar para configurar um dos renderizadores de áudio 22 como descrito acima. Embora ilustrado como

sinalizando a informação de renderização de áudio 2B uma única vez no bitstream 21B, a informação de renderização de áudio 2B pode ser sinalizada várias vezes no bitstream 21B ou pelo menos parcialmente ou totalmente em um canal separado fora da banda (como dados opcionais em alguns casos).

[0208] No exemplo da FIG. 8C, o bitstream 21C pode representar um exemplo do bitstream 21 mostrado nas FIGS. 2-4 acima. O bitstream 21C inclui a informação de renderização de áudio 2C que inclui um valor de sinal 554, que neste exemplo especifica um índice de algoritmo 554E. O bitstream 21C também inclui o conteúdo de áudio 558. O índice de algoritmo 554E pode ser definido utilizando dois a cinco bits, como referido acima, onde este índice de algoritmo 554E pode identificar um algoritmo de renderização a ser utilizado ao renderizar o conteúdo de áudio 558.

[0209] A unidade de extração 72 pode extrair o índice de algoritmo 550E e determinar se o índice de algoritmo 554E indica que a matriz está incluída no bitstream 21C (onde certos valores de índice, como 0000 ou 1111, podem sinalizar que a matriz é explicitamente especificada no bitstream 21C). No exemplo da FIG. 8C, o bitstream 21C inclui o índice de algoritmo 554E sinalizando que a matriz não é explicitamente especificada no bitstream 21C. Como resultado, a unidade de extração 72 encaminha o índice de algoritmo 554E para o sistema de reprodução de áudio 16, que seleciona um correspondente (se disponível) dos algoritmos de renderização (que são designados como renderizadores 22 no exemplo das FIGs 2-4). Embora ilustrado como sinalizando da informação de renderização de áudio 2C, uma única vez no bitstream 21C, no exemplo da FIG. 8C, a informação de renderização de áudio 2C pode ser

sinalizada várias vezes no bitstream 21B ou pelo menos parcialmente ou totalmente em um canal separado fora da banda (como dados opcionais em alguns casos).

[0210] No exemplo da FIG. 8D, o bitstream 21D pode representar um exemplo do bitstream 21 mostrado nas FIGS. 2-4 acima. O bitstream 21D inclui a informação de renderização de áudio 2D que inclui um valor de sinal 554, que neste exemplo especifica um índice de matriz 554F. O bitstream 21D também inclui o conteúdo de áudio 558. O índice de matriz 554F pode ser definido utilizando dois a cinco bits, como referido acima, onde este índice de matriz 554F pode identificar um algoritmo de renderização a ser utilizado ao renderizar o conteúdo de áudio 558.

[0211] A unidade de extração 72 pode extrair o índice de matriz 550F e determinar se o índice de matriz 554F indica que a matriz está incluída no bitstream 21D (onde certos valores de índice, como 0000 ou 1111, podem sinalizar que a matriz é explicitamente especificada no bitstream 21C). No exemplo da FIG. 8D, o bitstream 21D inclui o índice de matriz 554F sinalizando que a matriz não é explicitamente especificada no bitstream 21D. Como resultado, a unidade de extração 72 encaminha o índice de matriz 554F para um dispositivo de reprodução de áudio, que seleciona um correspondente (se disponível) dos renderizadores 22. Embora ilustrado como sinalização da informação de renderização de áudio 2D, um único momento no bitstream 21D, no exemplo da FIG. 8D, a informação de renderização de áudio 2D pode ser sinalizada várias vezes no bitstream 21D ou pelo menos parcialmente ou totalmente em um canal separado fora da banda (como dados opcionais em alguns casos).

[0212] As FIGs. 8E-8G são diagramas ilustrando porções do bitstream ou informação de canal lateral que

podem especificar os componentes espaciais comprimidos em mais detalhe. A FIG. 8E ilustra um primeiro exemplo de um quadro 249A' do bitstream 21. No exemplo da FIG. 8E, o quadro 249A' inclui campos de ChannelSideInfoData (CSID) 154A-154C, campos HOAGainCorrectionData (HOAGCD) e campos VVectorData 156A e 156B. O campo CSD 154A inclui a unitC 267, bb 266 e ba 265 juntamente com o ChannelType 269, cada um dos quais é ajustado para os valores correspondentes 01, 1, 0 e 01 mostrados no exemplo da FIG. 8E. O campo CSD 154B inclui a unitC 267, bb 266 e ba 265 juntamente com o ChannelType 269, cada um dos quais é ajustado para os valores correspondentes 01, 1, 0 e 01 mostrados no exemplo da FIG. 8E. O campo CSC 154C inclui o campo ChannelType 269 com um valor de 3. Cada um dos campos CSA 154A-154C corresponde ao respectivo dos canais de transporte 1, 2 e 3. Com efeito, cada campo CSID 154A-154C indica se a carga útil correspondente 156A e 156B são sinais com base na direção (quando o ChannelType correspondente é igual a zero), sinais baseados em vetor (quando o ChannelType correspondente é igual a um) um coeficiente HOA ambiente adicional (quando o ChannelType correspondente é igual a dois), ou vazio (quando o ChannelType é igual a três).

[0213] No exemplo da FIG. 8E, o quadro 249A inclui dois sinais baseados em vetor (dado o ChannelType 269 igual a 1 nos campos de CSID 154A e 154B) e um vazio (dado que o ChannelType 269 é igual a 3 no campo CSC 154C). Com base em uma porção de HOAconfig anterior (não mostrada para facilidade de ilustração), o dispositivo de decodificação de áudio 24 pode determinar que todos os 16 elementos de vetor V são codificados. Assim, os VVectorData 156A e 156B incluem, cada um, todos os 16 elementos de vetor, cada um deles uniformemente quanzado com 8 bits.

[0214] Como mostrado adicionalmente no exemplo da FIG. 8E, a estrutura 249A' não inclui um campo HOAPredictionInfo. O campo HOAPredictionInfo pode representar um campo correspondente a um segundo esquema de compressão com base direcional que pode ser removido de acordo com a técnica descrita nesta revelação quando o esquema de compressão baseado em vetor é utilizado para comprimir dados de áudio HOA.

[0215] A FIG. 8F é um diagrama que ilustra um quadro 249A" que é substancialmente semelhante ao quadro 249A exceto pelo fato de que o HOAGainCorrectionData foi removido de cada canal de transporte armazenado no quadro 249A". O campo HOAGainCorrectionData pode ser removido do quadro 249A" quando a correção de ganho é suprimida de acordo com vários aspectos das técnicas descritas acima.

[0216] A FIG. 8G é um diagrama que ilustra um quadro 249A'" que pode ser semelhante ao quadro 249A", exceto pelo fato de que o campo HOAPredictionInfo é removido. O quadro 249A'" representa um exemplo onde ambos os aspectos das técnicas podem ser aplicados conjuntamente para remover vários campos que podem não ser necessários em certas circunstâncias.

[0217] As técnicas anteriores podem ser realizadas com relação a qualquer número de contextos diferentes e ecossistemas de áudio. Uma série de contextos exemplares são descritos abaixo, embora as técnicas devam ser limitadas aos contextos de exemplo. Um exemplo de ecossistema de áudio pode incluir conteúdo de áudio, estúdios de cinema, estúdios de música, estúdios de áudio de jogos, conteúdo de áudio baseado em canais, mecanismos de codificação, hastes de áudio de jogo, mecanismos de codificação/renderização de áudio de jogo e sistemas de entrega.

[0218] Os estúdios de cinema, os estúdios de música e os 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 fornecer conteúdo de áudio baseado em canal (por exemplo, em 2.0, 5.1 e 7.1), como usando uma estação de trabalho de áudio digital (DAW). Os estúdios de música podem fornecer conteúdo de áudio baseado em canal (por exemplo, em 2.0 e 5.1), como usando uma DAW. Em qualquer dos casos, os mecanismos de codificação podem receber e codificar o conteúdo de áudio baseado em canal baseado em um ou mais codecs (por exemplo, AAC, AC3, Dolby True HD, Dolby Digital Plus e DTS Master Audio) para saída pelos sistemas de distribuição. Os estúdios de áudio de jogos podem produzir uma ou mais hastes de áudio do jogo, como pelo uso de uma DAW. Os mecanismos de codificação/renderização de áudio do jogo podem codificar e/ou renderizar as hastes de áudio para o conteúdo de áudio baseado em canal para saída pelos sistemas de entrega. Outro contexto de exemplo no qual as técnicas podem ser realizadas compreende um ecossistema de áudio que pode incluir a transmissão de objetos de áudio de gravação, sistemas de áudio profissionais, captura no dispositivo do consumidor, formato de áudio HOA, renderização no dispositivo, áudio de consumidor, TV e acessórios e sistemas de áudio para automóveis.

