



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

**(11) BR 112016001248-8 B1**



**(22) Data do Depósito:** 17/07/2014

**(45) Data de Concessão:** 16/11/2022

**(54) Título:** DECODIFICADOR DE ÁUDIO DE MÚLTIPLOS CANAIS E MÉTODO PARA FORNECER PELO MENOS DOIS SINAIS DE ÁUDIO DE SAÍDA COM BASE EM UMA REPRESENTAÇÃO CODIFICADA

**(51) Int.Cl.:** G10L 19/008; G10L 19/20.

**(30) Prioridade Unionista:** 18/10/2013 EP 13189309.1; 22/07/2013 EP 13177375.6.

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

**(72) Inventor(es):** SASCHA DICK; CHRISTIAN HELMRICH; JOHANNES HILPERT; ANDREAS HÖLZER.

**(86) Pedido PCT:** PCT EP2014065416 de 17/07/2014

**(87) Publicação PCT:** WO 2015/011020 de 29/01/2015

**(85) Data do Início da Fase Nacional:** 19/01/2016

**(57) Resumo:** DECODIFICADOR DE ÁUDIO DE MÚLTIPLOS CANAIS, CODIFICADOR DE ÁUDIO DE MÚLTIPLOS CANAIS, MÉTODOS E PROGRAMA DE COMPUTADOR COM O USO DE AJUSTE COM BASE EM SINAL RESIDUAL DE CONTRIBUIÇÃO DE SINAL DESCORRELACIONADO Trata-se de um decodificador de áudio de múltiplos canais, para fornecer pelo menos dois sinais de áudio de saída com base em uma representação codificada, que é configurado para realizar uma combinação ponderada de um sinal de downmix, um sinal descorrelacionado e um sinal residual para obter um dos sinais de áudio de saída. O decodificador de áudio de múltiplos canais é configurado para determinar um peso que descreve uma contribuição do sinal descorrelacionado na combinação ponderada que depende do sinal residual. Um codificador de áudio de múltiplos canais para fornecer uma representação codificada de um sinal de áudio de múltiplos canais é configurado para obter um sinal de downmix com base no sinal de áudio de múltiplos canais, para fornecer parâmetros que descrevem dependências entre os canais do sinal de áudio de múltiplos canais e para fornecer um sinal residual. O codificador de áudio de múltiplos canais é configurado para variar uma quantidade de sinal residual incluído na representação codificada que depende do sinal de áudio (...).

## **RELATÓRIO DESCRITIVO**

### **“DECODIFICADOR DE ÁUDIO DE MÚLTIPLOS CANAIS E MÉTODO PARA FORNECER PELO MENOS DOIS SINAIS DE ÁUDIO DE SAÍDA COM BASE EM UMA REPRESENTAÇÃO CODIFICADA”**

#### **CAMPO DA TÉCNICA**

**[001]** Uma modalidade de acordo com a invenção se refere a um decodificador de áudio de múltiplos canais para fornecer pelo menos dois sinais de áudio de saída com base em uma representação codificada.

**[002]** Outra modalidade de acordo com a invenção se refere a um codificador de áudio de múltiplos canais para fornecer uma representação codificada de um sinal de áudio de múltiplos canais.

**[003]** Outra modalidade de acordo com a invenção se refere a um método para fornecer pelo menos dois sinais de áudio de saída com base em uma representação codificada.

**[004]** Outra modalidade de acordo com a invenção se refere a um método para fornecer uma representação codificada de um sinal de áudio de múltiplos canais.

**[005]** Outra modalidade de acordo com a presente invenção se refere a um programa de computador para realizar um dos métodos.

**[006]** Geralmente, algumas modalidades de acordo com a invenção se referem a uma criptografia paramétrica e residual combinada.

#### **ANTECEDENTES DA INVENÇÃO**

**[007]** Atualmente, a demanda de armazenamento e transmissão de conteúdo de áudio tem aumentado progressivamente. Ademais, as exigências de qualidade para o armazenamento e transmissão de conteúdo de áudio também tem aumentado progressivamente. Em conformidade, os conceitos para a codificação e decodificação de conteúdo de áudio foram aprimorados. Por exemplo, a chamada “criptografia de áudio avançada” (AAC) foi desenvolvida, a qual é descrita, por exemplo, na norma internacional ISO/IEC 13818-7: 2003.

**[008]** Ademais, algumas extensões espaciais foram produzidas, como, por

exemplo, o chamado conceito “MPEG surround”, que é descrito, por exemplo, na norma internacional ISO/IEC 23003-1:2007. Ademais os aprimoramentos adicionais para a codificação e decodificação de informações espaciais de sinais de áudio são descritos na norma internacional ISO/IEC 23003-2:2010, que se refere à chamada criptografia de objeto de áudio espacial. Ademais, um conceito de codificação/decodificação de áudio flexível (comutável), que fornece a possibilidade de codificar tanto sinais de áudio gerais quanto sinais de voz com boa eficiência de criptografia e de lidar com sinais de áudio de múltiplos canais, é definido na norma internacional ISO/IEC 23003-3:2012, que descreve o chamado conceito de “criptografia de áudio e voz unificada”.

**[009]** No entanto, há um desejo de fornecer um conceito ainda mais avançado de uma codificação e decodificação eficiente de sinais de áudio de múltiplos canais.

## **SUMÁRIO DA INVENÇÃO**

**[010]** Uma modalidade de acordo com a invenção proporciona um decodificador de áudio de múltiplos canais para fornecer pelo menos dois sinais de áudio de saída com base em uma representação codificada. O decodificador de áudio de múltiplos canais é configurado para realizar uma combinação ponderada de um sinal de downmix, um sinal descorrelacionado e um sinal residual para obter um dos sinais de áudio de saída. O decodificador de áudio de múltiplos canais é configurado para determinar um peso que descreve uma contribuição do sinal descorrelacionado na combinação ponderada que depende do sinal residual.

**[011]** Essa modalidade de acordo com a invenção se baseia na constatação de que os sinais de áudio de saída podem ser obtidos com base em uma representação codificada de uma maneira muito eficiente se um peso que descreve uma contribuição do sinal descorrelacionado à combinação ponderada de um sinal de downmix, um sinal descorrelacionado e um sinal residual forem ajustados dependendo do sinal residual. Em conformidade, ajustando-se o peso que descreve a contribuição do sinal descorrelacionado na combinação

ponderada que depende do sinal residual, é possível mesclar (ou desvanecimento) entre uma criptografia paramétrica (ou uma criptografia principalmente paramétrica) e uma criptografia residual (ou criptografia principalmente residual) sem transmitir informações de controle adicionais. Ademais, constatou-se que o sinal residual, que é incluído na representação codificada, é uma boa indicação para o peso que descreve a contribuição do sinal descorrelacionado na combinação ponderada, visto que é tipicamente preferencial dispor um peso (comparativamente) maior no sinal descorrelacionado se o sinal residual for (comparativamente) fraco (ou insuficiente para uma reconstrução da energia desejada) e dispor um peso (comparativamente) menor no sinal descorrelacionado se o sinal residual for (comparativamente) forte (ou suficiente para reconstruir a energia desejada). Em conformidade, o conceito mencionado acima permite uma transição gradual entre uma criptografia paramétrica (em que, por exemplo, características de energia desejadas e/ou características de correlação são sinalizadas por parâmetros e reconstruídas adicionando-se um sinal descorrelacionado) e uma criptografia residual (em que o sinal residual é usado para reconstruir sinais de áudio de saída - em alguns casos, até mesmo forma de onda dos sinais de áudio de saída - com base em um sinal de downmix). Em conformidade, é possível adaptar a técnica para a reconstrução, e também a qualidade da reconstrução, para os sinais decodificados sem ter sobrecarga de sinalização adicional.

**[012]** Em uma modalidade preferencial, o decodificador de áudio de múltiplos canais é configurado para determinar o peso que descreve a contribuição do sinal descorrelacionado na combinação ponderada que (também) depende do sinal descorrelacionado. Determinando-se o peso que descreve a contribuição do sinal descorrelacionado na combinação ponderada tanto dependente do sinal residual quanto dependente do sinal descorrelacionado, o peso pode ser bem ajustado às características de sinal, de modo que uma boa qualidade de reconstrução dos pelo menos dois sinais de áudio de saída com base na representação codificada (em particular, com base

no sinal de downmix, no sinal descorrelacionado e no sinal residual) possa ser alcançada.

**[013]** Em uma modalidade preferencial, o decodificador de áudio de múltiplos canais é configurado para obter parâmetros de upmix com base na representação codificada e para determinar o peso que descreve a contribuição do sinal descorrelacionado na combinação ponderada que depende dos parâmetros de upmix. Considerando-se os parâmetros de upmix, é possível reconstruir características desejadas dos sinais de áudio de saída (como, por exemplo, uma correlação desejada entre os sinais de áudio de saída, e/ou características de energia desejadas dos sinais de áudio de saída) para obter um valor desejado.

**[014]** Em uma modalidade preferencial, o decodificador de áudio de múltiplos canais é configurado para determinar o peso que descreve a contribuição do sinal descorrelacionado na combinação ponderada, de modo que o peso do sinal descorrelacionado diminua com o aumento da energia dos um ou mais sinais residuais. Esse mecanismo permite ajustar a precisão da reconstrução dos pelo menos dois sinais de áudio de saída que dependem da energia do sinal residual. Se a energia dos sinais residuais é comparativamente alta, o peso da contribuição do sinal descorrelacionado é comparativamente pequeno, de modo que o sinal descorrelacionado não mais afete de modo prejudicial uma alta qualidade da reprodução que é causada usando-se o sinal residual. Em contrapartida, se a energia do sinal residual é comparativamente baixa, ou até mesmo zero, um peso grande é dado ao sinal descorrelacionado, de modo que o sinal descorrelacionado possa levar as características dos sinais de áudio de saída a valores desejados de modo eficiente.

**[015]** Em uma modalidade preferencial, o decodificador de áudio de múltiplos canais é configurado para determinar o peso que descreve a contribuição do sinal descorrelacionado na combinação ponderada, de modo que um peso máximo, que é determinado por um parâmetro de upmix de sinal descorrelacionado, seja associado ao sinal descorrelacionado se uma energia

do sinal residual for zero, e de modo que um peso zero seja associado ao sinal descorrelacionado se uma energia do sinal residual ponderada com o uso de um coeficiente de ponderação de sinal residual for maior ou igual a uma energia do sinal descorrelacionado, ponderada com o parâmetro de upmix de sinal descorrelacionado. Essa modalidade se baseia na constatação de que a energia desejada, que deve ser adicionada ao sinal de downmix, é determinada pela energia do sinal descorrelacionado, ponderada com o parâmetro de upmix de sinal descorrelacionado. Em conformidade, conclui-se que não é mais necessário adicionar o sinal descorrelacionado se a energia do sinal residual, ponderada com o coeficiente de ponderação de sinal residual, for maior ou igual à dita energia do sinal descorrelacionado, ponderada com o parâmetro de upmix de sinal descorrelacionado. Em outras palavras, o sinal descorrelacionado não é mais usado para fornecer os pelo menos dois sinais de áudio de saída caso se julgue que o sinal residual porta energia suficiente (por exemplo, suficiente a fim de alcançar uma energia total suficiente).

**[016]** Em uma modalidade preferencial, o decodificador de áudio de múltiplos canais é configurado para computar um valor de energia ponderado do sinal descorrelacionado, ponderada dependendo de um ou mais parâmetros de upmix de sinal descorrelacionado, e para computar um valor de energia ponderado do sinal residual, ponderado com o uso de um ou mais parâmetros de upmix de sinal residual (que pode ser igual ao coeficiente de ponderação de sinais residuais mencionado acima), para determinar um fator que depende do valor de energia ponderado do sinal descorrelacionado e do valor de energia ponderado do sinal residual, e para obter um peso que descreve a contribuição do sinal descorrelacionado a (pelo menos) um dos sinais de saída de áudio com base no fator. Constatou-se, que esse procedimento é bem adequado para uma computação eficiente do peso que descreve a contribuição do sinal descorrelacionado a um ou mais dos sinais de áudio de saída.

**[017]** Em uma modalidade preferencial, o decodificador de áudio de múltiplos canais é configurado para multiplicar o fator por um parâmetro de upmix

de sinal descorrelacionado, para obter o peso que descreve a contribuição do sinal descorrelacionado a (pelo menos) um dos sinais de áudio de saída. Usando-se tal procedimento, é possível considerar tanto um ou mais parâmetros que descrevem características de sinal desejadas dos pelo menos dois sinais de áudio de saída (que são descritas pelo parâmetro de upmix de sinal descorrelacionado) e a relação entre a energia de sinal descorrelacionado e a energia do sinal residual, a fim de determinar o peso que descreve a contribuição do sinal descorrelacionado na combinação ponderada. Desse modo, há tanto a possibilidade de mesclar (ou desvanecimento) entre uma criptografia paramétrica (ou criptografia predominantemente paramétrica) quanto uma criptografia residual (ou a criptografia predominantemente residual) enquanto ainda se consideram as características desejadas dos sinais de áudio de saída (que são refletidas pelo parâmetro de upmix de sinal descorrelacionado).

**[018]** Em uma modalidade preferencial, o decodificador de áudio de múltiplos canais é configurado para computar a energia do sinal descorrelacionado, ponderado com o uso dos parâmetros de upmix de sinal descorrelacionado, em uma pluralidade de canais de upmix e intervalos de tempo, para obter o valor de energia ponderado do sinal descorrelacionado. Em conformidade, é possível evitar fortes variações do valor de energia ponderado do sinal descorrelacionado. Desse modo, um ajuste estável do decodificador de áudio de múltiplos canais é alcançado.

**[019]** De modo semelhante, o decodificador de áudio de múltiplos canais é configurado para computar a energia do sinal residual, ponderado com o uso de parâmetros de upmix de sinal residual, em uma pluralidade de canais de upmix e intervalos de tempo, para obter o valor de energia ponderado do sinal residual. Em conformidade, um ajuste estável do decodificador de áudio de múltiplos canais é alcançado, visto que fortes variações do valor de energia ponderado do sinal residual são evitadas.

**[020]** No entanto, o período de referência pode ser selecionado suficientemente curto para permitir um ajuste dinâmico da ponderação.

**[021]** Em uma modalidade preferencial, o decodificador de áudio de múltiplos canais é configurado para computar o fator que depende de uma diferença entre o valor de energia ponderado do sinal descorrelacionado e o valor de energia ponderado do sinal residual. Uma computação, que “compara” o valor de energia ponderado do sinal descorrelacionado e o valor de energia ponderado do sinal residual permite suplementar o sinal residual (ou a versão ponderada do sinal residual) com o uso (da versão ponderada) do sinal descorrelacionado, em que o peso que descreve a contribuição do sinal descorrelacionado é ajustado às necessidades para o fornecimento dos pelo menos dois sinais de canal de áudio.