[0219] A transmissão de objetos de áudio de gravação, os sistemas de áudio profissionais e a captura no dispositivo do consumidor podem codificar sua saída usando o 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 usando a renderização no dispositivo, o áudio do consumidor, TV e

acessórios e os sistemas de áudio do carro. Em outras palavras, a representação única do conteúdo de áudio pode ser reproduzida em um sistema genérico de reprodução de áudio (isto é, em oposição a uma configuração específica como 5.1, 7.1, etc.), como o sistema de reprodução de áudio 16.

[0220] 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 fio e/ou sem fio (por exemplo, microfones Eigen), captura de som ambiente no 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 canais de comunicação com fios e/ou sem fios.

[0221] De acordo com uma ou mais técnicas desta revelação, o dispositivo móvel pode ser utilizado 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 ambiente no dispositivo (por exemplo, uma pluralidade de microfones integrados no dispositivo móvel). O dispositivo móvel pode então codificar o campo de som adquirido nos coeficientes 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 HOA.

[0222] O dispositivo móvel também pode utilizar um ou mais dos elementos de reprodução para

reproduzir o campo de som codificado HOA. Por exemplo, o dispositivo móvel pode decodificar o campo de som codificado HOA e emitir um sinal para um ou mais dos elementos de reprodução que fazem com que um ou mais dos elementos de reprodução recriem o campo de som. Como um exemplo, o dispositivo móvel pode utilizar os canais de comunicação sem fios e/ou com fios para enviar o sinal para um ou mais alto-falantes (por exemplo, conjuntos de alto-falantes, 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 ancoragem e/ou um ou mais alto-falantes encaixados (por exemplo, sistemas de som em carros e/ou casas inteligentes). Como outro exemplo, o dispositivo móvel pode utilizar processamento de fones de ouvido para emitir o sinal para um conjunto de fones de ouvido, por exemplo, para criar um som binaural realista.

[0223] Em alguns exemplos, um determinado dispositivo móvel pode adquirir um campo de som 3D e reproduzir o mesmo campo de som 3D posteriormente. 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.

[0224] Ainda outro contexto no qual 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, um ou mais DAWs podem incluir plugins HOA e/ou

ferramentas que podem ser configuradas para funcionar com (por exemplo, trabalhar com) um ou mais sistemas de áudio de jogos. Em alguns exemplos, os estúdios de jogos podem produzir novos formatos de haste que suportam HOA. Em qualquer caso, os estúdios de jogos podem produzir conteúdo de áudio codificado para os mecanismos de renderização que podem renderizar um campo de som para reprodução pelos sistemas de distribuição.

[0225] As técnicas também podem ser realizadas com relação aos dispositivos de aquisição de áudio exemplares. Por exemplo, as técnicas podem ser realizadas em relação a um microfone Eigen que pode incluir uma pluralidade de microfones que são coletivamente configurados para gravar um campo de som 3D. Em alguns exemplos, a pluralidade de microfones de microfone Eigen pode estar localizada na superfície de uma bola substancialmente esférica com um raio de aproximadamente 4 cm. Em alguns exemplos, o dispositivo de codificação de áudio 20 pode ser integrado no microfone de Eigen de modo a produzir um bitstream 21 diretamente a partir do microfone.

[0226] Outro contexto exemplificativo de aquisição de áudio pode incluir um caminhão de produção que pode ser configurado para receber um sinal de um ou mais microfones, tais como um ou mais microfones Eigen. O caminhão de produção também pode incluir um codificador de áudio, como o codificador de áudio 20 da FIG. 3.

[0227] O dispositivo móvel também pode, em alguns casos, incluir uma pluralidade de microfones que são coletivamente configurados para gravar um campo de som 3D. Por outras palavras, a pluralidade de microfones pode ter diversidade X, Y, Z. Em alguns exemplos, o dispositivo móvel pode incluir um microfone que pode ser rodado para fornecer diversidade X, Y, Z em relação a um ou mais

microfones do dispositivo móvel. O dispositivo móvel também pode incluir um codificador de áudio, como o codificador de áudio 20 da FIG. 3.

[0228] Um dispositivo de captura de vídeo reforçado pode ainda ser configurado para gravar um campo de som 3D. Em alguns exemplos, o dispositivo de captura de vídeo reforçado pode ser ligado a um capacete de um usuário envolvido em uma atividade. Por exemplo, o dispositivo de captura de vídeo reforçado pode ser ligado a um capacete de um usuário praticando rafting em águas claras. Desta forma, o dispositivo de captura de vídeo reforçado 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 na frente do usuário, etc ...).

[0229] As técnicas também podem ser realizadas com relação a um dispositivo móvel acessório melhorado, que pode ser configurado para gravar 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 mencionado para formar um dispositivo móvel acessório melhorado. Desta forma, o dispositivo móvel acessório melhorado pode capturar uma versão de qualidade superior do campo de som 3D do que apenas usando componentes de captura de som integrado ao dispositivo móvel acessório melhorado.

[0230] Exemplo de dispositivos de reprodução de áudio que podem executar vários aspectos das técnicas descritas nesta revelação são discutidos adicionalmente abaixo. De acordo com uma ou mais técnicas desta descrição, os alto-falantes e/ou as barras de som podem ser dispostos em qualquer configuração arbitrária enquanto

ainda reproduzem 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 de uma conexão com fios ou sem fios. De acordo com uma ou mais técnicas desta descrição, uma única representação genérica de um campo de som pode ser utilizada para renderizar o campo de som em qualquer combinação de alto-falantes, barras de som e dispositivos de reprodução de fones de ouvido.

[0231] Um certo número de diferentes ambientes de reprodução de áudio exemplares também podem ser adequados para executar vários aspectos das técnicas descritas nesta revelação. Por exemplo, um ambiente de reprodução de alto-falante 5.1, um ambiente de reprodução de alto-falante 2.0 (por exemplo, estéreo), um ambiente de reprodução de alto-falantes 9.1 com alto-falantes dianteiros de altura total, um ambiente de reprodução de alto-falantes 22.2, um dispositivo móvel com ambiente de reprodução de botão auricular pode ser ambiente adequado para executar vários aspectos das técnicas descritas nesta descrição.

[0232] De acordo com uma ou mais técnicas desta descrição, uma única representação genérica de um campo de som pode ser utilizada para renderizar o campo de som em qualquer dos ambientes de reprodução anteriormente mencionados. Adicionalmente, as técnicas desta revelação permitem a um renderizado renderizar um campo de som a partir de uma representação genérica para reprodução em ambientes de reprodução diferentes dos descritos acima. Por exemplo, se as considerações de projeto proibirem a colocação correta de alto-falantes de acordo com um ambiente de reprodução de alto-falante 7.1 (por exemplo, se não for possível colocar um alto-falante ambiente direito),

as técnicas desta revelação permitem que um renderizador compense com os outros 6 alto-falantes, de modo a reprodução pode ser alcançada em um ambiente de reprodução de alto-falante 6.1.

[0233] Além disso, um usuário pode assistir a um jogo esportivo enquanto usa fones de ouvido. De acordo com uma ou mais técnicas desta revelação, o campo de som 3D do jogo esportivo pode ser adquirido (por exemplo, um ou mais microfones Eigen podem ser colocados dentro e/ou ao redor do estádio de basebol), os coeficientes HOA correspondentes ao campo de som 3D podem ser obtidos e transmitidos a um decodificador, o decodificador pode reconstruir o campo de som 3D com base nos coeficientes HOA e emitir o campo de som 3D reconstruído para um renderizador, o renderizador pode obter uma indicação do tipo de ambiente de reprodução (por exemplo, fones de ouvido) e renderizar o campo de som 3D reconstruído em sinais que fazem com que os fones de ouvido produzam uma representação do campo de som 3D do jogo esportivo.

[0234] Em cada uma das várias instâncias descritas acima, deve ser entendido que o dispositivo de codificação de áudio 20 pode executar um método ou de outro modo compreender meios para executar 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 finalidade especial configurado por meio de instruções armazenadas em uma mídia de armazenamento legível por computador não transitória. Em outras palavras, vários aspectos das técnicas em cada um dos conjuntos de exemplos de codificação podem fornecer uma mídia de armazenamento legível por computador não transitória tendo armazenadas

nela instruções que, quando executadas, fazem com que um ou mais processadores executem o método para o qual o dispositivo de codificação de áudio 20 foi configurado para executar.

[0235] 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 em ou transmitidas através de uma ou mais instruções ou código em uma mídia legível por computador e executadas por uma unidade de processamento baseada em hardware. A mídia legível por computador pode incluir mídia de armazenamento legível por computador que corresponde a uma mídia tangível como mídia de armazenamento de dados. A mídia de armazenamento de dados pode ser qualquer mídia disponível que pode ser acessada 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 na presente revelação. Um produto de programa de computador pode incluir uma mídia legível por computador.