**[022]** Em uma modalidade preferencial, o decodificador de áudio de múltiplos canais é configurado para computar o fator que depende de uma razão entre uma diferença entre o valor de energia ponderado do sinal descorrelacionado e o valor de energia ponderado do sinal residual, e o valor de energia ponderado do sinal descorrelacionado. Constatou-se, que a computação do fator que depende dessa razão suscita bons resultados particulares. Ademais, deve-se observar que a razão descreve qual porção da energia total do sinal descorrelacionado (ponderada com o uso do parâmetro de upmix de sinal descorrelacionado) é necessária na presença do sinal residual a fim de alcançar uma boa impressão de audição (ou de modo equivalente, para ter substancialmente a mesma energia de sinal nos sinais de áudio de saída quando em comparação com o caso em que não há sinal residual).

**[023]** Em uma modalidade preferencial, o decodificador de áudio de múltiplos canais é configurado para determinar pesos que descrevem contribuições do sinal descorrelacionado a dois ou mais sinais de áudio de saída. Nesse caso, o decodificador de áudio de múltiplos canais é configurado para determinar uma contribuição do sinal descorrelacionado a um primeiro sinal de áudio de saída com base no valor de energia ponderado do sinal descorrelacionado e um parâmetro de upmix de sinal descorrelacionado de primeiro canal. Ademais, o decodificador de áudio de múltiplos canais é



configurado para determinar uma contribuição do sinal descorrelacionado a um segundo canal de áudio de saída com base no valor de energia ponderado do sinal descorrelacionado e um parâmetro de upmix de sinal descorrelacionado de segundo canal. Em conformidade, dois sinais de áudio de saída podem ser dotados de esforço moderado e boa qualidade de áudio, em que as diferenças entre os dois sinais de áudio de saída são consideradas com o uso de um parâmetro de upmix de sinal descorrelacionado de primeiro canal e um parâmetro de upmix de sinal descorrelacionado de segundo canal.

**[024]** Em uma modalidade preferencial, o decodificador de áudio de múltiplos canais é configurado para desabilitar uma contribuição do sinal descorrelacionado à combinação ponderada se uma energia residual exceder uma energia de descorrelacionador (isto é, uma energia do sinal descorrelacionado, ou de uma versão ponderada desse). Em conformidade, é possível comutar para uma criptografia residual pura, sem o uso do sinal descorrelacionado, se o sinal residual portar energia suficiente, se a energia residual exceder a energia do descorrelacionador.

**[025]** Em uma modalidade preferencial, o decodificador de áudio é configurado para determinar, no sentido da banda, o peso que descreve a contribuição do sinal descorrelacionado na combinação ponderada que depende de uma determinação em termos de banda de um valor de energia ponderado do sinal residual. Em conformidade, pode-se decidir de modo flexível, sem uma sobrecarga de sinalização adicional, em que bandas de frequência um refinamento dos pelo menos dois sinais de áudio de saída deve ser baseado (ou deve ser predominantemente baseado) em uma criptografia paramétrica, e em que bandas de frequência o refinamento dos pelo menos dois sinais de áudio de saída deve ser baseado (ou deve ser predominantemente baseado) em uma criptografia residual. Desse modo, pode-se decidir de modo flexível em que bandas de frequência uma reconstrução de forma de onda (ou pelo menos uma reconstrução de forma de onda parcial) deve ser realizada usando-se (pelo menos predominantemente) a criptografia residual ao mesmo tempo em que se

mantém o peso do sinal descorrelacionado comparativamente baixo. Desse modo, é possível obter uma boa qualidade de áudio aplicando-se seletivamente a criptografia paramétrica (que é principalmente baseada no fornecimento de um sinal descorrelacionado) e a criptografia residual (que é principalmente baseada no fornecimento de um sinal residual).

**[026]** Em uma modalidade preferencial, o decodificador de áudio é configurado para determinar o peso que descreve a contribuição do sinal descorrelacionado em uma combinação ponderada para cada quadro dos sinais de áudio de saída. Em conformidade, uma boa resolução de temporização pode ser obtida, que permite comutar de modo flexível entre uma criptografia paramétrica (ou criptografia predominantemente paramétrica) e a criptografia residual (ou criptografia predominantemente residual) entre quadros subsequentes. Em conformidade, a decodificação de áudio pode ser ajustada às características do sinal de áudio com uma boa resolução de tempo.

**[027]** Outra modalidade de acordo com a invenção proporciona um decodificador de áudio de múltiplos canais para fornecer pelo menos dois sinais de áudio de saída com base em uma representação codificada. O decodificador de áudio de múltiplos canais é configurado para obter (pelo menos) um dos sinais de áudio de saída com base em uma representação codificada de um sinal de downmix, uma pluralidade de parâmetros espaciais codificados e uma representação codificada de um sinal residual. O decodificador de áudio de múltiplos canais é configurado para mesclar entre uma criptografia paramétrica e a criptografia residual que depende do sinal residual. Em conformidade, um conceito de decodificação de áudio muito flexível é alcançado, em que o melhor modo de decodificação (decodificação e criptografia paramétrica versus decodificação e criptografia residual) pode ser selecionado sem sobrecarga de sinalização adicional. Ademais, a consideração explicada acima também é aplicada.

**[028]** Uma modalidade de acordo com a invenção proporciona um codificador de áudio de múltiplos canais para fornecer uma representação

codificada de um sinal de áudio de múltiplos canais. O codificador de áudio de múltiplos canais é configurado para obter um sinal de downmix com base no sinal de áudio de múltiplos canais. Ademais, o codificador de áudio de múltiplos canais é configurado para fornecer parâmetros que descrevem dependências entre os canais do sinal de áudio de múltiplos canais e para fornecer um sinal residual. Ademais, o codificador de áudio de múltiplos canais é configurado para variar uma quantidade de um sinal residual incluído na representação codificada que depende do sinal de áudio de múltiplos canais. Variando-se uma quantidade de sinal residual incluído para a representação codificada, é possível ajustar de modo flexível o processo de codificação às características do sinal. Por exemplo, é possível incluir uma quantidade comparativamente grande de sinal residual na representação codificada para porções (por exemplo, para porções temporais e/ou para porções de frequência) em que é desejável preservar, pelo menos parcialmente, a forma de onda do sinal de áudio decodificado. Desse modo, uma reconstrução baseada em sinal residual mais precisa do sinal de áudio de múltiplos canais é possibilitada pela possibilidade de variar a quantidade de sinal residual incluído na representação codificada. Ademais, deve-se observar que, em conjunto com o decodificador de áudio de múltiplos canais discutido acima, um conceito muito eficiente é produzido, visto que o decodificador de áudio de múltiplos canais descrito acima nem mesmo necessita de sinalização adicional para mesclar entre uma criptografia (predominantemente) paramétrica e uma criptografia (predominantemente) residual. Em conformidade, o codificador de múltiplos canais discutido no presente documento permite explorar os benefícios que são possíveis usando-se o codificador de áudio de múltiplos canais discutido acima.

**[029]** Em uma modalidade preferencial, o codificador de áudio de múltiplos canais é configurado para variar uma largura de banda do sinal residual que depende do sinal de áudio de múltiplos canais. Em conformidade, é possível ajustar o sinal residual, de modo que o sinal residual auxilie na reconstrução de faixas de frequência ou bandas de frequência mais importantes de modo

psicoacústico.

**[030]** Em uma modalidade preferencial, o codificador de áudio de múltiplos canais é configurado para selecionar bandas de frequência para as quais o sinal residual é incluído na representação codificada que depende do sinal de áudio de múltiplos canais. Em conformidade, o codificador de áudio de múltiplos canais pode decidir para quais bandas de frequência é necessário, ou mais benéfico, incluir um sinal residual (em que o sinal residual tipicamente resulta em reconstrução de forma de onda pelo menos parcial). Por exemplo, as bandas de frequência significativas de modo psicoacústico podem ser consideradas. Além disso, a presença de eventos transientes também pode ser considerada, visto que um sinal residual tipicamente auxilia no aprimoramento da renderização de transientes em um decodificador de áudio. Ademais, a taxa de bits disponível também pode ser levada em consideração para decidir que quantidade de sinal residual é incluída na representação codificada.

**[031]** Em uma modalidade preferencial, o codificador de áudio de múltiplos canais é configurado para incluir seletivamente o sinal residual na representação codificada para bandas de frequência para as quais o sinal de áudio de múltiplos canais é tonal, ao mesmo tempo em que se omite a inclusão do sinal residual na representação codificada para bandas de frequência em que o sinal de áudio de múltiplos canais é não tonal. Essa modalidade se baseia na consideração de que uma qualidade de áudio obténível no lado de um decodificador de áudio pode ser aprimorada se bandas de frequência tonais forem reproduzidas com qualidade particularmente alta e preferencialmente com o uso de reconstrução de forma de onda pelo menos parcial. Em conformidade, é vantajoso incluir seletivamente o sinal residual na representação codificada para bandas de frequência para as quais o sinal de áudio de múltiplos canais é tonal, visto que isso resulta em um bom comprometimento entre taxa de bits e qualidade de áudio.

**[032]** Em uma modalidade preferencial, o codificador de áudio de múltiplos canais é configurado para incluir seletivamente o sinal residual na representação

codificada para porções de tempo e/ou banda de frequência, em que a formação do sinal de downmix resulta em um cancelamento de componentes de sinal do sinal de áudio de múltiplos canais. Constatou-se, que é difícil ou até mesmo impossível reconstruir de modo adequado múltiplos sinais de áudio com base em um sinal de downmix se houver um cancelamento de componentes do sinal de áudio de múltiplos canais, devido ao fato de que até mesmo uma decorrelação ou uma previsão não pode recuperar componentes de sinal que foram cancelados durante a formação do sinal de downmix. Em tal caso, o uso de um sinal residual é uma maneira eficiente de evitar uma degradação significativa do sinal de áudio de múltiplos canais reconstruído. Desse modo, esse conceito auxilia no aprimoramento da qualidade do áudio ao mesmo tempo em que se evita um esforço de sinalização (por exemplo, quando considerado em conjunto com o decodificador de áudio descrito acima).

**[033]** Em uma modalidade preferencial, o codificador de áudio de múltiplos canais é configurado para detectar um cancelamento de componentes de sinal do sinal de áudio de múltiplos canais no sinal de downmix, e o decodificador de áudio de múltiplos canais também é configurado para ativar o fornecimento do sinal residual em resposta a um resultado da detecção. Em conformidade, há uma maneira eficiente de evitar uma má qualidade de áudio.

**[034]** Em uma modalidade preferencial, o codificador de áudio de múltiplos canais é configurado para computar o sinal residual com o uso de uma combinação linear de pelo menos dois sinais de canal do sinal de áudio de múltiplos canais e uma dependência de coeficientes de upmix a serem usados no lado de um decodificador de múltiplos canais. Consequentemente, o sinal residual é computado em de uma maneira eficiente e bem adaptada para uma reconstrução do sinal de áudio de múltiplos canais no lado de um decodificador de áudio de múltiplos canais.

**[035]** Em uma modalidade, o codificador de áudio de múltiplos canais é configurado para codificar os coeficientes de upmix com o uso dos parâmetros que descrevem dependências entre os canais do sinal de áudio de múltiplos

canais, ou para derivar os coeficientes de upmix a partir dos parâmetros que descrevem dependências entre os canais do sinal de áudio de múltiplos canais. Em conformidade, o fornecimento do sinal residual pode ser realizado de modo eficiente com base em parâmetros, que também são usados para uma criptografia paramétrica.

**[036]** Em uma modalidade preferencial, o codificador de áudio de múltiplos canais é configurado para determinar, de modo variante no tempo, a quantidade de sinal residual incluído na representação codificada com o uso de um modelo psicoacústico. Em conformidade, uma quantidade comparativamente alta de sinal residual pode ser incluída para porções (porções temporais, ou porções de frequência, ou porções de frequência de tempo) do sinal de áudio de múltiplos canais que compreendem uma relevância psicoacústica comparativamente alta, enquanto uma quantidade (comparativamente) menor de sinal residual pode ser incluída para porções temporais ou porções de frequência ou porções de frequência de tempo do sinal de áudio de múltiplos canais com uma relevância psicoacústica comparativamente baixa. Em conformidade, uma boa troca de entre taxa de bits e qualidade de áudio pode ser alcançada.

**[037]** Em uma modalidade preferencial, o codificador de áudio de múltiplos canais é configurado para determinar, de modo variante no tempo, a quantidade de sinal residual incluído na representação codificada que depende de uma taxa de bits atualmente disponível. Em conformidade, a qualidade do áudio pode ser adaptada à taxa de bits disponível, que permite alcançar a melhor qualidade de áudio possível para a taxa de bits atualmente disponível.

**[038]** Uma modalidade de acordo com a invenção proporciona um método para fornecer pelo menos dois sinais de áudio de saída com base em uma representação codificada. O método compreende realizar uma combinação ponderada de um sinal de downmix, um sinal descorrelacionado e um sinal residual para obter um dos sinais de áudio de saída. Um peso que descreve uma contribuição do sinal descorrelacionado na combinação ponderada é determinado dependendo do sinal residual. Esse método se baseia nas mesmas

considerações que o decodificador de áudio descrito acima.

**[039]** Outra modalidade de acordo com a invenção proporciona um método para fornecer pelo menos dois sinais de áudio de saída com base em uma representação codificada. O método compreende obter (pelo menos) um dos sinais de áudio de saída com base em uma representação codificada de um sinal de downmix, uma pluralidade de parâmetros espaciais codificados e uma representação codificada de um sinal residual. Uma mescla (ou desvanecimento) é realizada entre uma criptografia paramétrica e uma criptografia residual que depende do sinal residual. Esse método também é baseado nas mesmas considerações que o decodificador de áudio descrito acima.

**[040]** Outra modalidade de acordo com a invenção proporciona um método para fornecer uma representação codificada de um sinal de áudio de múltiplos canais. O método compreende obter um sinal de downmix com base no sinal de áudio de múltiplos canais, fornecer parâmetros que descrevem dependências entre os canais do sinal de áudio de múltiplos canais e fornecer um sinal residual. Uma quantidade de sinal residual incluído na representação codificada é variada dependendo do sinal de áudio de múltiplos canais. Esse método se baseia nas mesmas considerações que o codificador de áudio descrito acima.

**[041]** Modalidades adicionais, de acordo com a invenção proporcionam programas de computador para realizar os métodos descritos no presente documento.

## **BREVE DESCRIÇÃO DAS FIGURAS**

**[042]** Modalidades de acordo com a invenção serão subsequentemente descritas com referência às figuras anexas, em que

**[043]** Figura 1 mostra um diagrama de blocos esquemático de um codificador de áudio de múltiplos canais, de acordo com uma modalidade de a invenção;

**[044]** Figura 2 mostra um diagrama de blocos esquemático de um decodificador de áudio de múltiplos canais, de acordo com uma modalidade de

a invenção;

**[045]** Figura 3 mostra um diagrama de blocos esquemático de um decodificador de áudio de múltiplos canais, de acordo com a outra modalidade da presente invenção;

**[046]** Figura 4 mostra um fluxograma de um método para fornecer uma representação codificada de um sinal de áudio de múltiplos canais, de acordo com uma modalidade de a invenção;

**[047]** Figura 5 mostra um fluxograma de um método para fornecer pelo menos dois sinais de áudio de saída com base em uma representação codificada, de acordo com uma modalidade de a invenção;

**[048]** Figura 6 mostra um fluxograma de um método para fornecer pelo menos dois sinais de áudio de saída com base em uma representação codificada, de acordo com outra modalidade de a invenção; e

**[049]** Figura 7 mostra um diagrama de fluxo de um decodificador, de acordo com uma modalidade da presente invenção; e

**[050]** Figura 8 mostra uma representação esquemática de um Decodificador Residual Híbrido.

## **DESCRIÇÃO DETALHADA DAS MODALIDADES**

**[051]** Codificador de áudio de múltiplos canais de acordo com a Figura 1

**[052]** A Figura 1 mostra um diagrama de blocos esquemático de um codificador de áudio de múltiplos canais 100 para fornecer uma representação codificada de um sinal de múltiplos canais.