[0236] De modo similar, cada uma das várias instâncias descritas acima, deve ser entendido que o dispositivo de decodificação de áudio 24 pode executar um método ou de outro modo compreender meios para executar 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 em uma mídia de armazenamento legível por computador não transitória. Em outras palavras, vários aspectos das técnicas em cada um dos conjuntos de exemplos

de codificação podem fornecer uma mídia de armazenamento legível por computador não transitória tendo armazenadas nela instruções que, quando executadas, fazem com que um ou mais processadores executem o método para o qual o dispositivo de decodificação de áudio 24 foi configurado para executar.

[0237] A título de exemplo, e não como limitação, tais mídias de armazenamento legíveis por computador podem compreender RAM, ROM, EEPROM, CD-ROM ou outro armazenamento em disco ótico, armazenamento em disco magnético ou outros dispositivos de armazenamento magnéticos, memória flash ou qualquer outro meio que possa ser utilizado para armazenar código de programa desejado sob a forma de instruções ou estruturas de dados e que pode ser acessado por um computador. Deve-se compreender, no entanto, que a mídia de armazenamento legível por computador e mídia de armazenamento de dados não incluem conexões, ondas transportadoras, sinais ou outra mídia transitória, mas são, ao invés, direcionadas à mídia de armazenamento tangível, não transitória. Disco e disquete, como aqui utilizados, incluem disco compacto (CD), disco a laser, disco ótico, disco versátil digital (DVD), disquete e disco Blu-ray onde os disquetes geralmente reproduzem dados magneticamente, enquanto que os discos reproduzem dados oticamente com lasers. Combinações dos anteriores também devem ser incluídas dentro do escopo de mídias legíveis por computador.

[0238] As instruções podem ser executadas por um ou mais processadores, como um ou mais processadores de sinal digital (DSPs), microprocessadores de uso geral, circuitos integrados de aplicação específica (ASIC), arranjos de lógica programáveis em campo (FPGA), ou outros circuitos lógicos ou discretos equivalentes.

Consequentemente, o termo "processador" como aqui utilizado pode referir-se a qualquer uma das estruturas precedentes ou qualquer outra estrutura adequada para implementação 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 a codificação e decodificação, ou incorporados em um codec combinado. Também, as técnicas podem ser totalmente implementadas em um ou mais circuitos ou elementos lógicos.

[0239] As técnicas da presente revelação podem ser implementadas em uma vasta variedade de dispositivos ou aparelhos, incluindo um monofone sem fios, um circuito integrado (IC) ou um conjunto de ICs (por exemplo, um conjunto de chips). Vários componentes, módulos ou unidades são descritos na presente revelação para enfatizar os aspectos funcionais dos dispositivos configurados para executar as técnicas divulgadas, mas não precisam necessariamente da realização por diferentes unidades de hardware. Em vez disso, como descrito acima, várias unidades podem ser combinadas em uma unidade de hardware de codec ou fornecida por uma coleta de unidades de hardware interoperativas, incluindo um ou mais processadores, conforme descrito acima, em conjunto com o software e/ou firmware adequado.

[0240] Vários aspectos das técnicas foram descritos. Esses e outros aspectos das técnicas estão dentro do escopo das reivindicações a seguir.

REIVINDICAÇÕES

1. Dispositivo configurado para renderizar coeficientes ambissônicos de ordem superior, **caracterizado** pelo fato de que compreende:

um ou mais processadores configurados para:

obter, a partir de um fluxo de bits que inclui uma versão codificada dos coeficientes ambissônicos de ordem superior, informações de esparsidade que indicam esparsidade de uma matriz usada para renderizar os coeficientes ambissônicos de ordem superior para gerar uma pluralidade de alimentações de alto-falantes; e

obter, a partir do fluxo de bits, informações de simetria de sinal que indicam a simetria de sinal da matriz;

obter, a partir do fluxo de bits, um número reduzido de bits usado para representar a matriz; e

com base nas informações de esparsidade, nas informações de simetria de sinal e no número reduzido de bits, reconstruir a matriz; e

uma memória acoplada ao um ou mais processadores e configurada para armazenar as informações de esparsidade.

2. Dispositivo, de acordo com a reivindicação 1, **caracterizado** pelo fato de que o um ou mais processadores são adicionalmente configurados para determinar um layout de alto-falante para o qual a matriz deve ser usada para renderizar a pluralidade de alimentações de alto-falante a partir dos coeficientes ambissônicos de ordem superior.

3. Dispositivo, de acordo com a reivindicação 1, **caracterizado** pelo fato de que compreende adicionalmente um alto-falante configurado para reproduzir um campo de som representado pelos coeficientes ambissônicos de ordem superior com base na pluralidade de alimentações de alto-falante.

4. Dispositivo, de acordo com a reivindicação 1, **caracterizado** pelo fato de que o um ou mais processadores são adicionalmente configurados para obter, a partir do fluxo de bits, informações de renderização de áudio indicativas de um valor de sinal que identifica um renderizador de áudio usado ao gerar a pluralidade de alimentações de alto-falante, e renderizar a pluralidade de alimentações de alto-falante com base nas informações de renderização de áudio.

5. Dispositivo, de acordo com a reivindicação 4, **caracterizado** pelo fato de que o valor de sinal inclui um índice associado à matriz usada para renderizar os coeficientes ambissônicos de ordem superior para a pluralidade de alimentações de alto-falante, e

em que o um ou mais processadores são configurados para renderizar a pluralidade de alimentações de alto-falante com base na matriz associada com o índice incluído no valor de sinal.

6. Método para renderizar coeficientes ambissônicos de ordem superior, **caracterizado** pelo fato de que compreende:

obter, a partir de um fluxo de bits que inclui uma versão codificada dos coeficientes ambissônicos de ordem superior, informações de esparsidão que indicam esparsidão de uma matriz usada para renderizar os coeficientes ambissônicos de ordem superior para gerar uma pluralidade de alimentações de alto-falantes; e

obter, a partir do fluxo de bits, informações de simetria de sinal que indicam a simetria de sinal da matriz;

obter, a partir do fluxo de bits, um número reduzido de bits usado para representar a matriz; e

com base nas informações de esparsidão, nas informações de simetria de sinal e no número reduzido de bits, reconstruir a matriz.

7. Método, de acordo com a reivindicação 6, **caracterizado** pelo fato de que compreende adicionalmente determinar um layout de alto-falante para o qual a matriz deve ser usada para renderizar os dados de áudio de múltiplos canais a partir dos coeficientes ambissônicos de ordem superior.

8. Método, de acordo com a reivindicação 6, **caracterizado** pelo fato de que compreende adicionalmente reproduzir um campo de som representado pelos coeficientes ambissônicos de ordem superior com base na pluralidade de alimentações de alto-falante.

9. Método, de acordo com a reivindicação 6, **caracterizado** pelo fato de que compreende adicionalmente:

obter, do fluxo de bits, informações de renderização de áudio indicativas de um valor de sinal que identifica um renderizador de áudio usado ao gerar a pluralidade de alimentações de alto-falante; e

renderizar a pluralidade de alimentações de alto-falante com base nas informações de renderização de áudio.

10. Método, de acordo com a reivindicação 9, **caracterizado** pelo fato de que o valor de sinal inclui um índice associado à matriz usada para renderizar os coeficientes ambissônicos de ordem superior para gerar a pluralidade de alimentações de alto-falante, e

em que o método compreende adicionalmente renderizar a pluralidade de alimentações de alto-falante com base na matriz associada ao índice incluída no valor de sinal.

11. Dispositivo configurado para produzir um fluxo de bits, **caracterizado** pelo fato de que compreende:

uma memória configurada para armazenar uma matriz usada para renderizar os coeficientes ambissônicos de ordem

superior para gerar uma pluralidade de alimentações de alto-falante; e

um ou mais processadores acoplados à memória, e configurados para:

obter informações de simetria de sinal indicativas de simetria de sinal da matriz;

obter informações de esparsidade que indicam uma esparsidade da matriz; e

com base nas informações de simetria de sinal e nas informações de esparsidade, determinar um número reduzido de bits usado para representar a matriz;

gerar o fluxo de bits para incluir uma versão codificada dos coeficientes ambissônicos de ordem superior, das informações de simetria de sinal, das informações de esparsidade e do número reduzido de bits.

12. Dispositivo, de acordo com a reivindicação 11, **caracterizado** pelo fato de que o um ou mais processadores são adicionalmente configurados para determinar um layout de alto-falante para o qual a matriz deve ser usada para renderizar a pluralidade de alimentações de alto-falante a partir dos coeficientes ambissônicos de ordem superior.