**[053]** O codificador de áudio de múltiplos canais 100 é configurado para receber um sinal de áudio de múltiplos canais 110 e para fornecer, com base nesse, uma representação codificada 112 do sinal de áudio de múltiplos canais 110. O codificador de áudio de múltiplos canais 100 compreende um processador (ou dispositivo de processamento) 120, que é configurado para receber o sinal de áudio de múltiplos canais e para obter um sinal de downmix 122 com base no sinal de áudio de múltiplos canais 110. O processador 120 é adicionalmente configurado para fornecer parâmetros 124 que descrevem



dependências entre os canais do sinal de áudio de múltiplos canais 110. Ademais, o processador 120 é configurado para fornecer um sinal residual 126. Ademais, o codificador de áudio de múltiplos canais compreende um processamento de sinal residual 130, que é configurado para variar uma quantidade de sinal residual incluído na representação codificada 112 que depende do sinal de áudio de múltiplos canais 110.

**[054]** No entanto, deve-se observar que não é necessário que o decodificador de áudio de múltiplos canais compreenda um processador 120 separado e um processamento de sinal residual 130 separado. Em vez disso, é suficiente se o codificador de áudio de múltiplos canais for, de algum modo, configurado para realizar a funcionalidade do processador 120 e do processamento de sinal residual 130.

**[055]** Considerando-se a funcionalidade do codificador de áudio de múltiplos canais 100, pode-se observar que os sinais de canal do sinal de áudio de múltiplos canais 110 são tipicamente codificados com o uso de uma codificação de múltiplos canais, em que a representação codificada 112 tipicamente compreende (em uma forma codificada) o sinal de downmix 122, os parâmetros 124 que descrevem dependências entre canais (ou sinais de canal) do sinal de áudio de múltiplos canais 110 e o sinal residual 126. O sinal de downmix 122 pode, por exemplo, ser baseado uma combinação (por exemplo, combinação linear) dos sinais de canal do sinal de áudio de múltiplos canais. No entanto, um sinal de downmix 122 pode ser fornecido com base em uma pluralidade de sinais de canal do sinal de áudio de múltiplos canais. No entanto, alternativamente, dois ou mais sinais de downmix podem ser associados a um número maior (tipicamente maior do que o número de sinais de downmix) de sinais de canal do sinal de áudio de múltiplos canais 110. Os parâmetros 124 podem descrever dependências (por exemplo, uma correlação, uma covariância, uma relação de nível ou similares) entre canais (ou sinais de canal) do sinal de áudio de múltiplos canais 110. Em conformidade, os parâmetros 124 servem para o propósito de derivar uma versão reconstruída dos sinais de canal do sinal

de áudio de múltiplos canais 110 com base no sinal de downmix 122 no lado de um decodificador de áudio. Para essa finalidade, os parâmetros 124 descrevem características desejadas (por exemplo, características individuais ou características relativas) dos sinais de canal do sinal de áudio de múltiplos canais, de modo que um codificador de áudio, que usa uma decodificação paramétrica, pode reconstruir sinais de canal com base nos um ou mais sinais de downmix 122.

**[056]** Além disso, o decodificador de áudio de múltiplos canais 100 fornece o sinal residual 126, que tipicamente representa componentes de sinal que, de acordo com o que se espera ou se estima do codificador de áudio de múltiplos canais, não podem ser reconstruídos por um decodificador de áudio (por exemplo, por um decodificador de áudio que segue uma determinada regra de processamento) com base no sinal de downmix 122 e nos parâmetros 124. Em conformidade, o sinal residual 126 pode tipicamente ser considerado como um sinal de refinamento, que permite uma reconstrução de forma de onda, ou pelo menos uma reconstrução de forma de onda parcial, no lado de um decodificador de áudio.

**[057]** No entanto, o codificador de áudio de múltiplos canais 100 é configurado para variar uma quantidade de sinal residual incluído na representação codificada 112 que depende do sinal de áudio de múltiplos canais 110. Em outras palavras, o codificador de áudio de múltiplos canais pode, por exemplo, decidir quanto à intensidade (ou a energia) do sinal residual 126 que é incluída na representação codificada 112. Adicional ou alternativamente, o codificador de áudio de múltiplos canais 100 pode decidir em relação a quais bandas de frequência e/ou a quantas bandas de frequência o sinal residual é incluído na representação codificada 112. Variando-se a “quantidade” de sinal residual 126 incluída na representação codificada 112 que depende do sinal de áudio de múltiplos canais (e/ou que depende de uma taxa de bits disponível), o codificador de áudio de múltiplos canais 100 pode determinar de modo flexível com que precisão os sinais de canal do sinal de áudio de múltiplos canais 110

podem ser reconstruídos no lado de um decodificador de áudio com base na representação codificada 112. Desse modo, a precisão com que os sinais de canal do sinal de áudio de múltiplos canais 110 pode ser reconstruída pode ser adaptada a uma relevância psicoacústica de diferentes porções de sinal dos sinais de canal do sinal de áudio de múltiplos canais 110 (como, por exemplo, porções temporais, porções de frequência e/ou porções de frequência/tempo). Desse modo, porções de sinal de alta relevância psicoacústica (como, por exemplo, porções de sinal tonal ou porções de sinal que compreendem eventos transientes podem ser codificadas com resolução particularmente alta incluindo-se uma “grande quantidade” do sinal residual 126 na representação codificada. Por exemplo, pode-se conseguir que um sinal residual com uma energia comparativamente alta seja incluída na representação codificada 112 para porções de sinal de alta relevância psicoacústica. Ademais, pode-se conseguir que um sinal residual de alta energia seja incluído na representação codificada 112 se o sinal de downmix 122 compreender uma “qualidade insuficiente”, por exemplo, se houver um cancelamento substancial de componentes de sinal ao se combinarem os sinais de canal do sinal de áudio de múltiplos canais 112 no sinal de downmix 122. Em outras palavras, o decodificador de áudio de múltiplos canais 100 pode seletivamente inserir uma “maior quantidade” de sinal residual (por exemplo, um sinal residual que tem uma energia comparativamente alta) na representação codificada 112 para porções de sinal do sinal de áudio de múltiplos canais 110 para que o fornecimento de uma quantidade comparativamente grande do sinal residual forneça um aprimoramento significativo dos sinais de canal reconstruídos (reconstruídos no lado de um decodificador de áudio).

**[058]** Em conformidade, a variação da quantidade de sinal residual incluído na representação codificada que depende do sinal de áudio de múltiplos canais 110 permite adaptar a representação codificada 112 (por exemplo, o sinal residual 126, que é incluído na representação codificada em uma forma codificada) do sinal de áudio de múltiplos canais 110, de modo que uma boa

troca entre a eficiência taxa de bits e a qualidade de áudio do sinal de áudio de múltiplos canais reconstruído (reconstruído no lado de um decodificador de áudio) possa ser alcançada.

**[059]** Deve-se observar que o codificador de áudio de múltiplos canais 100 pode ser opcionalmente aprimorado de muitas maneiras diferentes. Por exemplo, o codificador de áudio de múltiplos canais pode ser configurado para variar uma largura de banda do sinal residual 126 (que é incluída na representação codificada) que depende do sinal de áudio de múltiplos canais 110. Em conformidade, a quantidade de sinal residual incluído na representação codificada 112 pode ser adaptada às bandas de frequência perceptualmente mais importantes.

**[060]** Opcionalmente, o decodificador de áudio de múltiplos canais pode ser configurado para selecionar bandas de frequência para as quais o sinal residual 126 é incluído na representação codificada 112 que depende do sinal de áudio de múltiplos canais 110. Em conformidade, a representação codificada 120 (mais precisamente, a quantidade de sinal residual incluído na representação codificada 112) pode ser adaptada ao sinal de áudio de múltiplos canais, por exemplo, às bandas de frequência perceptualmente mais importantes do sinal de áudio de múltiplos canais 110.

**[061]** Opcionalmente, o codificador de áudio de múltiplos canais pode ser configurado para incluir o sinal residual 126 na representação codificada para bandas de frequência para as quais o sinal de áudio de múltiplos canais é tonal. Além disso, o codificador de áudio de múltiplos canais pode ser configurado para não incluir o sinal residual 126 na representação codificada 112 para bandas de frequência em que o sinal de áudio de múltiplos canais é non-tonal (a menos que qualquer outra condição específica seja preenchida que cause uma inclusão do sinal residual na representação codificada para uma banda de frequência específica). Desse modo, o sinal residual pode ser seletivamente incluído na representação codificada para bandas de frequência tonais perceptualmente importantes.

**[062]** Opcionalmente, o codificador de áudio de múltiplos canais 100 pode ser configurado para incluir seletivamente o sinal residual na representação codificada para porções de tempo e/ou para bandas de frequência em que a formação do sinal de downmix resulta em um cancelamento de componentes de sinal do sinal de áudio de múltiplos canais. Por exemplo, o codificador de áudio de múltiplos canais pode ser configurado para detectar um cancelamento de componentes de sinal do sinal de áudio de múltiplos canais 110 no sinal de downmix 122, e para ativar o fornecimento do sinal residual 126 (por exemplo, a inclusão do sinal residual 126 na representação codificada 112) em resposta ao resultado da detecção. Em conformidade, se o downmix (ou qualquer outra combinação tipicamente linear) de sinais de canal do sinal de áudio de múltiplos canais 110 no sinal de downmix 122 resultar em um cancelamento de componentes de sinal do sinal de áudio de múltiplos canais 112 (que pode ser causado, por exemplo, por componentes de sinal de diferentes sinais de canal que são variados em fase em 180 graus), o sinal residual 126, que auxilia na superação do efeito prejudicial desse cancelamento durante a reconstrução do sinal de áudio de múltiplos canais 110 em um decodificador de áudio, será incluído na representação codificada 112. Por exemplo, o sinal residual 126 pode ser seletivamente incluído na representação codificada 112 para bandas de frequência para as quais há tal cancelamento.

**[063]** Opcionalmente, o codificador de áudio de múltiplos canais pode ser configurado para computar o sinal residual com o uso de uma combinação linear de pelo menos dois sinais de canal do sinal de áudio de múltiplos canais e que depende de coeficientes de upmix a serem usados no lado de um decodificador de áudio de múltiplos canais. Tal computação de um sinal residual é eficiente e permite uma simples reconstrução dos sinais de canal no lado de um decodificador de áudio.

**[064]** Opcionalmente, o codificador de áudio de múltiplos canais pode ser configurado para codificar os coeficientes de upmix com o uso do parâmetro 124 que descreve dependências entre os canais do sinal de áudio de múltiplos

canais, ou para derivar os coeficientes de upmix a partir dos parâmetros que descrevem dependências entre os canais do sinal de áudio de múltiplos canais. Em conformidade, os parâmetros 124 (que pode, por exemplo, ser parâmetros de diferença de nível intracanal, parâmetros de correlação intracanal ou similares) podem ser usados tanto para a criptografia paramétrica (codificação ou decodificação) quanto para a criptografia assistida por sinal residual (codificação ou decodificação). Desse modo, o uso do sinal residual 126 não resulta em uma sobrecarga de sinalização adicional. Em vez disso, os parâmetros 124, que são usados para a criptografia paramétrica (codificação/decodificação) de qualquer modo, são reusados também para a criptografia residual (codificação/decodificação). Desse modo alta eficiência de criptografia pode ser alcançada.

**[065]** Opcionalmente, o decodificador de áudio de múltiplos canais pode ser configurado para determinar, de modo variante no tempo, a quantidade de sinal residual incluído na representação codificada com o uso de um modelo psicoacústico. Em conformidade, a precisão de codificação pode ser adaptada a características psicoacústicas do sinal, que tipicamente resultam em uma boa eficiência de taxa de bits.

**[066]** No entanto, deve-se observar que o codificador de áudio de múltiplos canais pode opcionalmente ser suplementado por qualquer um dos recursos ou funcionalidades descritas no presente documento (tanto na descrição e quanto nas reivindicações). Ademais, o codificador de áudio de múltiplos canais também pode ser adaptado paralelamente ao decodificador de áudio descrito no presente documento para cooperar com o decodificador de áudio.

**[067]** 2. Decodificador de áudio de múltiplos canais, de acordo com a Figura 2

**[068]** A Figura 2 mostra um diagrama de blocos esquemático de um decodificador de áudio de múltiplos canais 200 de acordo com uma modalidade da presente invenção.

**[069]** O decodificador de áudio de múltiplos canais 200 é configurado para receber uma representação codificada 210 e para fornecer, com base nesse, pelo menos dois sinais de áudio de saída 212, 214. O decodificador de áudio de múltiplos canais 200 pode, por exemplo, compreender um combinador de ponderação 220, que é configurado para realizar uma combinação ponderada de um sinal de downmix 222, um sinal descorrelacionado 224 e um sinal residual 226, para obter (pelo menos) um dos sinais de saída, por exemplo, o primeiro sinal de áudio de saída 212. Deve-se observar, no presente documento, que o sinal de downmix 212, o sinal descorrelacionado 224 e o sinal residual 226 podem, por exemplo, ser derivados da representação codificada 210, em que a representação codificada 210 pode portar uma representação codificada do sinal de downmix 220 e uma representação codificada do sinal residual 226. Ademais, o sinal descorrelacionado 224 pode, por exemplo, ser derivado do sinal de downmix 222 ou pode ser derivado com o uso de informações adicionais incluídas na representação codificada 210. No entanto, o sinal descorrelacionado também pode ser fornecido sem quaisquer informações específicas da representação codificada 210.

**[070]** O decodificador de áudio de múltiplos canais 200 também é configurado para determinar um peso que descreve uma contribuição do sinal descorrelacionado 224 na combinação ponderada que depende do sinal residual 226. Por exemplo, o decodificador de áudio de múltiplos canais 200 pode compreender um determinador de peso 230, que é configurado para determinar um peso 232 que descreve a contribuição do sinal descorrelacionado 224 na combinação ponderada (por exemplo, a contribuição do sinal descorrelacionado 224 ao primeiro sinal de áudio de saída 212) com base no sinal residual 226.