13. Dispositivo, de acordo com a reivindicação 11, **caracterizado** pelo fato de que comprehende adicionalmente um microfone configurado para capturar um campo de som representado pelos coeficientes ambissônicos de ordem superior.

14. Método para produzir um fluxo de bits, **caracterizado** pelo fato de que comprehende:

obter informações de esparsidade indicativas de uma esparsidade de uma matriz utilizada para renderizar os coeficientes ambissônicos de ordem superior para gerar uma pluralidade de alimentações de alto-falantes;

obter informações de simetria de sinal que indicam uma simetria de sinal da matriz; e

com base nas informações de simetria de sinal e nas informações de esparsidade, determinar um número reduzido de bits usados para representar a matriz;

gerar o fluxo de bits para incluir uma versão codificada dos coeficientes ambissônicos de ordem superior, das informações de simetria de sinal, das informações de esparsidade e do número reduzido de bits.

15. Método, de acordo com a reivindicação 14, **caracterizado** pelo fato de que compreende adicionalmente determinar um layout de alto-falante para o qual a matriz deve ser usada para renderizar os dados de áudio de múltiplos canais a partir dos coeficientes ambissônicos de ordem superior.

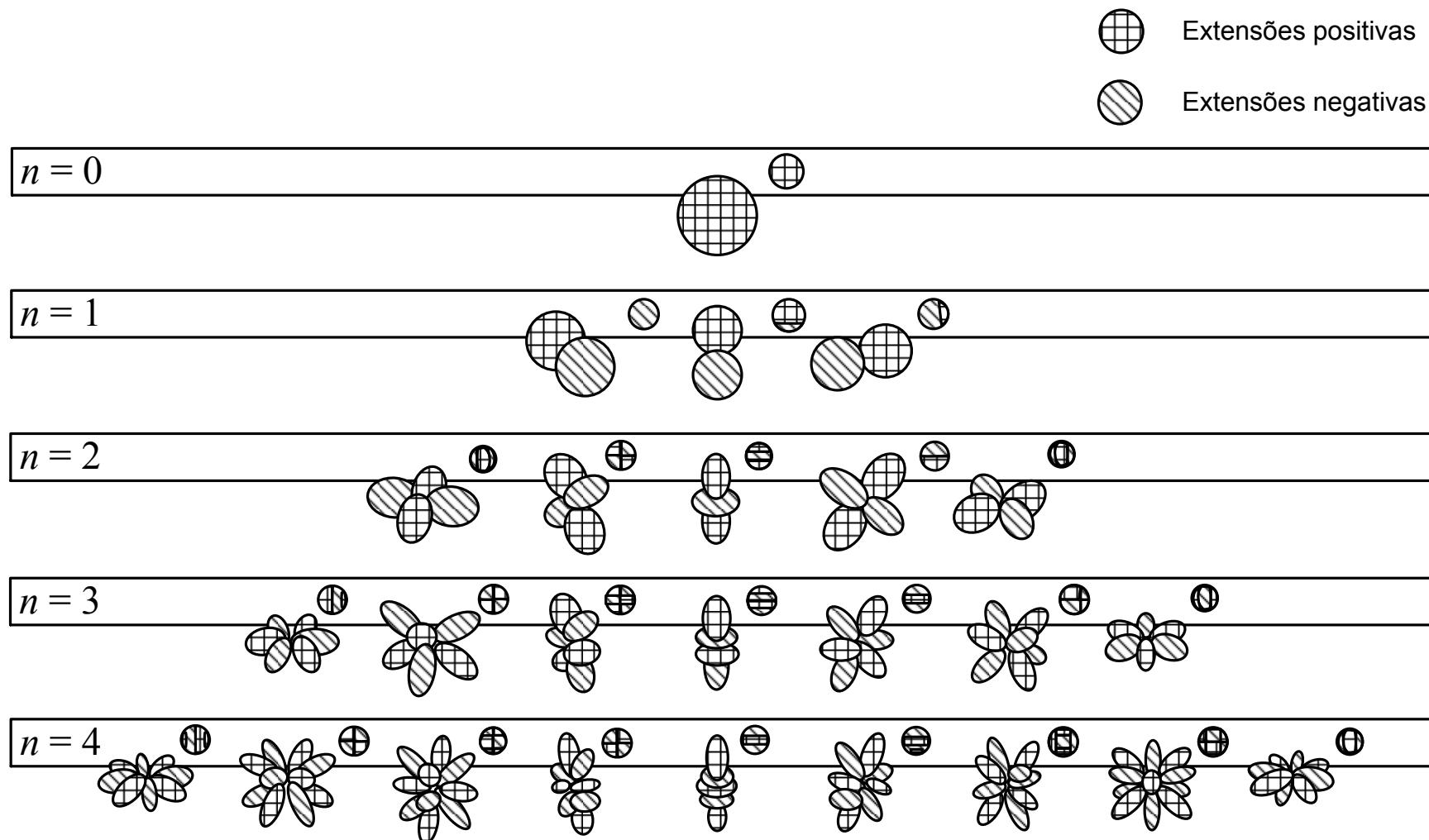
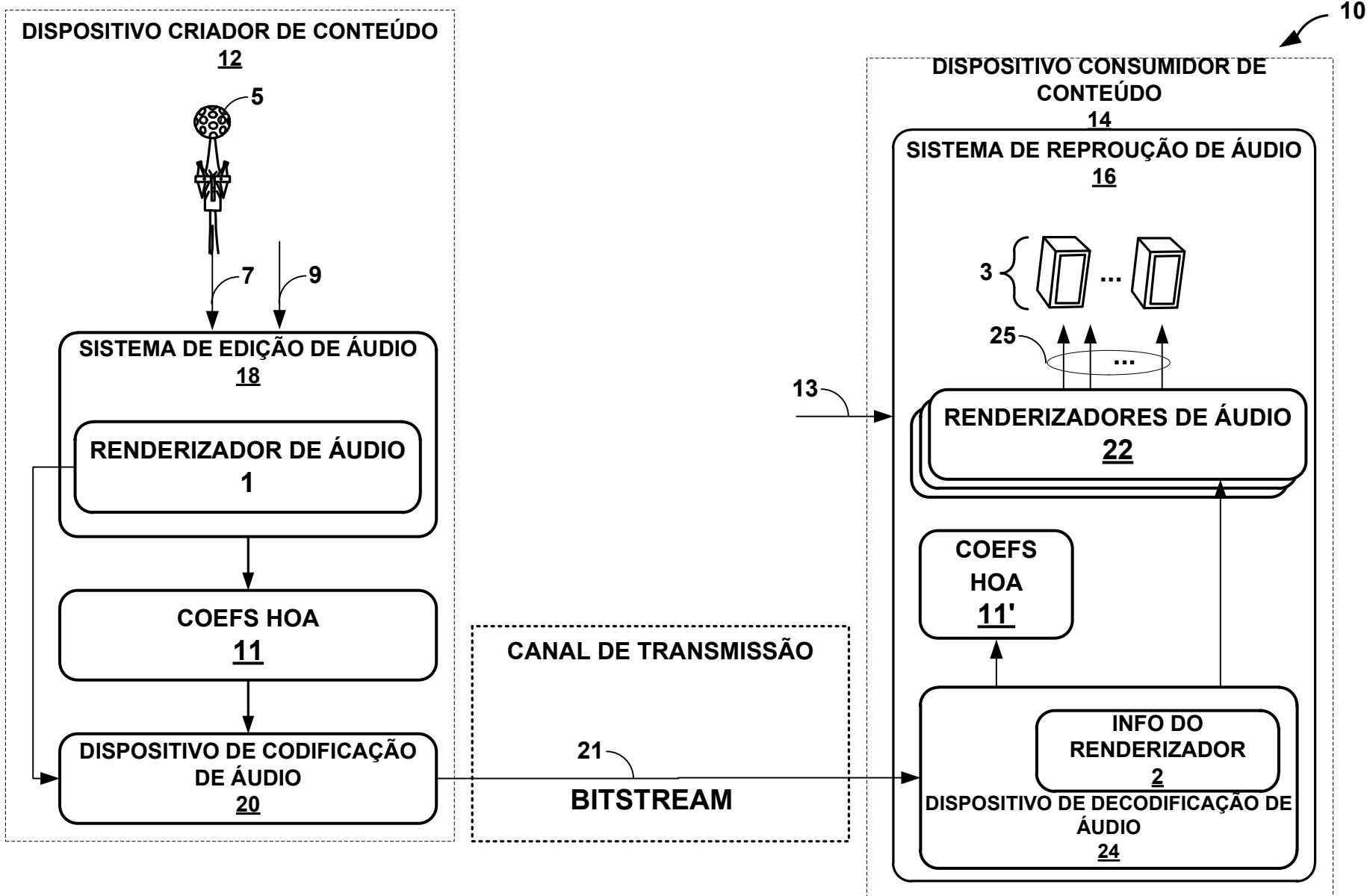
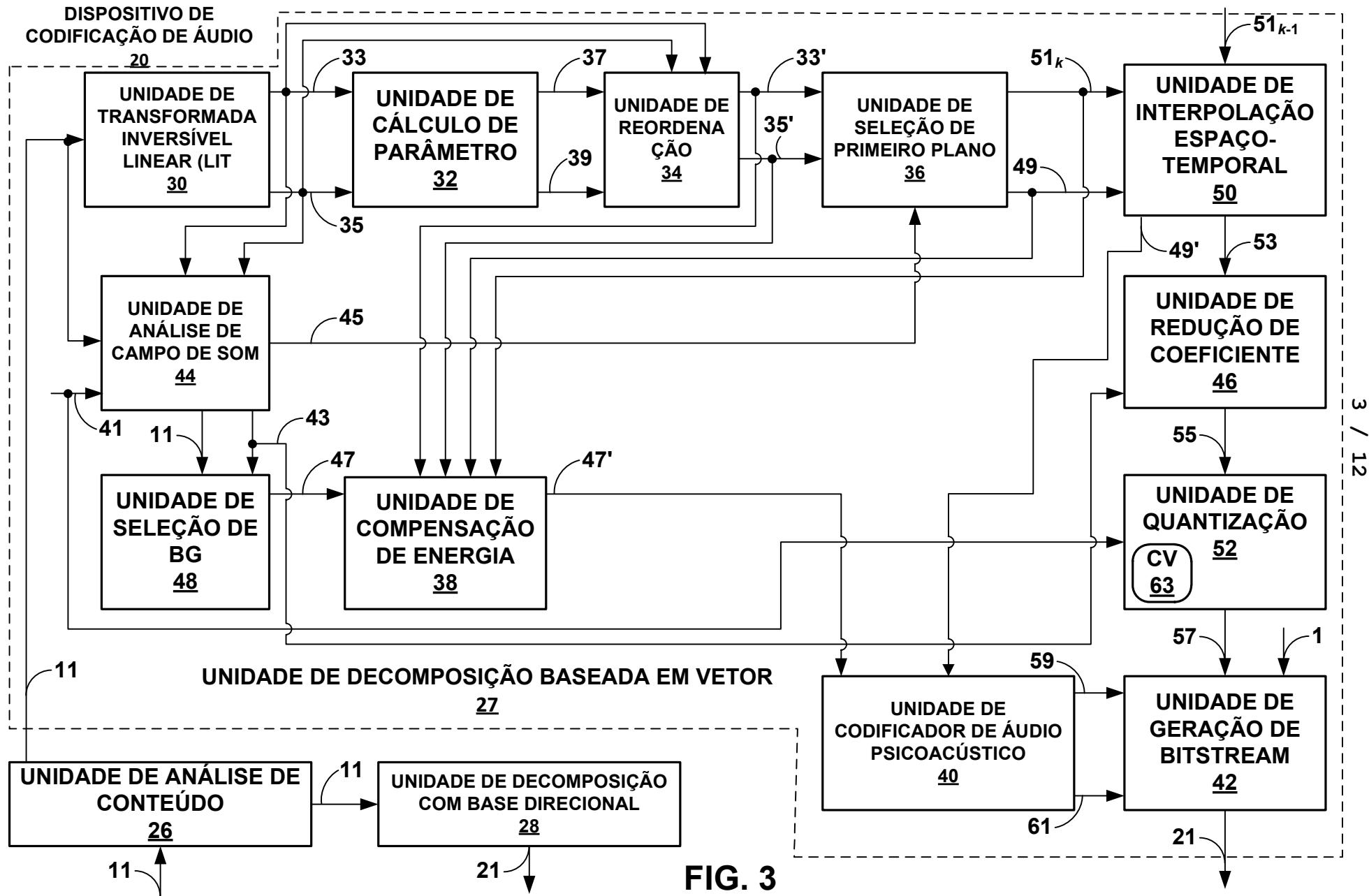