**[071]** Considerando-se a funcionalidade do decodificador de áudio de múltiplos canais 200, deve-se observar que a contribuição do sinal descorrelacionado 224 à combinação ponderada e, consequentemente ao primeiro sinal de áudio de saída 212, é ajustada de uma maneira flexível (por exemplo, com variação temporal e dependente de frequência) dependendo do

signal residual 226, sem sobrecarga de sinalização adicional. Em conformidade, a quantidade de signal descorrelacionado 224, que é incluída no primeiro signal de áudio de saída 212, é adaptada dependendo da quantidade de signal residual 226 que é incluída no primeiro signal de áudio de saída 212, de modo que uma boa qualidade do primeiro signal de áudio de saída 212 seja alcançada. Em conformidade, é possível obter uma ponderação apropriada do signal descorrelacionado 224 em quaisquer circunstâncias e sem uma sobrecarga de sinalização adicional. Desse modo, com o uso do decodificador de áudio de múltiplos canais 200, uma boa qualidade do signal de áudio de saída decodificado 212 pode ser alcançada com taxa de bits moderada. Uma precisão da reconstrução pode ser ajustada de modo flexível por um codificador de áudio, em que o codificador de áudio pode determinar uma quantidade de signal residual 226 que é incluída na representação codificada 212 (por exemplo, a quantidade da energia do signal residual 226 incluída na representação codificada 210, ou a quantas bandas de frequência o signal residual 226 incluído na representação codificada 210 se refere), e o decodificador de áudio de múltiplos canais 200 pode reagir em conformidade e ajustar a ponderação do signal descorrelacionado 224 para corresponder à quantidade de signal residual 226 incluída na representação codificada 210. Consequentemente, se há uma grande quantidade de signal residual 226 incluída na representação codificada 210 (por exemplo, para uma banda de frequência específica, ou para uma porção temporal específica), a combinação ponderada 220 pode predominantemente (ou exclusivamente) considerar o signal residual 226 ao mesmo tempo em que fornece pouco peso (ou nenhum peso) ao signal descorrelacionado 224. Em contrapartida, se há apenas uma quantidade menor de um signal residual 226 incluída na representação codificada 210, a combinação ponderada 220 pode predominantemente (ou exclusivamente) considerar o signal descorrelacionado 224, mas apenas a um grau comparativamente pequeno (ou nenhum) o signal residual 226 além do signal de downmix 222. Desse modo, o decodificador de áudio de múltiplos canais 200 pode cooperar de modo flexível com um



codificador de áudio de múltiplos canais apropriado e ajustar a combinação ponderada 220 para alcançar a melhor qualidade de áudio possível em quaisquer circunstâncias (independentemente de se uma quantidade menor ou uma quantidade maior de sinal residual 226 é incluída na representação codificada 210).

**[072]** Deve-se observar que o segundo sinal de áudio de saída 214 pode ser gerado de uma maneira semelhante. No entanto, não é necessário aplicar os mesmos mecanismos ao segundo sinal de áudio de saída 214, por exemplo, se houver diferentes exigências de qualidade em relação ao segundo sinal de áudio de saída.

**[073]** Em um aprimoramento opcional, o decodificador de áudio de múltiplos canais pode ser configurado para determinar o peso 232 que descreve a contribuição do sinal descorrelacionado 224 na combinação ponderada que depende do sinal descorrelacionado 224. Em outras palavras, o peso 232 pode depender tanto do sinal residual 226 quanto do sinal descorrelacionado 224. Em conformidade, o peso 232 pode ser ainda mais bem adaptado a um sinal de áudio atualmente decodificado sem sobrecarga de sinalização adicional.

**[074]** Como outro aprimoramento opcional, o decodificador de áudio de múltiplos canais pode ser configurado para obter parâmetros de upmix com base na representação codificada 212 e para determinar o peso 232 que descreve a contribuição do sinal descorrelacionado na combinação ponderada que depende dos parâmetros de upmix. Em conformidade, o peso 232 pode adicionalmente depender dos parâmetros de upmix, de modo que uma adaptação ainda melhor do peso 232 possa ser alcançada.

**[075]** Como outro aprimoramento opcional, o decodificador de áudio de múltiplos canais pode ser configurado para determinar o peso que descreve a contribuição do sinal descorrelacionado na combinação ponderada de modo que o peso do sinal descorrelacionado diminua com o aumento da energia do sinal residual. Em conformidade, uma mescla ou desvanecimento pode ser realizada entre uma decodificação que é predominantemente baseada no sinal

descorrelacionado 224 (além de um sinal de downmix 222) e uma decodificação que é predominantemente baseada no sinal residual 226 (além de um sinal de downmix 222).

**[076]** Como outro aprimoramento opcional, o decodificador de áudio de múltiplos canais 200 pode ser configurado para determinar o peso 232 de modo que um peso máximo, que é determinado por um parâmetro de upmix de sinal descorrelacionado (que pode ser incluído ou derivado da representação codificada 210) seja associado ao sinal descorrelacionado 224 se uma energia do sinal residual 226 for zero, e que de modo que um peso zero seja associado ao sinal descorrelacionado 224 se uma energia do sinal residual 226, ponderada com o coeficiente de ponderação de sinal residual (ou um parâmetro de upmix de sinal residual), for maior ou igual a uma energia do sinal descorrelacionado 224, ponderada com o parâmetro de upmix de sinal descorrelacionado. Em conformidade, é possível mesclar (ou desvanecimento) completamente entre uma decodificação baseada no sinal descorrelacionado 224 e uma decodificação baseada no sinal residual 226. Se o sinal residual 226 for julgado como sendo forte o suficiente (por exemplo, quando a energia do sinal residual ponderado for igual ou maior do que a energia do sinal descorrelacionado ponderado 224), a combinação ponderada pode se basear totalmente no sinal residual 226 para refinar o sinal de downmix 222 ao mesmo tempo em que se deixa o sinal descorrelacionado 224 fora de consideração. Nesse caso, uma reconstrução de forma de onda (pelo menos parcial) no lado do decodificador de áudio de múltiplos canais 200 pode ser realizada, visto que a consideração do sinal descorrelacionado 224 tipicamente evita uma reconstrução de forma de onda particularmente boa, enquanto o uso do sinal residual 226 tipicamente permite uma boa reconstrução de forma de onda.

**[077]** Em outro aprimoramento opcional, o decodificador de áudio de múltiplos canais 200 pode ser configurado para computar um valor de energia ponderado de um sinal descorrelacionado, ponderada dependendo de um ou mais parâmetros de upmix de sinal descorrelacionado, e para computar um valor

de energia ponderado do sinal residual, ponderado com o uso de um ou mais parâmetros de upmix de sinal residual. Nesse caso, o decodificador de áudio de múltiplos canais pode ser configurado para determinar um fator que depende do valor de energia ponderado do sinal descorrelacionado e do valor de energia ponderado do sinal residual e para obter um peso que descreve a contribuição do sinal descorrelacionado 224 a um dos sinais de áudio de saída (por exemplo, o primeiro sinal de áudio de saída 212) com base no fator. Desse modo, a determinação de peso 230 pode fornecer valores de ponderação particularmente bem adaptados 232.

**[078]** Em um aprimoramento opcional, o decodificador de áudio de múltiplos canais 200 (ou o determinador de peso 230 desse) pode ser configurado para multiplicar o fator pelo parâmetro de upmix de sinal descorrelacionado (que pode ser incluído na representação codificada 210, ou derivado da representação codificada 210), para obter o peso (ou valor de ponderação) 232 que descreve a contribuição do sinal descorrelacionado 224 a um dos sinais de áudio de saída (por exemplo, o primeiro sinal de áudio de saída 212).

**[079]** Em um aprimoramento opcional, o decodificador de áudio de múltiplos canais (ou o determinador de peso 230 desse) pode ser configurado para computar a energia do sinal descorrelacionado 224, ponderado com o uso de parâmetros de upmix de sinal descorrelacionado (que pode ser incluído na representação codificada 210, ou que pode ser derivado da representação codificada 210), em uma pluralidade de canais de upmix e intervalos de tempo, para obter o valor de energia ponderado do sinal descorrelacionado.

**[080]** Como um aprimoramento opcional adicional, o decodificador de áudio de múltiplos canais 200 pode ser configurado para computar a energia do sinal residual 224, ponderado com o uso de parâmetros de upmix de sinal residual (que pode ser incluído na representação codificada 210 ou que pode ser derivado da representação codificada 210) em relação a uma pluralidade de canais de upmix e intervalos de tempo, para obter o valor de energia ponderado do sinal residual.

**[081]** Como outro aprimoramento opcional, o decodificador de áudio de múltiplos canais 200 (ou o determinador de peso 232 desse) pode ser configurado para computar o fator mencionado acima que depende de uma diferença entre o valor de energia ponderado do sinal descorrelacionado e o valor de energia ponderado do sinal residual. Constatou-se, que tal computação é uma solução eficiente para determinar os valores de ponderação 232.

**[082]** Como um aprimoramento opcional, o decodificador de áudio de múltiplos canais pode ser configurado para computar o fator que depende de uma razão entre uma diferença entre o valor de energia ponderado do sinal descorrelacionado 224 e o valor de energia ponderado do sinal residual 226, e o valor de energia ponderado do sinal descorrelacionado 224. Constatou-se, que tal computação para o fator proporciona bons resultados para mesclar entre um refinamento baseado em sinal predominantemente de descorrelação do sinal de downmix 222 e refinamento baseado em sinal predominantemente residual do sinal de downmix 222.

**[083]** Como um aprimoramento opcional, o decodificador de áudio de múltiplos canais 200 pode ser configurado para determinar pesos que descrevem contribuições dos sinais descorrelacionados a dois ou mais sinais de áudio de saída, como, por exemplo, o primeiro sinal de áudio de saída 212 e o segundo sinal de áudio de saída 214. Nesse caso, o decodificador de áudio de múltiplos canais pode ser configurado para determinar uma contribuição do sinal descorrelacionado 224 ao primeiro sinal de áudio de saída 212 com base no valor de energia ponderado do sinal descorrelacionado 224 e um parâmetro de upmix de sinal descorrelacionado de primeiro canal. Ademais, o decodificador de áudio de múltiplos canais pode ser configurado para determinar uma contribuição do sinal descorrelacionado 224 ao segundo sinal de áudio de saída 214 com base no valor de energia ponderado do sinal descorrelacionado 224 e um parâmetro de upmix de sinal descorrelacionado de segundo canal. Em outras palavras, diferentes parâmetros de upmix de sinal descorrelacionado podem ser usados para fornecer o primeiro sinal de áudio de saída 212 e o segundo sinal

de áudio de saída 214. No entanto, o mesmo valor de energia ponderado do sinal descorrelacionado pode ser usado para determinar a contribuição do sinal descorrelacionado ao primeiro sinal de áudio de saída 212 e a contribuição do sinal descorrelacionado ao segundo sinal de áudio de saída 214. Desse modo, um ajuste eficiente é possível, em que, entretanto, diferentes características dos dois sinais de áudio de saída 212, 214 podem ser consideradas por diferentes parâmetros de upmix de sinal descorrelacionado.

**[084]** Como um aprimoramento opcional, o decodificador de áudio de múltiplos canais 200 pode ser configurado para desabilitar uma contribuição do sinal descorrelacionado 224 à combinação ponderada se uma energia residual (por exemplo, uma energia do sinal residual 226 ou da versão ponderada do sinal residual 226) exceder uma energia descorrelacionada (por exemplo, uma energia do sinal descorrelacionado 224 ou da versão ponderada do sinal descorrelacionado 224).

**[085]** Como um aprimoramento opcional adicional, o decodificador de áudio pode ser configurado para determinar, no sentido da banda, o peso 232 que descreve uma contribuição do sinal descorrelacionado 224 na combinação ponderada que depende de uma determinação, no sentido da banda, de um valor de energia ponderado do sinal residual. Em conformidade, um ajuste fino do decodificador de áudio de múltiplos canais 200 aos sinais a serem decodificados pode ser realizado.

**[086]** Em outro aprimoramento opcional, o decodificador de áudio pode ser configurado para determinar o peso que descreve uma contribuição do sinal descorrelacionado na combinação ponderada para cada quadro do sinal de áudio de saída 212, 214. Em conformidade, uma boa resolução temporal pode ser alcançada.

**[087]** Em um aprimoramento adicionalmente opcional, a determinação do valor de ponderação 232 pode ser realizada de acordo com algumas das equações fornecidas abaixo.

**[088]** Ademais, deve-se observar que o decodificador de áudio de múltiplos

canais 200 pode ser suplementado por qualquer um dos recursos ou funcionalidades descritas no presente documento, também em relação a outras modalidades.

**[089]**     3. Decodificador de áudio de múltiplos canais, de acordo com a Figura 3

**[090]**     A Figura 3 mostra um diagrama de blocos esquemático de um decodificador de áudio de múltiplos canais 300 de acordo com uma modalidade de a invenção. O decodificador de áudio de múltiplos canais 300 é configurado para receber uma representação codificada 310 e para fornecer, com base nesse, dois ou mais sinais de áudio de saída 312, 314. A representação codificada 310 pode, por exemplo, compreender uma representação codificada de um sinal de downmix, uma representação codificada de um ou mais parâmetros espaciais e uma representação codificada de um sinal residual. O decodificador de áudio de múltiplos canais 300 é configurado para obter (pelo menos) um dos sinais de áudio de saída, por exemplo, um primeiro sinal de áudio de saída 312 e/ou um segundo sinal de áudio de saída 314, com base na representação codificada do sinal de downmix, uma pluralidade de parâmetros espaciais codificados e uma representação codificada do sinal residual.

**[091]**     Em particular, o decodificador de áudio de múltiplos canais 300 é configurado para mesclar entre uma criptografia paramétrica e uma criptografia residual que depende do sinal residual (que é incluído, em uma forma codificada, na representação codificada 310). Em outras palavras, o decodificador de áudio de múltiplos canais 300 pode mesclar entre um modo de decodificação em que o fornecimento dos sinais de áudio de saída 312, 314 é realizado com base no sinal de downmix e com o uso de parâmetros espaciais que descrevem uma relação desejada entre os sinais de áudio de saída 312, 314 (por exemplo, uma diferença de nível intercanal desejada ou uma correlação intercanal desejada dos sinais de áudio de saída 312, 314), e um modo de decodificação em que os sinais de áudio de saída 312, 314 são reconstruídos com base no sinal de downmix com o uso do sinal residual. Desse modo, a intensidade (por exemplo,

energia) do sinal residual, que é incluída na representação codificada 310, pode determinar se a decodificação é principalmente (ou exclusivamente) baseada em os parâmetros espaciais (além do sinal de downmix) ou se a decodificação é principalmente (ou exclusivamente) baseada no sinal residual (além do sinal de downmix), ou se um estado intermediário é considerado em que tanto os parâmetros espaciais quanto o sinal residual afetam o refinamento do sinal de downmix, para derivar os sinais de áudio de saída 312, 314 do sinal de downmix.

**[092]** Ademais, o decodificador de áudio de múltiplos canais 300 permite uma decodificação que é bem adaptada ao conteúdo de áudio atual sem alta sobrecarga de sinalização mesclando-se entre a criptografia paramétrica, (em que, tipicamente, um peso comparativamente alto é fornecido a um sinal descorrelacionado quando se fornecem os sinais de áudio de saída 312, 314) e uma criptografia residual (em que, tipicamente, um peso comparativamente baixo é fornecido a um sinal descorrelacionado) que depende do sinal residual.