FIG. 1

**FIG. 2**



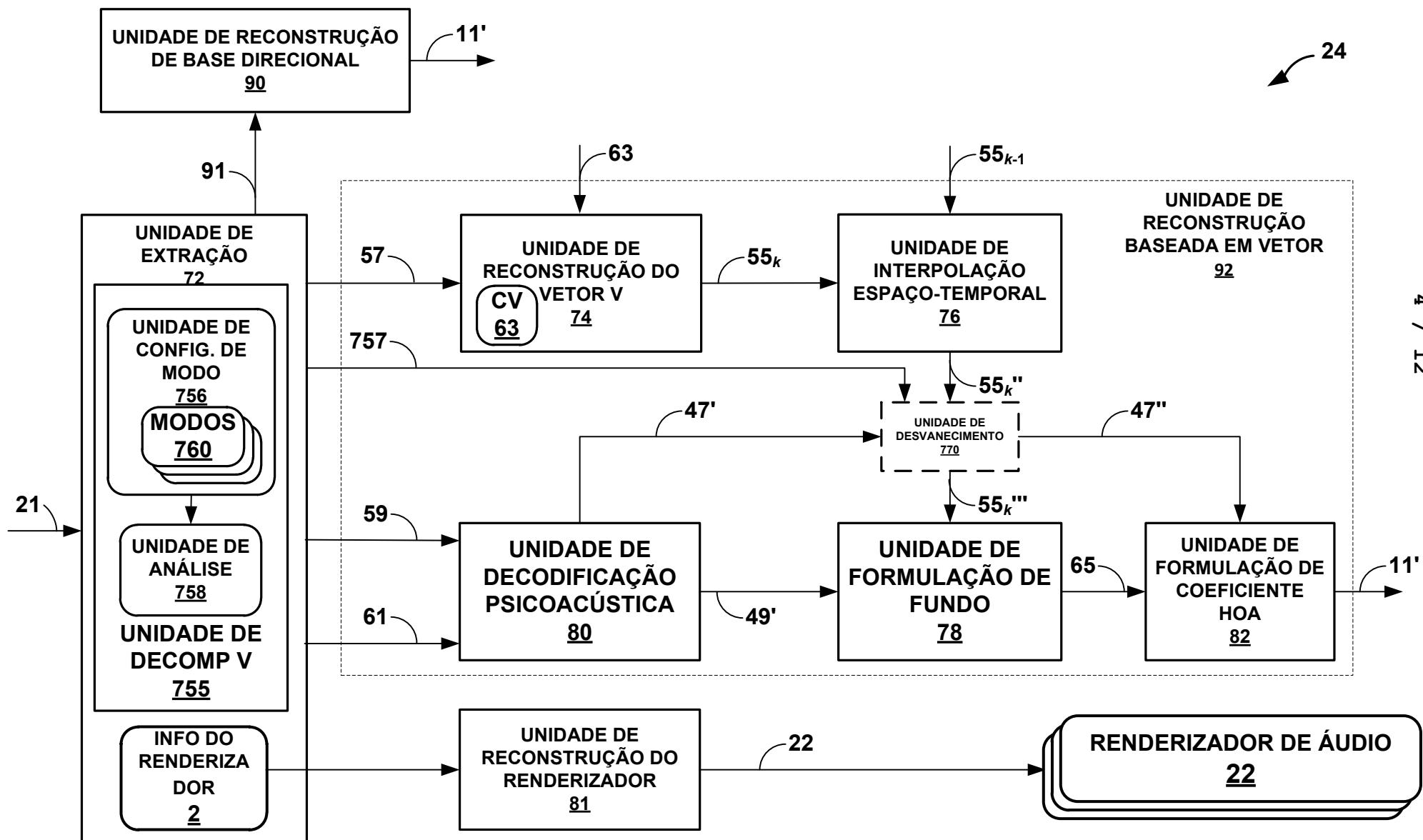
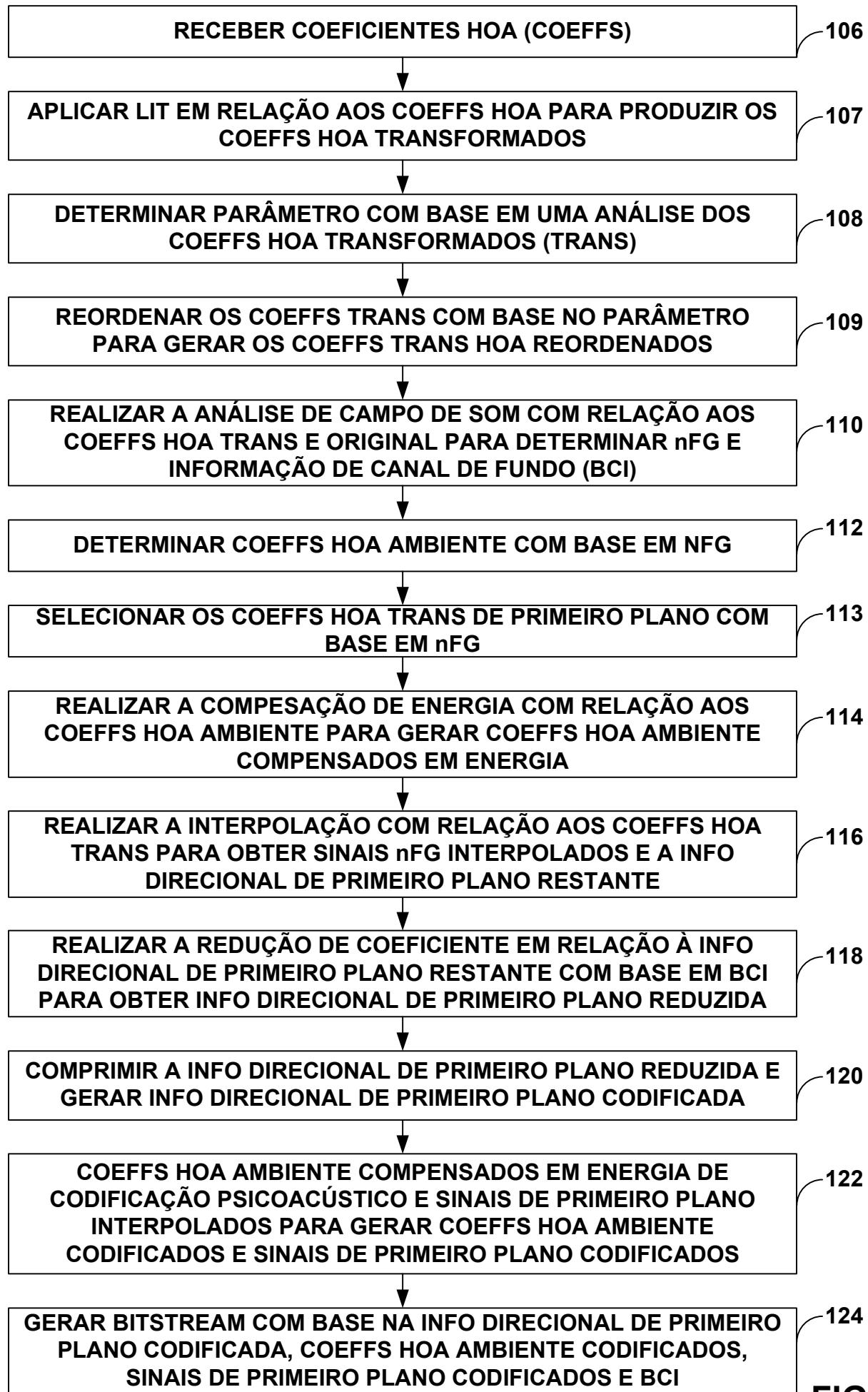
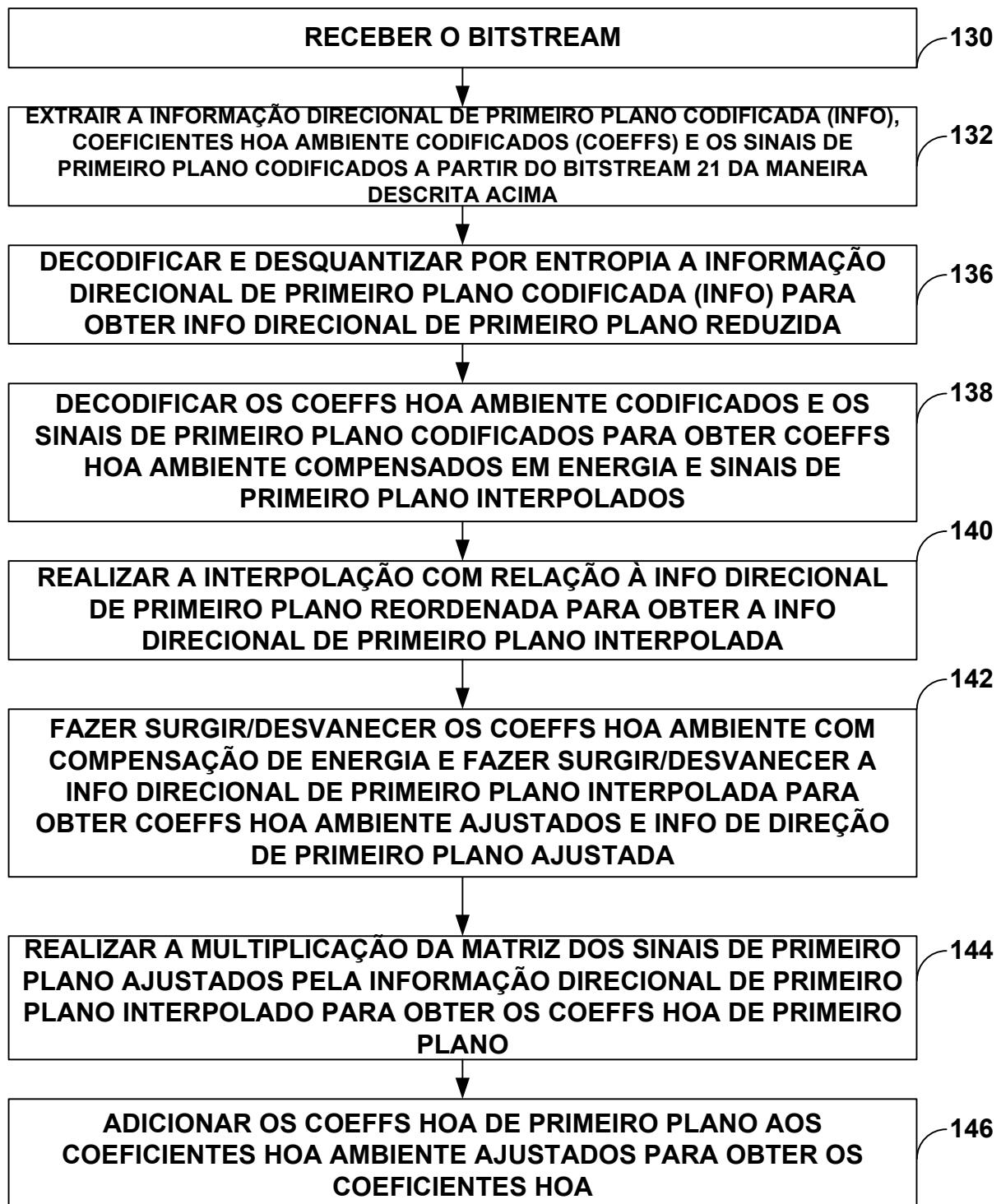