**[093]** Ademais, deve-se observar que o decodificador de áudio de múltiplos canais 300 se baseia em considerações semelhantes ao decodificador de áudio de múltiplos canais 200 e que aprimoramentos opcionais descritos acima em relação ao decodificador de áudio de múltiplos canais 200 também podem ser aplicados ao decodificador de áudio de múltiplos canais 300.

**[094]** 4. Método para fornecer uma representação codificada de um sinal de áudio de múltiplos canais de acordo com a Figura 4

**[095]** A Figura 4 mostra um fluxograma de um método 400 para fornecer uma representação codificada de um sinal de áudio de múltiplos canais.

**[096]** O método 400 compreende uma etapa 410 de obter um sinal de downmix com base em um sinal de áudio de múltiplos canais. O método 400 também compreende uma etapa 420 de fornecer parâmetros que descrevem dependências entre os canais do sinal de áudio de múltiplos canais. Por exemplo, parâmetros de diferença de nível intercanal e/ou parâmetros de correlação intercanal (ou parâmetros de covariância) podem ser fornecidos, os quais descrevem dependências entre canais do sinal de áudio de múltiplos

canais. O método 400 também compreende uma etapa 430 de fornecer um sinal residual. Ademais, o método compreende uma etapa 440 de variar uma quantidade de sinal residual incluído na representação codificada que depende do sinal de áudio de múltiplos canais.

**[097]** Deve-se observar que o método 400 se baseia nas mesmas considerações que o codificador de áudio 100 de acordo com a Figura 1. Ademais, o método 400 pode ser suplementado por qualquer um dos recursos e funcionalidades descritos no presente documento em relação aos aparelhos da invenção.

**[098]** 5. Método para fornecer pelo menos dois sinais de áudio de saída com base em uma representação codificada de acordo com a Figura 5.

**[099]** A Figura 5 mostra um fluxograma de um método 500 para fornecer pelo menos dois sinais de áudio de saída com base em uma representação codificada. O método 500 compreende determinar 510 um peso que descreve uma contribuição de um sinal descorrelacionado em uma combinação ponderada que depende de um sinal residual. O método 500 também compreende realizar 520 uma combinação ponderada de um sinal de downmix, um sinal descorrelacionado e um sinal residual para obter um dos sinais de áudio de saída.

**[100]** Deve-se observar que o método 500 pode ser suplementado por qualquer um dos recursos e funcionalidades descritos no presente documento em relação aos aparelhos da invenção.

**[101]** 6. Método para fornecer pelo menos dois sinais de áudio de saída com base em uma representação codificada de acordo com a Figura 6.

**[102]** A Figura 6 mostra um fluxograma de um método 600 para fornecer pelo menos dois sinais de áudio de saída com base em uma representação codificada. O método 600 compreende obter 610 um dos sinais de áudio de saída com base em uma representação codificada de um sinal de downmix, uma pluralidade de parâmetros espaciais codificados e uma representação codificada de um sinal residual. Obter 610 um dos sinais de áudio de saída compreende



realizar 620 uma mescla entre uma criptografia paramétrica e uma criptografia residual que depende do sinal residual.

**[103]** Deve-se observar que o método 600 pode ser suplementado por qualquer um dos recursos e funcionalidades descritos no presente documento em relação aos aparelhos da invenção.

**[104]** 7. Modalidades adicionais

**[105]** A seguir, algumas considerações gerais e algumas modalidades adicionais serão descritas.

**[106]** 7.1 Considerações gerais

**[107]** Modalidades de acordo com a invenção são baseadas na ideia de que, em vez de usar uma largura de banda residual fixa, um decodificador (por exemplo, um decodificador de áudio de múltiplos canais) detecta a quantidade de sinal residual transmitido medindo-se its energia por banda para cada quadro (ou, geralmente, pelo menos para uma pluralidade de faixas de frequência e/ou para uma pluralidade de porções temporais). Dependendo dos parâmetros espaciais transmitidos, uma saída descorrelacionada é adicionada quando a energia residual “está ausente”, para alcançar uma quantidade necessária (ou desejada) de descorrelação e energia de saída. Isso permite uma largura de banda residual variável bem como sinais residuais do estilo passa-banda. Por exemplo, é possível apenas usar criptografia residual para bandas tonais. Para se ter a capacidade de usar o downmix simplificado para criptografia paramétrica bem como para criptografia de preservação de forma de onda (que também é designada como criptografia residual), um sinal residual para o downmix simplificado é definido no presente documento.

**[108]** 7.2 Cálculo do sinal residual para o downmix simplificado

**[109]** A seguir, algumas considerações em relação ao cálculo do sinal residual e em relação à construção de sinais de canal de um sinal de áudio de múltiplos canais serão descritas.

**[110]** Em criptografia de áudio e voz unificada (USAC), não há sinal residual definido quando um chamado “downmix simplificado” é usado. Desse modo,

nenhuma criptografia de preservação de forma de onda de modo parcial é possível. No entanto, a seguir, um método para calcular um sinal residual para o chamado “downmix simplificado” será descrito.

**[111]** Ponderações de “downmix simplificado”  $d_1$ ,  $d_2$  são calculadas por banda de fator de escala, enquanto que coeficientes de upmix paramétricos  $u_{d1}$ ,  $u_{d2}$  são calculados por banda de parâmetro. Desse modo, coeficientes  $w_{r1}$ ,  $w_{r2}$ , para calcular o sinal residual não podem ser diretamente computados a partir dos parâmetros espaciais (como é o caso para um MPEG surround clássico), mas podem precisar ser determinados por de fator de escala por banda a partir dos coeficientes de downmix e upmix.

**[112]** Com  $L$ ,  $R$  sendo os canais de entrada e  $D$  sendo o canal de downmix, um sinal residual  $res$  deve preencher as propriedades a seguir:

$$D = d_1 L + d_2 R \quad (1)$$

$$L = u_{d,1} D + u_{r,1} res \quad (2)$$

$$R = u_{d,2} D + u_{r,2} res \quad (3)$$

**[113]** Isso é alcançado calculando-se o residual como

$$res = w_{r,1} L + w_{r,2} R \quad (4)$$

**[114]** com o uso dos pesos de downmix

$$w_{r,1} = \frac{1}{2} \left( \frac{1 - u_{d,1} d_1}{u_{r,1}} - \frac{u_{d,2} d_1}{u_{r,2}} \right) \quad (5)$$

$$w_{r,2} = \frac{1}{2} \left( \frac{1 - u_{d,2} d_2}{u_{r,2}} - \frac{u_{d,1} d_2}{u_{r,1}} \right). \quad (6)$$

**[115]** Os coeficientes de upmix residuais  $u_{r,1}$ ,  $u_{r,2}$  usados pelo decodificador são preferencialmente escolhidos de maneira a garantir decodificação robusta. Visto que o downmix simplificado tem propriedades assimétricas (em oposição a MPEG Surround com pesos fixos) um upmix que depende dos parâmetros espaciais é aplicado, por exemplo, com o uso dos seguintes coeficientes de upmix:

$$u_{r,1} = \max \{u_{d,1}, 0.5\} \quad (7)$$

$$u_{r,2} = -\max \{u_{d,2}, 0.5\} \quad (8)$$

[116] Outra opção é definir os coeficientes de upmix residuais para serem ortogonais aos coeficientes de upmix de sinal de downmix, de modo que:

$$\left\langle \begin{pmatrix} u_{d,1} \\ u_{d,2} \end{pmatrix}, \begin{pmatrix} u_{r,1} \\ u_{r,2} \end{pmatrix} \right\rangle \stackrel{!}{=} 0 \quad (9)$$

[117] Em outras palavras, um decodificador de áudio pode obter o sinal de downmix D com o uso de uma combinação linear de um sinal de canal esquerdo L (primeiro sinal de canal) e o sinal de canal direito R (segundo sinal de canal). De modo semelhante, o sinal residual res é obtido com o uso de uma combinação linear do canal esquerdo L e do sinal de canal direito R (ou, geralmente, de um primeiro sinal de canal e um segundo sinal de canal do sinal de áudio de múltiplos canais).

[118] Pode-se verificar, por exemplo, nas Equações (5) e (6), os pesos de downmix  $w_{r,1}$  e  $w_{r,2}$  para obter o sinal residual res podem ser obtidos quando os pesos de downmix simplificado  $d_1$ ,  $d_2$ , os coeficientes de upmix paramétricos  $u_{d,1}$  e  $u_{d,2}$  e os coeficientes de upmix residuais  $u_{r,1}$  e  $u_{r,2}$  são determinados. Ademais, pode-se verificar que  $u_{r,1}$  e  $u_{r,2}$  podem ser derivados de  $u_{d,1}$  e  $u_{d,2}$  com o uso das equações (7) e (8) ou da equação (9). Os pesos de downmix simplificado  $d_1$  e  $d_2$ , bem como os coeficientes de upmix paramétricos  $u_{d,1}$  e  $u_{d,2}$  podem ser obtidos da maneira usual.

### [119] 7.3 Processo de codificação

[120] A seguir, alguns detalhes em relação ao processo de codificação serão descritos. A codificação pode, por exemplo, ser realizada pelo codificador de áudio de múltiplos canais 100 ou por quaisquer outros programas de computador ou meios apropriados.

[121] Preferencialmente, a quantidade de um residual que é transmitido é determinada por um modelo psicoacústico do codificador (por exemplo, codificador de áudio de múltiplos canais), dependendo do sinal de áudio (por exemplo, dependendo dos sinais de canal do sinal de áudio de múltiplos canais 110) e uma taxa de bits disponível. O sinal residual transmitido pode, por exemplo, ser usado para preservação de forma de onda parcial ou para evitar

cancelamento de sinal causado pelo método de downmix usado (por exemplo, o método de downmix descrito pela equação (1) acima).

**[122]**      7.3.1 Preservação de forma de onda parcial

**[123]**      A seguir, descreve-se como a preservação de forma de onda parcial pode ser alcançada. Por exemplo, o residual calculado (por exemplo, o residual res de acordo com a equação (4)) é transmitido em modo de banda total ou banda limitada para fornecer preservação de forma de onda parcial dentro da largura de banda residual. Partes residuais, que são detectadas como perceptualmente irrelevantes pelo modelo psicoacústico podem, por exemplo, ser quantizadas para zero (por exemplo, ao fornecer a representação codificada 112 com base no sinal residual 126). Isso inclui, mas não se limita à redução da largura de banda residual transmitida no tempo de execução (que pode ser considerado como variando uma quantidade de sinal residual que é incluída na representação codificada). Esse sistema também pode permitir que a exclusão do estilo passa-banda de partes de sinal residual, como energia de sinal ausente, seja reconstruída pelo decodificador (por exemplo, pelo decodificador de áudio de múltiplos canais 200 ou o decodificador de áudio de múltiplos canais 300). Desse modo, por exemplo, a criptografia residual pode ser apenas aplicada a componentes tonais do sinal, preservando suas relações de fase, enquanto que o ruído de fundo pode ser parametricamente criptografado para reduzir a taxa de bits residual. Em outras palavras, o sinal residual 126 pode apenas ser incluído na representação codificada 112 (por exemplo, pelo processamento de sinal residual 130) para bandas de frequência e/ou porções temporais para as quais o sinal de áudio de múltiplos canais 110 (ou pelo menos um dos sinais de canal do sinal de áudio de múltiplos canais 110) é constatado como sendo tonal. Em contrapartida, o sinal residual 126 pode não ser incluído na representação codificada 112 para bandas de frequência e/ou porções temporais para as quais o sinal de áudio de múltiplos canais 110 (ou pelo menos um ou mais sinais de canal do sinal de áudio de múltiplos canais 110) é identificado como sendo similar a ruído. Desse modo, uma quantidade de sinal residual incluído na

representação codificada é variada dependendo do sinal de áudio de múltiplos canais.

**[124]**      7.3.2 Prevenção de cancelamento de sinal em downmix

**[125]**      A seguir, será descrito como um cancelamento de sinal no downmix pode ser prevenido (ou compensado).

**[126]**      Para aplicações de baixa taxa de bits, a criptografia paramétrica (que se baseia predominante ou exclusivamente nos parâmetros 124 que descrevem dependências entre canais do sinal de áudio de múltiplos canais) em vez de criptografia de preservação de forma de onda (que, por exemplo, se baseia predominantemente no sinal residual 126, além do sinal de downmix 122) é aplicada. No presente documento, o sinal residual 126 é apenas usado para compensar cancelamentos de sinal no downmix 122, para minimizar o uso de bits do residual. Desde que nenhum cancelamento de sinal no downmix 122 seja detectado, o sistema é executado em modo paramétrico com o uso de descorrelacionadores (no lado do decodificador de áudio). Quando os cancelamentos de sinal ocorrem, por exemplo, para faseamento de sinais tonais, um sinal residual 126 é transmitido para as partes de sinal desemparelhadas (por exemplo, bandas de frequência e/ou porções temporais). Desse modo, o sinal energia pode ser restaurados pelo decodificador.

**[127]**      7.4 Processo de decodificação

**[128]**      7.4.1 Visão geral

**[129]**      No decodificador (por exemplo, no decodificador de áudio de múltiplos canais 200 ou no decodificador de áudio de múltiplos canais 300), os e sinais de downmix e residuais transmitidos (por exemplo, sinal de downmix 222 ou sinal residual 226) são decodificados por um decodificador de núcleo e alimentados a um decodificador de MPEG surround juntamente com a carga de MPEG surround decodificada. Os coeficientes de upmix residuais para o downmix de MPS clássico são imutáveis, e o coeficiente de upmix residual para o downmix simplificado são definidos nas equações (7) e (8) e/ou (9). Adicionalmente, as saídas de descorrelacionador e seus coeficientes de ponderação são calculados,

como para decodificação paramétrica. O sinal residual e as saídas de descorrelacionador são ponderados e ambos misturados ao sinal de saída. Portanto, fatores de ponderação são determinados medindo-se as energias dos sinais de descorrelacionador e residuais.

**[130]** Em outras palavras, fatores de upmix residuais (ou coeficientes) podem ser determinados medindo-se as energias dos sinais descorrelacionados e residuais.

**[131]** Por exemplo, o sinal de downmix 222 é fornecido com base na representação codificada 210, e o sinal descorrelacionado 224 é derivado do sinal de downmix 222 ou gerado com base em parâmetros incluídos na representação codificada 210 (ou de outro modo). Os coeficientes de upmix residuais podem, por exemplo, ser derivados dos coeficientes de upmix paramétricos  $u_{d,1}$  e  $u_{d,2}$  de acordo com as equações (7) e (8) pelo decodificador, em que os coeficientes de upmix paramétricos  $u_{d,1}$   $u_{d,2}$  podem ser obtidos com base na representação codificada 210, por exemplo, diretamente ou derivando-os de dados espaciais incluídos na representação codificada 210 (por exemplo, a partir de coeficientes de correlação intercanal e coeficientes de diferença de nível intercanal, ou de coeficientes de correlação interobjeto e diferenças de nível interobjeto).