FIG. 4

**FIG. 5**

**FIG. 6**

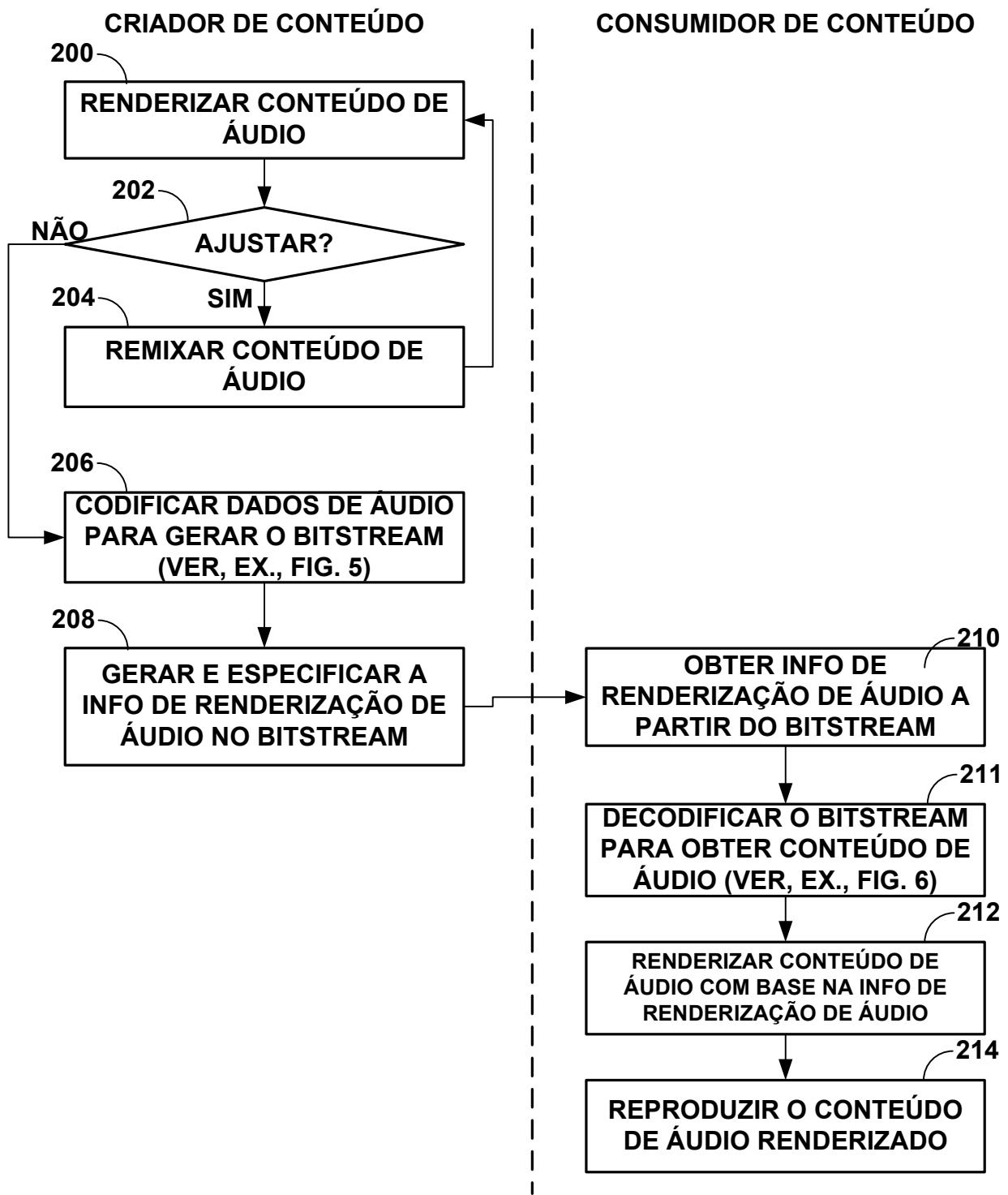


FIG. 7

INFORMAÇÃO DE RENDERIZAÇÃO DE
ÁUDIO

2A

VALOR DO SINAL <u>554</u>	CONTEÚDO DE ÁUDIO <u>558</u>
---------------------------------	------------------------------------

21A

FIG. 8A

INFORMAÇÃO DE RENDERIZAÇÃO DE ÁUDIO

2B

ÍNDICE <u>554A</u>	TAMANHO DE FILEIRA <u>554B</u>	TAMANHO DE COLUNA <u>554C</u>	COEFICIENTES DE MATRIZ <u>554D</u>	CONTEÚDO DE ÁUDIO <u>558</u>
-----------------------	--------------------------------------	-------------------------------------	---------------------------------------	------------------------------------

21B

FIG. 8B

INFORMAÇÃO DE RENDERIZAÇÃO DE

ÁUDIO

2C

ÍNDICE DE ALGORITMO <u>554E</u>	CONTEÚDO DE ÁUDIO <u>558</u>
---------------------------------------	------------------------------------

21C

INFORMAÇÃO DE RENDERIZAÇÃO DE ÁUDIO

2D

ÍNDICE DE MATRIZ <u>554F</u>	CONTEÚDO DE ÁUDIO <u>558</u>
------------------------------------	------------------------------------

21D

FIG. 8C

FIG. 8D

249A'

Quadro k

Neste exemplo os canais de Transporte flexíveis contêm dois canais de Primeiro Plano baseados em vetor no quadro k e três Canais de Primeiro Plano baseados em vetor no segundo quadro k+1

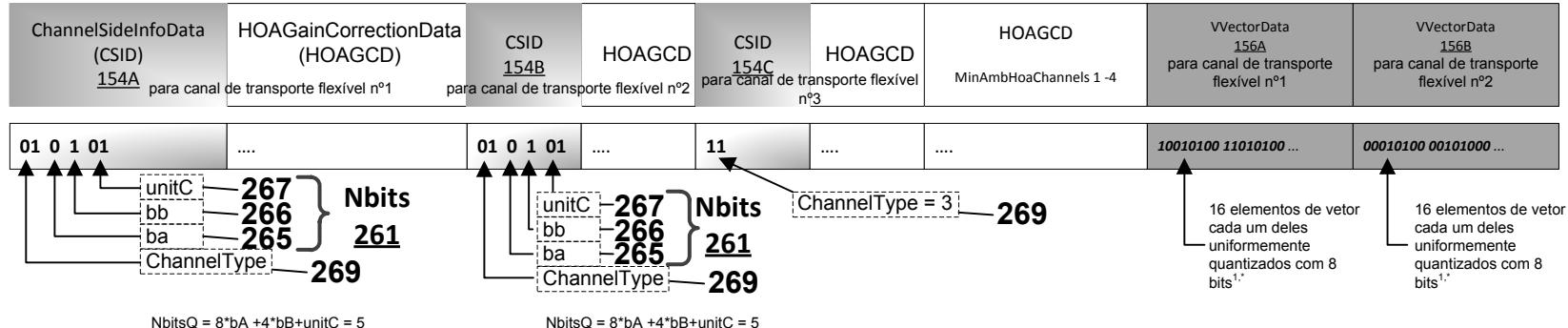


FIG. 8E

9 / 12

249A''

Quadro k

Neste exemplo os canais de Transporte flexíveis contêm dois canais de Primeiro Plano baseados em vetor no quadro k e três Canais de Primeiro Plano baseados em vetor no segundo quadro k+1

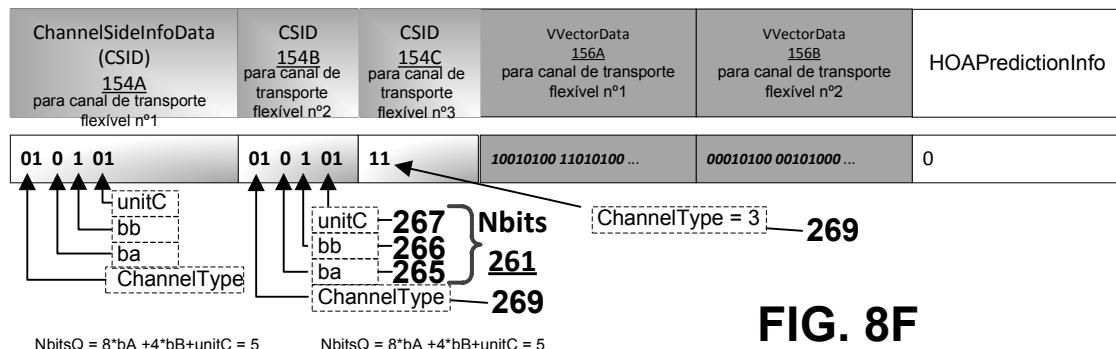


FIG. 8F

249A'''

Quadro k

Neste exemplo os canais de Transporte flexíveis contêm dois canais de Primeiro Plano baseados em vetor no quadro k e três Canais de Primeiro Plano baseados em vetor no segundo quadro k+1

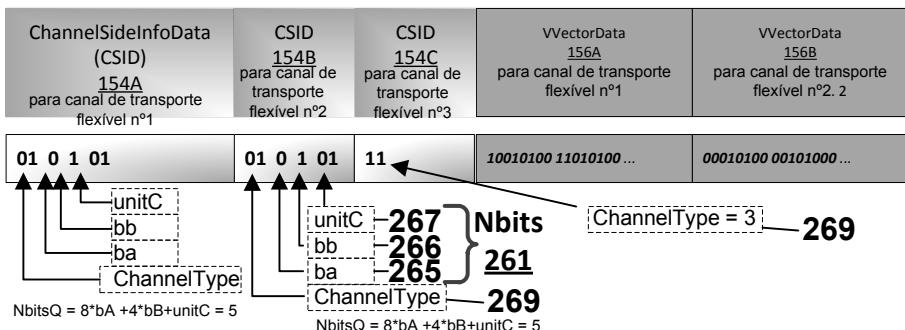
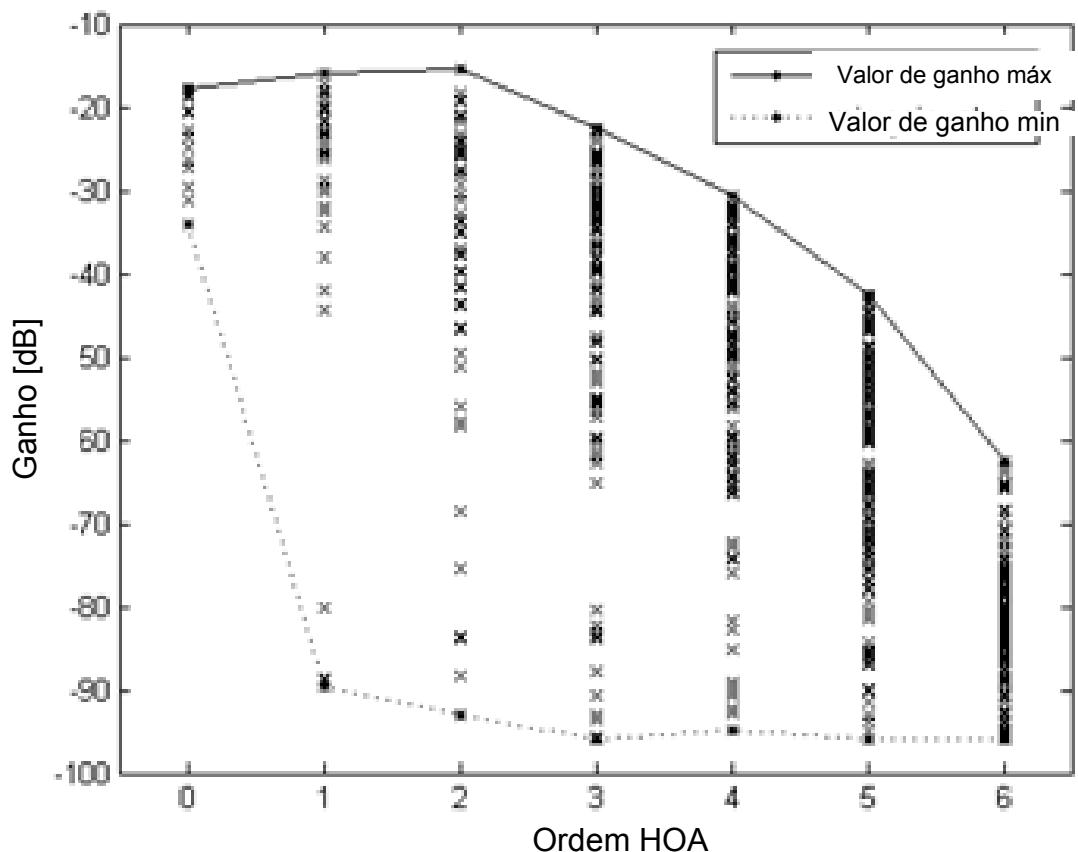
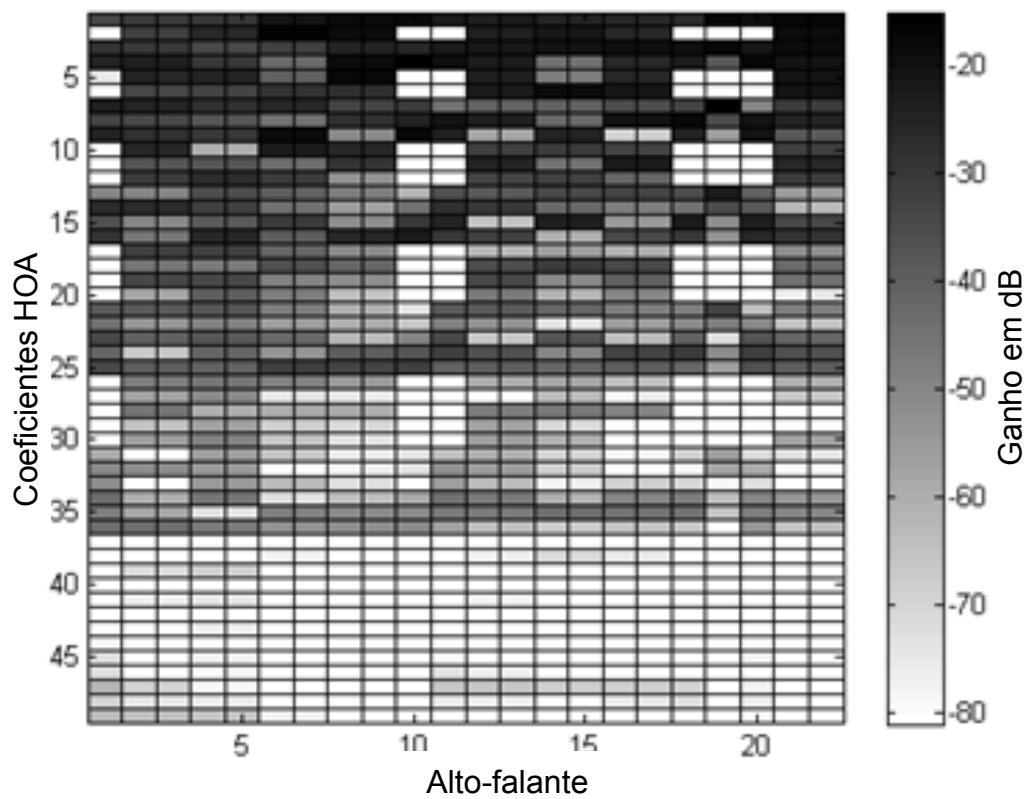


FIG. 8G

**FIG. 9****FIG. 10**

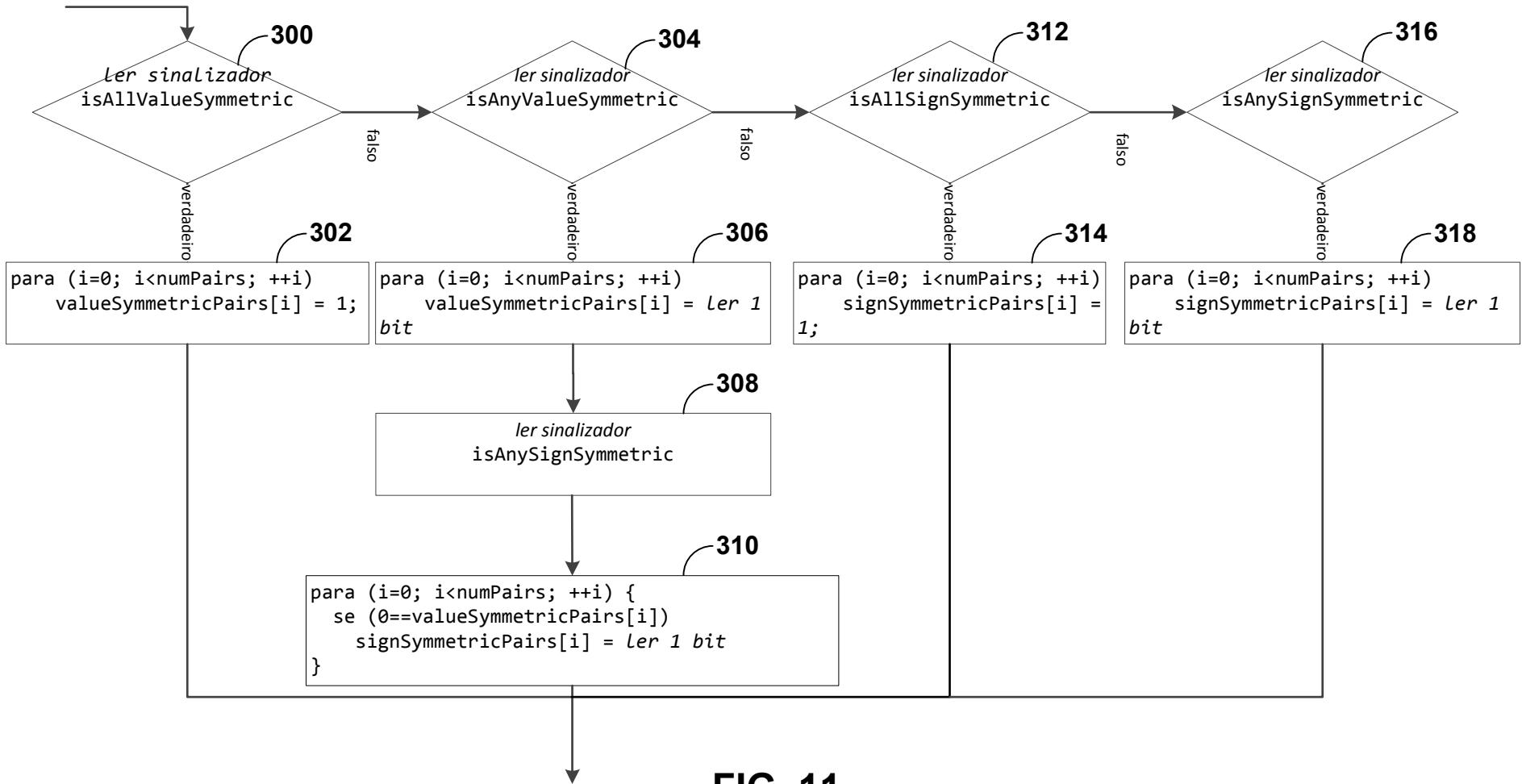


FIG. 11