**[132]** Os coeficientes de upmix para a saída de descorrelacionador (ou saídas) podem ser obtidos como para decodificação de MPEG surround convencional. No entanto, os fatores de ponderação para ponderar a saída de descorrelacionador (ou saídas de descorrelacionador) podem ser determinados com base nas energias do sinal residual (e possivelmente também com base nas energias do sinal ou sinais de descorrelacionador) de modo que um peso que descreve uma contribuição do sinal descorrelacionado na combinação ponderada seja determinado dependendo do sinal residual.

**[133]** 7.4.2 Implementação exemplificativa

**[134]** A seguir, uma implementação exemplificativa será descrita em referência à Figura 7. No entanto, deve-se observar que o conceito descrito no

presente documento também pode ser aplicado nos decodificadores de áudio de múltiplos canais 200 ou 300 de acordo com as Figuras 2 e 3.

**[135]** A Figura 7 mostra um diagrama de blocos esquemático (ou diagrama de fluxo) de um decodificador (por exemplo, de um decodificador de áudio de múltiplos canais). O decodificador de acordo com a Figura 7 é designado com 700 em sua totalidade. O decodificador 700 é configurado para receber um fluxo de bits 710 e para fornecer, com base nesse, um primeiro sinal de canal de saída 712 e um segundo sinal de canal de saída 714. O decodificador 700 compreende um decodificador de núcleo 720, que é configurado para receber o fluxo de bits 710 e para fornecer, com base nesse, um sinal de downmix 722, um sinal residual 724 e dados espaciais 726. Por exemplo, o decodificador no núcleo 720 pode fornecer, como o sinal de downmix, uma representação de domínio de tempo ou representação de domínio de transformada (por exemplo, representação de domínio de frequência, representação de domínio de MDCT, representação de domínio de QMF) do sinal de downmix representado pelo fluxo de bits 710. De modo semelhante, o decodificador no núcleo 720 pode fornecer uma representação de domínio de tempo ou representação de domínio de transformada do sinal residual 724, que é representado pelo fluxo de bits 710. Ademais, o decodificador no núcleo 720 pode fornecer um ou mais parâmetros espaciais 726, como, por exemplo, um ou mais parâmetro de correlação intercanal, parâmetros de diferença de nível intercanal ou similares.

**[136]** O decodificador 700 também compreende um descorrelacionador 730, que é configurado para fornecer um sinal descorrelacionado 732 com base no sinal de downmix 722. Qualquer um dos conceitos de descorrelacionamento conhecidos podem ser usados pelo descorrelacionador 730. Ademais, o decodificador 700 também compreende um calculador de coeficiente de upmix 740, que é configurado para receber dados espaciais 726 e para fornecer parâmetros de upmix (por exemplo, parâmetros de upmix  $U_{dmx,1}$ ,  $U_{dmx,2}$ ,  $U_{dec,1}$  e  $U_{dec,2}$ ). Ademais, o decodificador 700 compreende um upmixer 750, que é configurado para aplicar os parâmetros de upmix 742 (também designados como

coeficientes de upmix) que são fornecidos pelo calculador de coeficiente de upmix 740 com base nos dados espaciais 726. Por exemplo, o upmixer 750 pode escalonar o sinal de downmix 722 com o uso de dois coeficientes de upmix de sinal de downmix (por exemplo,  $u_{dmx,1}$ ,  $u_{dmx,2}$ ), para obter duas versões submetidas a upmix 752, 754 do sinal de downmix 722. Ademais, o upmixer 750 também é configurado para aplicar um ou mais parâmetros de upmix (por exemplo, dois parâmetros de upmix) ao sinal descorrelacionado 732 fornecido pelo descorrelacionador 730, para obter uma primeira versão submetida a upmix (escalonada) 756 e uma segunda versão submetida a upmix (escalonada) 758 do sinal descorrelacionado 732. Ademais, o upmixer 750 é configurado para aplicar um ou mais coeficientes de upmix (por exemplo, dois coeficientes de upmix) ao sinal residual 724, para obter uma primeira versão submetida a upmix (escalonada) 760 e uma segunda versão submetida a upmix (escalonada) 762 do sinal residual 724.

**[137]** O decodificador 700 também compreende um calculador de peso 770, que é configurado para medir as energias das versões submetidas a upmix (escalonadas) 756, 758 do sinal descorrelacionado 732 e da versão submetida a upmix (escalonada) 760, 762 do sinal residual 724. Ademais, o calculador de peso 770 é configurado para fornecer um ou mais valores de ponderação 772 a um ponderador 780. O ponderador 780 é configurado para obter uma primeira versão submetida a upmix (escalonada) e ponderada 782 do sinal descorrelacionado 732, uma segunda versão submetida a upmix (escalonada) e ponderada 784 do sinal descorrelacionado 732, uma primeira versão submetida a upmix (escalonada) e ponderada 786 do sinal residual 724 e uma segunda versão submetida a upmix (escalonada) e ponderada 788 do sinal residual 724 com o uso de um ou mais valores de ponderação 772 fornecidos pelo calculador de peso 770. O decodificador também compreende um primeiro adicionador 790, que é configurado para adicionar a primeira versão submetida a upmix (escalonada) 752 do sinal de downmix 720, a primeira versão submetida a upmix (escalonada) e ponderada 782 do sinal descorrelacionado 732 e a primeira



versão submetida a upmix (escalonada) e ponderada 786 do sinal residual 724, para obter o primeiro sinal de canal de saída 712. Ademais, o decodificador compreende um segundo adicionador 792, que é configurado para adicionar a segunda versão submetida a upmix 754 do sinal de downmix 720, a segunda versão submetida a upmix (escalonada) e ponderada 784 do sinal descorrelacionado 732 e a segunda versão submetida a upmix (escalonada) e ponderada 788 do sinal residual 724, para obter o segundo sinal de canal de saída 714.

**[138]** No entanto, deve-se observar que não é necessário que o ponderador 780 pese todos os sinais 756, 758, 760, 762. Por exemplo, em algumas modalidades, pode ser suficiente ponderar apenas os sinais 756, 758, ao mesmo tempo em que se deixam os sinais 760, 762 não afetados (de modo que, efetivamente, os sinais 760, 762 sejam diretamente aplicados aos adicionadores 790, 792). Alternativamente, no entanto, a ponderação dos sinais residuais 760, 762 pode ser variada com o tempo. Por exemplo, os sinais residuais podem ser desvanecidos no início ou desvanecidos no final. Por exemplo, a ponderação (ou os fatores de ponderação) dos sinais descorrelacionados pode ser suavizada com o tempo, e os sinais residuais podem ser desvanecidos no início ou desvanecidos no final de modo correspondente.

**[139]** Ademais, deve-se observar que a ponderação, que é realizada pelo ponderador 780, e o upmix, que é aplicado pelo upmixer 750, também podem ser realizados como uma operação combinada, em que o cálculo do peso pode ser realizado diretamente com o uso do sinal descorrelacionado 732 e o sinal residual 724.

**[140]** A seguir, alguns detalhes adicionais em relação à funcionalidade do decodificador 700 serão descritas.

**[141]** Um modo de criptografia paramétrica e residual combinadas pode, por exemplo, ser sinalizado de um modo compatível semirreverso, por exemplo, sinalizando-se a largura de banda residual de uma banda de parâmetro no fluxo de bits. Desse modo, um decodificador preexistente ainda passará e decodificará

o fluxo de bits ao comutar para decodificação paramétrica acima da primeira banda de parâmetro. Fluxos de bits preexistentes com o uso de uma largura de banda residual de um não conteriam energia residual acima da primeira banda de parâmetro, o que leva a uma decodificação paramétrica no novo decodificador proposto.

**[142]** No entanto, dentro de um sistema de codec de áudio 3D, a criptografia paramétrica e residual combinada pode ser usada em conjunto com outras ferramentas de decodificador de núcleo como um elemento de canal quádruplo, que possibilita que o decodificador explicitamente detecte fluxos de bits preexistentes e os decodifique em modo de criptografia residual limitada por banda regular. Uma largura de banda residual real é preferencialmente não explicitamente sinalizada, conforme é determinado pelo decodificador no tempo de execução. O cálculo dos coeficientes de upmix é estabelecido para o modo paramétrico em vez de um modo de criptografia residual. As energias da saída de descorrelacionador ponderada  $E_{\text{dec}}$  e sinal residual ponderado  $E_{\text{res}}$  são calculadas pela banda híbrida  $hb$  em todos os intervalos de tempo  $ts$  e canais de upmix  $ch$  para cada quadro:

$$E_{\text{dec}}(hb) = \sum_{ch} \sum_{ts} \|u_{\text{dec}}(hb, ts, ch) \cdot x_{\text{dec}}(hb, ts, ch)\| \quad (10)$$

$$E_{\text{res}}(hb) = \sum_{ch} \sum_{ts} \|u_{\text{res}}(hb, ts, ch) \cdot x_{\text{res}}(hb, ts, ch)\| \quad (11)$$

**[143]** No presente documento,  $u_{\text{dec}}$  designa um parâmetro de upmix de sinal descorrelacionado para uma banda de frequência  $hb$ , para um intervalo de tempo  $ts$  e para um canal de upmix  $ch$ ,  $\sum_{ch}$  designa uma soma de canais de upmix, e  $\sum_{ts}$  designa uma soma de intervalos de tempo.  $x_{\text{dec}}$  designa um valor (por exemplo, um valor de domínio de transformada complexa) do sinal descorrelacionado para uma banda de frequência  $hb$ , para um intervalo de tempo  $ts$  e para um canal de upmix  $ch$ .

**[144]** O sinal residual (por exemplo, o sinal residual submetido a upmix 760 ou o sinal residual submetido a upmix 762) é adicionado aos canais de saída

(por exemplo, aos canais de saída 712, 714) com um peso de um. O sinal de descorrelacionador (por exemplo, o sinal de descorrelacionador submetido a upmix 756 ou o sinal de descorrelacionador submetido a upmix 758) pode ser ponderado com um fator  $r$  (por exemplo, pelo ponderador 780) que é calculado como

$$r = \sqrt{\left| \frac{E_{\text{dec}}(\text{hb}) - E_{\text{res}}(\text{hb})}{E_{\text{dec}}(\text{hb})} \right|} \quad (12)$$

(13)

**[145]** em que  $E_{\text{dec}}(\text{hb})$  representa um valor de energia ponderado do sinal descorrelacionado  $x_{\text{dec}}$  para uma banda de frequência  $\text{hb}$ , e em que  $E_{\text{res}}(\text{hb})$  representa um valor de energia ponderado do sinal residual  $x_{\text{res}}$  para uma banda de frequência  $\text{hb}$ .

**[146]** Se nenhum residual (por exemplo, nenhum sinal residual 724) tiver sido transmitido, por exemplo, se  $E_{\text{res}} = 0$ ,  $r$  (o fator que pode ser aplicado pelo ponderador 780, e que pode ser considerado como um valor de ponderação 772) se torna 1, que é equivalente a uma decodificação puramente paramétrica. Se a energia residual (por exemplo, a energia do sinal residual submetido a upmix 760 e/ou do sinal residual submetido a upmix 762) exceder a energia do descorrelacionador (por exemplo, a energia do sinal descorrelacionado submetido a upmix 756 ou do sinal descorrelacionado submetido a upmix 758), por exemplo, se  $E_{\text{res}} > E_{\text{dec}}$ , o fator  $r$  pode ser estabelecido como zero, desabilitando, desse modo, o descorrelacionador e habilitando parcialmente a decodificação de preservação de forma de onda (que pode ser considerada como criptografia residual). No processo de upmix, a saída de descorrelacionador ponderada (por exemplo, sinais 782 e 784) e o sinal residual (por exemplo, sinais 786, 788 ou sinais 760, 762) são ambos adicionados aos canais de saída (por exemplo, sinais 712, 714).

**[147]** Em conclusão, isso leva a uma regra de upmix em forma de matriz

$$\begin{pmatrix} \text{ch}_1 \\ \text{ch}_2 \end{pmatrix} = \begin{bmatrix} u_{\text{dmx},1} & r \cdot u_{\text{dec},1} & \max\{u_{\text{dmx},1}, 0,5\} \\ u_{\text{dmx},2} & r \cdot u_{\text{dec},2} & -\max\{u_{\text{dmx},2}, 0,5\} \end{bmatrix} \cdot \begin{pmatrix} x_{\text{dmx}} \\ x_{\text{dec}} \\ x_{\text{res}} \end{pmatrix} \quad (14)$$

**[148]** em que  $ch1$  representa uma ou mais amostras de domínio de tempo ou amostras de domínio de transformada de um primeiro sinal de áudio de saída, em que  $ch2$  representa uma ou mais amostras de domínio de tempo ou amostras de domínio de transformada de um segundo sinal de áudio de saída, em que  $x_{dmx}$  representa uma ou mais amostras de domínio de tempo ou amostras de domínio de transformada de um sinal de downmix, em que  $x_{dec}$  representa uma ou mais amostras de domínio de tempo ou amostras de domínio de transformada de um sinal descorrelacionado, em que  $x_{res}$  representa uma ou mais amostras de domínio de tempo ou amostras de domínio de transformada de um sinal residual, em que  $u_{dmx,1}$  representa um parâmetro de upmix de sinal de downmix para o primeiro sinal de áudio de saída, em que  $u_{dmx,2}$  representa um parâmetro de upmix de sinal de downmix para o segundo sinal de áudio de saída, em que  $u_{dec,1}$  representa um parâmetro de upmix de sinal descorrelacionado para o primeiro sinal de áudio de saída, em que  $u_{dec,2}$  representa um parâmetro de upmix de sinal descorrelacionado para o segundo sinal de áudio de saída, em que  $\max$  representa um operador máximo, e em que  $r$  representa um fator que descreve uma ponderação do sinal descorrelacionado que depende do sinal residual.

**[149]** Os coeficientes de upmix  $U_{dmx,1}$ ,  $U_{dmx,2}$ ,  $U_{dec,1}$ ,  $U_{dec,2}$  são calculados como para o modo paramétrico MPS dois-um-dois (2-1-2). Para detalhes, faz-se referência à norma mencionada acima do conceito de MPEG surround.

**[150]** Em resumo, uma modalidade de acordo com a invenção proporciona um conceito para fornecer sinais de canal de saída com base em um sinal de downmix, um sinal residual e dados espaciais, em que uma ponderação do sinal descorrelacionado é ajustada de modo flexível sem qualquer sobrecarga de sinalização significativa.

**[151]** 7.5 Alternativas de implementação

**[152]** Embora alguns aspectos tenham sido descritos no contexto de um aparelho, é claro que esses aspectos também representam uma descrição do método correspondente, em que um bloco ou dispositivo corresponde a uma

etapa do método ou um recurso de uma etapa do método. De modo análogo, aspectos descritos no contexto de uma etapa do método também representam uma descrição de um bloco ou item ou recurso correspondente de um aparelho correspondente. Algumas ou todas as etapas do método podem ser executadas por (ou com o uso de) um aparelho de hardware, como, por exemplo, um microprocessador, um computador programável ou um circuito eletrônico. Em algumas modalidades, alguma ou mais das etapas do método mais importantes podem ser executadas por tal aparelho.

**[153]** O sinal de áudio codificado da invenção pode ser armazenado em um meio de armazenamento digital ou pode ser transmitido em um meio de transmissão tal como um meio de transmissão sem fio ou um meio de transmissão com fio tal como a Internet.

**[154]** Dependendo de determinadas exigências de implementação, modalidades da invenção podem ser implementadas em hardware ou em software. A implementação pode ser realizada com o uso de um meio de armazenamento digital, por exemplo, um disco flexível, um DVD, um Blu-Ray, um CD, uma ROM, uma PROM, uma EPROM, uma EEPROM ou uma memória FLASH, que tem sinais de controle eletronicamente legíveis armazenados nesse, que cooperam (ou tenham capacidade de cooperar) com um sistema de computador programável de modo que o respectivo método seja realizado. Portanto, o meio de armazenamento digital pode ser legível por computador.

**[155]** Algumas modalidades de acordo com a invenção compreendem uma portadora de dados que tem sinais de controle eletronicamente legíveis, com capacidade de cooperar com um sistema de computador programável, de modo que um dos métodos descritos no presente documento seja realizado.

**[156]** Geralmente, modalidades da presente invenção podem ser implementadas como um produto de programa de computador com um código de programa, sendo que o código de programa é operável para realizar um dos métodos quando o produto de programa de computador é executado em um computador. O código de programa pode, por exemplo, ser armazenado em uma

portadora legível por máquina.

**[157]** Outras modalidades compreendem o programa de computador para realizar um dos métodos descritos no presente documento, armazenados em uma portadora legível por máquina.

**[158]** Em outras palavras, uma modalidade do método da invenção é, portanto, um programa de computador que tem um código de programa para realizar um dos métodos descritos no presente documento, quando o programa de computador é executado em um computador.

**[159]** Uma modalidade adicional dos métodos da invenção é, portanto, uma portadora de dados (ou um meio de armazenamento digital, ou um meio legível por computador) que compreende, gravado nesse, o programa de computador para realizar um dos métodos descritos no presente documento. A portadora de dados, o meio de armazenamento digital ou o meio gravado são tipicamente tangíveis e/ou não transitórios.

**[160]** Uma modalidade adicional do método da invenção é, portanto, um fluxo de dados ou uma sequência de sinais que representa o programa de computador para realizar um dos métodos descritos no presente documento. O fluxo de dados ou a sequência de sinais pode, por exemplo, ser configurado para ser transferido por meio de uma conexão de comunicação de dados, por exemplo, por meio da Internet.

**[161]** Uma modalidade adicional compreende um meio de processamento, por exemplo, um computador, ou um dispositivo de lógica programável, configurado ou adaptado para realizar um dos métodos descritos no presente documento.

**[162]** Uma modalidade adicional compreende a computador que tem instalado nesse o programa de computador para realizar um dos métodos descritos no presente documento.

**[163]** Uma modalidade adicional de acordo com a invenção compreende um aparelho ou um sistema configurado para transferir (por exemplo, eletrônica ou opticamente) um programa de computador para realizar um dos métodos

descritos no presente documento para um receptor. O receptor pode, por exemplo, ser um computador, um dispositivo móvel, um dispositivo de memória ou similares. O aparelho ou sistema pode, por exemplo, compreender um servidor de arquivos para transferir o programa de computador para o receptor.

**[164]** Em algumas modalidades, um dispositivo de lógica programável (por exemplo, um arranjo de portas programável em campo) pode ser usado para realizar algumas ou todas as funcionalidades dos métodos descritos no presente documento. Em algumas modalidades, um arranjo de portas programável em campo pode cooperar com um microprocessador a fim de realizar um dos métodos descritos no presente documento. Geralmente, os métodos são preferencialmente realizados por qualquer aparelho de hardware.

**[165]** As modalidades acima descritas são meramente ilustrativas para os princípios da presente invenção. Entende-se que modificações e variações das disposições e dos detalhes descritos no presente documento estarão claras para outros versados na técnica. Pretende-se, portanto, que sejam limitadas apenas pelo escopo das reivindicações da patente anexas e não pelos detalhes específicos apresentados a título de descrição e explicação das modalidades no presente documento.

**[166]** 7.6 Modalidade adicional

**[167]** A seguir, outra modalidade de acordo com a invenção será descrita em referência à Figura 8, que mostra um diagrama de blocos esquemático de um chamado Decodificador Residual Híbrido.

**[168]** O Decodificador Residual Híbrido 800 de acordo com a Figura 8 é muito semelhante ao decodificador 700 de acordo com a Figura 7, de modo que se faça referência às explicações acima. No entanto, no Decodificador Residual Híbrido 800, uma ponderação adicional (além da aplicação dos parâmetros de upmix) é apenas aplicada aos sinais descorrelacionados submetidos a upmix (que correspondem aos sinais 756, 758 no decodificador 700), mas não aos sinais residuais submetidos a upmix (que correspondem aos sinais 760, 762 no decodificador 700). Desse modo, o ponderador no Decodificador Residual

Híbrido 800 é, de algum modo, mais simples do que o ponderador no decodificador 700, mas está em concordância, por exemplo, com a ponderação de acordo com a equação (14).

**[169]** A seguir, a Decodificação Residual e Paramétrica combinada (Criptografia Residual Híbrida) de acordo com a Figura 8 será explicada em mais alguns detalhes.

**[170]** No entanto, primeiramente, uma visão geral será fornecida.

**[171]** Além do uso da criptografia residual ou de upmix mono para estéreo com base em descorrelacionador conforme descrito em ISO/IEC 23003-3, subitem 7.11.1, a Criptografia Residual Híbrida permite uma combinação dependente de sinal de ambos os modos. O sinal residual e saída de descorrelacionador são mesclados entre si, com o uso de fatores de ponderação dependentes de tempo e frequência dependendo das energias de sinal e dos parâmetros espaciais, conforme ilustrado na Figura 8.

**[172]** A seguir, o processo de decodificação será descrito.

**[173]** O modo de Criptografia Residual Híbrida é indicado pelos elementos de sintaxe `bsResidualCoding == 1` e `bsResidualBands == 1` em `Mps212Config()`. Em outras palavras, o uso da criptografia Residual Híbrida pode ser sinalizado com o uso de um elemento do fluxo de bits da representação codificada. O cálculo da matriz de mistura M2 é realizado como se `bsResidualCoding == 0`, seguindo o cálculo em ISO/IEC 23003-3, subitem 7.11.2.3. A matriz  $R_2^{l,m}$  para a parte com base em descorrelacionador é definida como

$$R_2^{l,m} = \begin{bmatrix} H11_{OTT}^{l,m} & H12_{OTT}^{l,m} \\ H21_{OTT}^{l,m} & H22_{OTT}^{l,m} \end{bmatrix}$$

**[174]**

**[175]** O processo de upmix é dividido em Downmix, saída de descorrelacionador e residual. O Downmix submetido a upmix  $u_{dmx}$  é calculado com o uso de:

$$R_{2, dmx}^{l,m} = \begin{bmatrix} H11_{OTT}^{l,m} & 0 \\ H21_{OTT}^{l,m} & 0 \end{bmatrix}$$

**[176]**



[177] A saída de descorrelacionador submetida a upmix  $u_{dec}$  é calculada com o uso de:

[178] 
$$R_{2,dec}^{l,m} = \begin{bmatrix} 0 & H12_{OTT}^{l,m} \\ 0 & H22_{OTT}^{l,m} \end{bmatrix}$$

[179] O sinal residual submetido a upmix  $u_{res}$  é calculado com o uso de:

[180] 
$$R_{2,res}^{l,m} = \begin{bmatrix} 0 & H12_{RES}^{l,m} \\ 0 & H22_{RES}^{l,m} \end{bmatrix} = \begin{bmatrix} 0 & \max\{0,5, H11_{OTT}^{l,m}\} \\ 0 & -\max\{0,5, H21_{OTT}^{l,m}\} \end{bmatrix}$$

[181] As energias do sinal residual submetido a upmix  $E_{res}$  e da saída de descorrelacionador submetida a upmix  $E_{dec}$  são calculadas por banda híbrida como soma de canais de saída  $ch$  e todos os intervalos de tempo  $ts$  quanto de um quadro as:

[182] 
$$E_{res} = \sum_{ch} \sum_{ts} \|u_{res}(ch, ts)\|$$

[183] 
$$E_{dec} = \sum_{ch} \sum_{ts} \|u_{dec}(ch, ts)\|$$

[184] A saída de descorrelacionador submetida a upmix é ponderada com o uso de um fator de ponderação  $r_{dec}$  calculado para cada banda híbrida por quadro como:

$$r_{dec} = \begin{cases} 0 & se \quad E_{res} > E_{dec} \\ 1 & se \quad E_{res} < \varepsilon \\ \sqrt{\frac{E_{dec} - E_{res} + \varepsilon}{E_{dec} + \varepsilon}} & ou \end{cases}$$

[185] sendo que  $\varepsilon$  é um número pequeno para evitar a divisão por zero (por exemplo,  $\varepsilon = 1e-9$ , ou  $0 < \varepsilon \leq 1e-5$ ). No entanto, em algumas modalidades,  $\varepsilon$  pode ser estabelecido como zero (substituindo " $E_{res} < \varepsilon$ " por " $E_{res} = 0$ ").

**[186]** Todos os três sinais de upmix são adicionados para formar o sinal de saída decodificado.

**[187]** 8. Conclusões

**[188]** Em conclusão, modalidades de acordo com a invenção produzem uma criptografia paramétrica e residual combinada.

**[189]** A presente invenção proporciona um método para uma combinação dependente de sinal de criptografia residual e paramétrica para a criptografia de estéreo conjunto, que se baseia na ferramenta para estéreo unificado de USAC. Em vez de usar uma largura de banda residual fixa, a quantidade de residual transmitida é determinada como dependente de sinal por um codificador, com variação de tempo e frequência. No lado do decodificador, a quantidade necessária de descorrelação entre os canais de saída é gerada misturando-se o sinal residual e saída de descorrelacionador. Desse modo, um sistema de criptografia/decodificação correspondente tem capacidade para mesclar entre criptografia residual de preservação de forma de onda e criptografia paramétrica total, dependendo do sinal codificado.

**[190]** Modalidades de acordo com a invenção superam as soluções convencionais. Por exemplo, em USAC, um sistema de MPEG surround dois-um-dois (2-1-2) é usado para criptografia estéreo paramétrica, ou estéreo unificado, que transmite a sinal residual de largura de banda total ou limitado por banda para a preservação de forma de onda parcial. Se um residual limitado por banda é transmitido, o upmix paramétrico com o uso de descorrelacionadores é aplicado acima da largura de banda residual. A desvantagem desse método é que a largura de banda residual é estabelecida como um valor fixo na inicialização do codificador.

**[191]** Em contrapartida, modalidades de acordo com a invenção permitem uma adaptação dependente de sinal da largura de banda residual ou a comutação para criptografia paramétrica. Ademais, se o processo de downmix em modo de criptografia paramétrica produz cancelamentos de sinal para relações de fase condicionadas equivocadamente, modalidades de acordo com

a invenção permitem reconstruir as partes de sinal ausentes (por exemplo, fornecendo um sinal residual apropriado). Deve-se observar que o método de downmix simplificado produz menos cancelamentos de sinal do que o downmix MPS clássico para criptografia paramétrica. No entanto, embora o downmix simplificado convencional não possa ser usado para a preservação de forma de onda parcial, visto que nenhum sinal residual é definido em USAC, modalidades de acordo com a invenção permitem uma reconstrução de forma de onda (por exemplo, uma reconstrução de forma de onda parcial seletiva para porções de sinal em que a reconstrução de forma de onda parcial parece ser importante).

**[192]** Ainda em conclusão, modalidades de acordo com a invenção produzem um aparelho, um método ou um programa de computador para codificação ou decodificação de áudio conforme descrito no presente documento.

## REIVINDICAÇÕES

1. Decodificador de áudio de múltiplos canais (200; 300; 700; 800) para fornecer pelo menos dois sinais de áudio de saída (212, 214; 312, 314; 712, 714) com base em uma representação codificada (210; 310; 710),

em que o decodificador de áudio de múltiplos canais é configurado para realizar uma combinação ponderada (220; 780, 790, 792) de um sinal de downmix (222; 752, 754), um sinal descorrelacionado (224; 756, 758) e um sinal residual (226; 760, 762; res) para obter um dos sinais de áudio de saída (212, 214; 712, 714),

### **caracterizado por:**

o decodificador de áudio de múltiplos canais ser configurado para determinar um peso (232;  $r$ ;  $r_{dec}$ ) que descreve uma contribuição do sinal descorrelacionado na combinação ponderada que depende de ambos o sinal residual e do sinal descorrelacionado;

o decodificador de áudio ser configurado para determinar em banda o peso (232;  $r$ ;  $r_{dec}$ ) que descreve uma contribuição do sinal descorrelacionado na combinação ponderada que depende de uma determinação em banda dos valores de energia ponderada do sinal residual.

2. Decodificador de áudio de múltiplos canais, de acordo com a reivindicação 1, sendo que o decodificador de áudio de múltiplos canais é **caracterizado por** ser configurado para obter parâmetros de upmix ( $U_{dmx,1}$ ,  $U_{dmx,2}$ ,  $U_{dec,1}$ ,  $U_{dec,2}$ ,  $U_{r,1}$ ,  $U_{r,2}$ ) com base na representação codificada, e para determinar o peso (232;  $r$ ;  $r_{dec}$ ) que descreve a contribuição do sinal descorrelacionado na combinação ponderada que depende dos parâmetros de upmix.

3. Decodificador de áudio de múltiplos canais, de acordo com a reivindicação 1 ou 2, sendo que o decodificador de áudio de múltiplos canais é **caracterizado por** ser configurado para determinar o peso (232;  $r$ ;  $r_{dec}$ ) que descreve a contribuição do sinal descorrelacionado na combinação ponderada, de modo que o peso do sinal descorrelacionado diminua com o aumento da

energia do sinal residual.

4. Decodificador de áudio de múltiplos canais, de acordo com qualquer uma das reivindicações 1 a 3, sendo que o decodificador de áudio de múltiplos canais é **caracterizado por** ser configurado para determinar o peso ( $232; r; r_{dec}$ ) que descreve a contribuição do sinal descorrelacionado na combinação ponderada, de modo que um peso máximo, que é determinado por um parâmetro de upmix de sinal descorrelacionado ( $U_{dec,1}, U_{dec,2}; U_{dec}(hb,ts,ch); U_{dec}(ch,ts)$ ), seja associado ao sinal descorrelacionado se uma energia do sinal residual for zero, e de modo que um peso zero seja associado ao sinal descorrelacionado se uma energia do sinal residual ponderada com um coeficiente de ponderação de sinal residual ( $U_{r,1}, U_{r,2}; U_{res}(hb,ts,ch); U_{res}(ch,ts)$ ) for maior ou igual a uma energia do sinal descorrelacionado, ponderada com o parâmetro de upmix de sinal descorrelacionado.

5. Decodificador de áudio de múltiplos canais, de acordo com qualquer uma das reivindicações 1 a 4, sendo que o decodificador de áudio de múltiplos canais é **caracterizado por** ser configurado para computar um valor de energia ponderado ( $E_{dec}(hb); E_{dec}$ ) do sinal descorrelacionado, ponderado dependendo de um ou mais parâmetros de upmix de sinal descorrelacionado, e para computar um valor de energia ponderado ( $E_{res}(hb); E_{res}$ ) do sinal residual, ponderado com o uso de um ou mais parâmetros de upmix de sinal residual, para determinar um fator ( $r, r_{dec}$ ) que depende do valor de energia ponderado do sinal descorrelacionado e do valor de energia ponderado do sinal residual, e para obter o peso que descreve a contribuição do sinal descorrelacionado a um dos sinais de áudio de saída com base no fator ou para usar o fator como o peso que descreve a contribuição do sinal descorrelacionado a um dos sinais de áudio de saída.

6. Decodificador de áudio de múltiplos canais, de acordo com a reivindicação 5, sendo que o decodificador de áudio de múltiplos canais é **caracterizado por** ser configurado para multiplicar o fator ( $r$ ) por um parâmetro de upmix de sinal descorrelacionado ( $U_{dec,1}, U_{dec,2}; U_{dec}(hb,ts,ch); U_{dec}(ch,ts)$ ) para

obter o peso que descreve a contribuição do sinal descorrelacionado a um dos sinais de áudio de saída.

7. Decodificador de áudio de múltiplos canais, de acordo com a reivindicação 5 ou 6, sendo que o decodificador de áudio de múltiplos canais é **caracterizado por** ser configurado para computar a energia do sinal descorrelacionado, ponderado com o uso de parâmetros de upmix de sinal descorrelacionado, em uma pluralidade de canais de upmix (ch) e intervalos de tempo (ts), para obter o valor de energia ponderado ( $E_{dec}(hb)$ ;  $E_{dec}$ ) do sinal descorrelacionado.

8. Decodificador de áudio de múltiplos canais, de acordo com qualquer uma das reivindicações 5 a 7, sendo que o decodificador de áudio de múltiplos canais é **caracterizado por** ser configurado para computar a energia do sinal residual, ponderado com o uso de parâmetros de upmix de sinal residual, em uma pluralidade de canais de upmix (ch) e intervalos de tempo (ts), para obter o valor de energia ponderado ( $E_{res}(hb)$ ;  $E_{res}$ ) do sinal residual.

9. Decodificador de áudio de múltiplos canais, de acordo com qualquer uma das reivindicações 5 a 8, sendo que o decodificador de áudio de múltiplos canais é **caracterizado por** ser configurado para computar o fator ( $r$ ;  $r_{dec}$  que depende de uma diferença entre o valor de energia ponderado ( $E_{dec}(hb)$ ;  $E_{dec}$ ) do sinal descorrelacionado e o valor de energia ponderado ( $E_{res}(hb)$ ;  $E_{res}$ ) do sinal residual.

10. Decodificador de áudio de múltiplos canais, de acordo com a reivindicação 9, sendo que o decodificador de áudio de múltiplos canais é **caracterizado por** ser configurado para computar o fator ( $r$ ;  $r_{dec}$ ) que depende de uma razão entre

- uma diferença entre o valor de energia ponderado ( $E_{dec}(hb)$ ;  $E_{dec}$ ) do sinal descorrelacionado e o valor de energia ponderado do sinal residual ( $E_{res}(hb)$ ;  $E_{res}$ ), e
- o valor de energia ponderado ( $E_{dec}(hb)$ ;  $E_{dec}$ ) do sinal descorrelacionado.

11. Decodificador de áudio de múltiplos canais, de acordo com qualquer uma das reivindicações 5 a 10, sendo que o decodificador de áudio de múltiplos canais é **caracterizado por** ser configurado para determinar pesos que descrevem contribuições do sinal descorrelacionado a dois ou mais sinais de áudio de saída,

em que o decodificador de áudio de múltiplos canais é configurado para determinar uma contribuição do sinal descorrelacionado a um primeiro sinal de áudio de saída com base no valor de energia ponderado ( $E_{\text{dec}}(\text{hb}); E_{\text{dec}}$ ) do sinal descorrelacionado e um parâmetro de upmix de sinal descorrelacionado de primeiro canal ( $u_{\text{dec},1}$ ), e

em que o decodificador de áudio de múltiplos canais é configurado para determinar uma contribuição do sinal descorrelacionado a um segundo canal de áudio de saída com base no valor de energia ponderado ( $E_{\text{dec}}(\text{hb}); E_{\text{dec}}$ ) do sinal descorrelacionado e a parâmetro de upmix de sinal descorrelacionado de segundo canal ( $u_{\text{dec},2}$ ).

12. Decodificador de áudio de múltiplos canais, de acordo com qualquer uma das reivindicações 1 a 11, sendo que o decodificador de áudio de múltiplos canais é **caracterizado por** ser configurado para desabilitar uma contribuição do sinal descorrelacionado à combinação ponderada se uma energia residual ( $E_{\text{res}}(\text{hb}); E_{\text{res}}$ ) exceder uma energia de descorrelacionador ( $E_{\text{dec}}(\text{hb}); E_{\text{dec}}$ ).

13. Decodificador de áudio de múltiplos canais, de acordo com qualquer uma das reivindicações 1 a 12, sendo que o decodificador de áudio de múltiplos canais é **caracterizado por** ser configurado para computar dois sinais de áudio de saída ch1, ch2 de acordo com

$$\begin{pmatrix} \text{ch}_1 \\ \text{ch}_2 \end{pmatrix} = \begin{bmatrix} u_{\text{dmx},1} & r \cdot u_{\text{dec},1} & \max \{u_{\text{dmx},1}, 0,5\} \\ u_{\text{dmx},2} & r \cdot u_{\text{dec},2} & -\max \{u_{\text{dmx},2}, 0,5\} \end{bmatrix} \cdot \begin{pmatrix} x_{\text{dmx}} \\ x_{\text{dec}} \\ x_{\text{res}} \end{pmatrix}$$

em que ch1 representa uma ou mais amostras de domínio de tempo ou amostras de domínio de transformada de um primeiro sinal de áudio de saída, em que ch2 representa uma ou mais amostras de domínio de tempo

ou amostras de domínio de transformada de um segundo sinal de áudio de saída,

em que  $x_{dmx}$  representa uma ou mais amostras de domínio de tempo ou amostras de domínio de transformada de um sinal de downmix;

em que  $x_{dec}$  representa uma ou mais amostras de domínio de tempo ou amostras de domínio de transformada de um sinal descorrelacionado;

em que  $x_{res}$  representa uma ou mais amostras de domínio de tempo ou amostras de domínio de transformada de um sinal residual;

em que  $u_{dmx,1}$  representa um parâmetro de upmix de sinal de downmix para o primeiro sinal de áudio de saída;

em que  $u_{dmx,2}$  representa um parâmetro de upmix de sinal de downmix para o segundo sinal de áudio de saída;

em que  $u_{dec,1}$  representa um parâmetro de upmix de sinal descorrelacionado para o primeiro sinal de áudio de saída;

em que  $u_{dec,2}$  representa um parâmetro de upmix de sinal descorrelacionado para o segundo sinal de áudio de saída;

em que  $\max$  representa um operador máximo; e

em que  $r$  representa um fator que descreve uma ponderação do sinal descorrelacionado que depende do sinal residual.

14. Decodificador de áudio de múltiplos canais, de acordo com a reivindicação 13, sendo que o decodificador de áudio de múltiplos canais é **caracterizado por** ser configurado para computar o fator  $r$  de acordo com

$$r = \sqrt{\left| \frac{E_{dec}(hb) - E_{res}(hb)}{E_{dec}(hb)} \right|}$$

ou de acordo com

$$r = \begin{cases} 0 & \text{if } E_{res} > E_{dec} \\ 1 & \text{if } E_{res} < \varepsilon \\ \sqrt{\frac{E_{dec} - E_{res} + \varepsilon}{E_{dec} + \varepsilon}} & \text{else} \end{cases}$$



em que  $E_{\text{dec}}(\text{hb})$  ou  $E_{\text{dec}}$  representa um valor de energia ponderado do sinal descorrelacionado  $x_{\text{dec}}$  para uma banda de frequência  $\text{hb}$ , e

em que  $E_{\text{res}}(\text{hb})$  ou  $E_{\text{res}}$  representa um valor de energia ponderado do sinal residual  $x_{\text{res}}$  para uma banda de frequência  $\text{hb}$ .

15. Decodificador de áudio de múltiplos canais, de acordo com a reivindicação 14, sendo que o decodificador de áudio de múltiplos canais é **caracterizado por** ser configurado para computar o valor de energia ponderado do sinal descorrelacionado de acordo com

$$E_{\text{dec}}(\text{hb}) = \sum_{\text{ch}} \sum_{\text{ts}} \|u_{\text{dec}}(\text{hb}, \text{ts}, \text{ch}) \cdot x_{\text{dec}}(\text{hb}, \text{ts}, \text{ch})\|$$

em que  $u_{\text{dec}}$  designa um parâmetro de upmix de sinal descorrelacionado para uma banda de frequência  $\text{hb}$ , para um intervalo de tempo  $\text{ts}$  e para um canal de upmix  $\text{ch}$ ,

em que  $x_{\text{dec}}$  representa uma amostra de domínio de tempo ou amostra de domínio de transformada de um sinal descorrelacionado para uma banda de frequência  $\text{hb}$ , para um intervalo de tempo  $\text{ts}$  e para um canal de upmix  $\text{ch}$ ,

em que  $\sum_{\text{ch}}$  designa uma soma de canais de upmix  $\text{ch}$ , e

em que  $\sum_{\text{ts}}$  designa uma soma de intervalos de tempo  $\text{ts}$ ,

em que  $\| \cdot \|$  designa um operador normal,

em que o decodificador de áudio de múltiplos canais é configurado para computar o valor de energia ponderado do sinal residual de acordo com

$$E_{\text{res}}(\text{hb}) = \sum_{\text{ch}} \sum_{\text{ts}} \|u_{\text{res}}(\text{hb}, \text{ts}, \text{ch}) \cdot x_{\text{res}}(\text{hb}, \text{ts}, \text{ch})\|$$

em que  $u_{\text{res}}$  designa um sinal residual parâmetro de upmix para uma banda de frequência  $\text{hb}$ , para um intervalo de tempo  $\text{ts}$  e para um canal de upmix  $\text{ch}$ ,

em que  $x_{\text{res}}$  representa uma amostra de domínio de tempo ou amostra de domínio de transformada de um sinal descorrelacionado para uma banda de frequência  $\text{hb}$ , para um intervalo de tempo  $\text{ts}$  e para um canal de upmix  $\text{ch}$ .

16. Decodificador de áudio, de acordo com qualquer uma das

reivindicações 1 a 15, sendo que o decodificador de áudio de múltiplos canais é **caracterizado por** ser configurado para ajustar de modo variado um peso que descreve uma contribuição do sinal residual na combinação ponderada.

17. Método (500) para fornecer pelo menos dois sinais de áudio de saída com base em uma representação codificada, que compreende:

realizar (520) uma combinação ponderada de um sinal de downmix, um sinal descorrelacionado e um sinal residual para obter um dos sinais de áudio de saída,

**caracterizado por**

um peso que descreve uma contribuição do sinal descorrelacionado na combinação ponderada ser determinado (510) dependendo de ambos o sinal residual e do sinal descorrelacionado;

o peso (232;  $r$ ;  $r_{dec}$ ) que descreve uma contribuição do sinal descorrelacionado na combinação ponderada ser determinado em banda que depende de uma determinação em banda dos valores de energia ponderada do sinal residual.

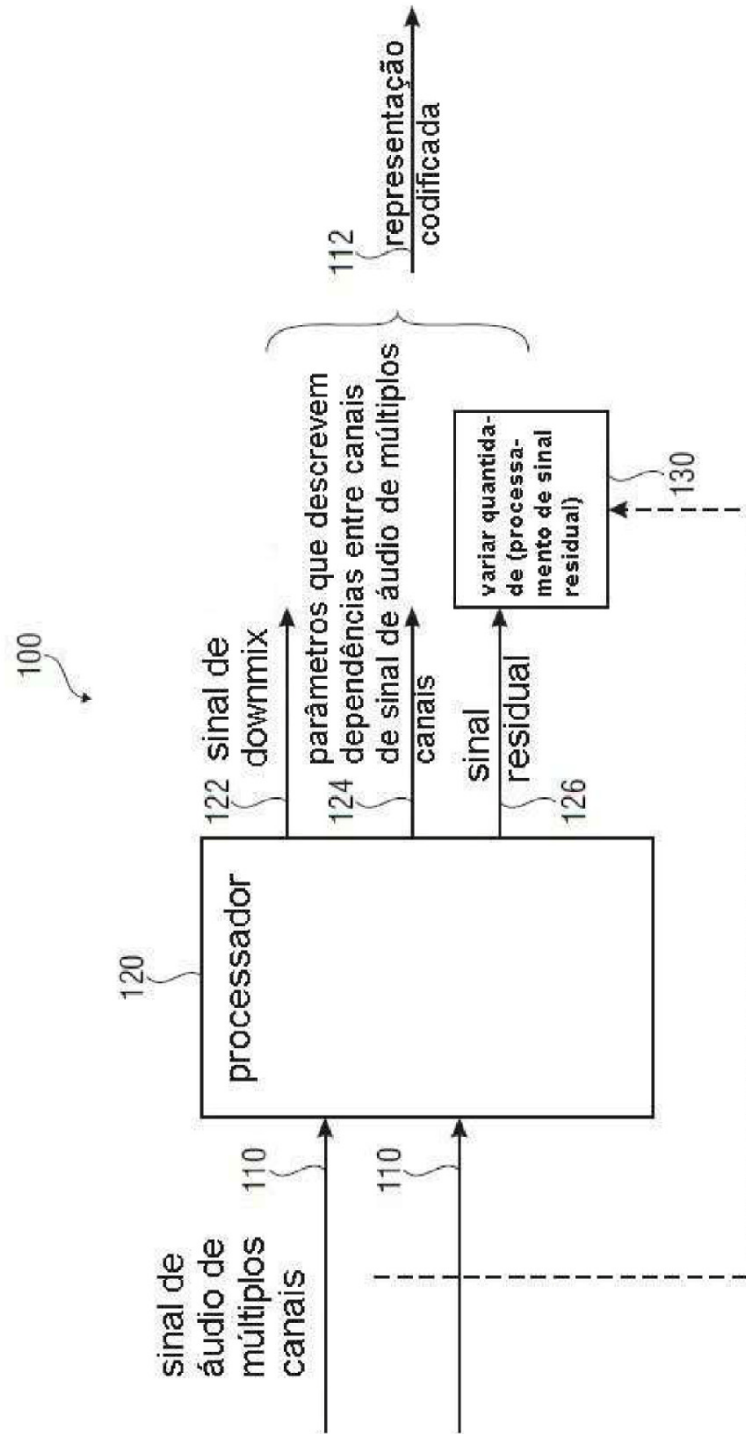


FIG. 1

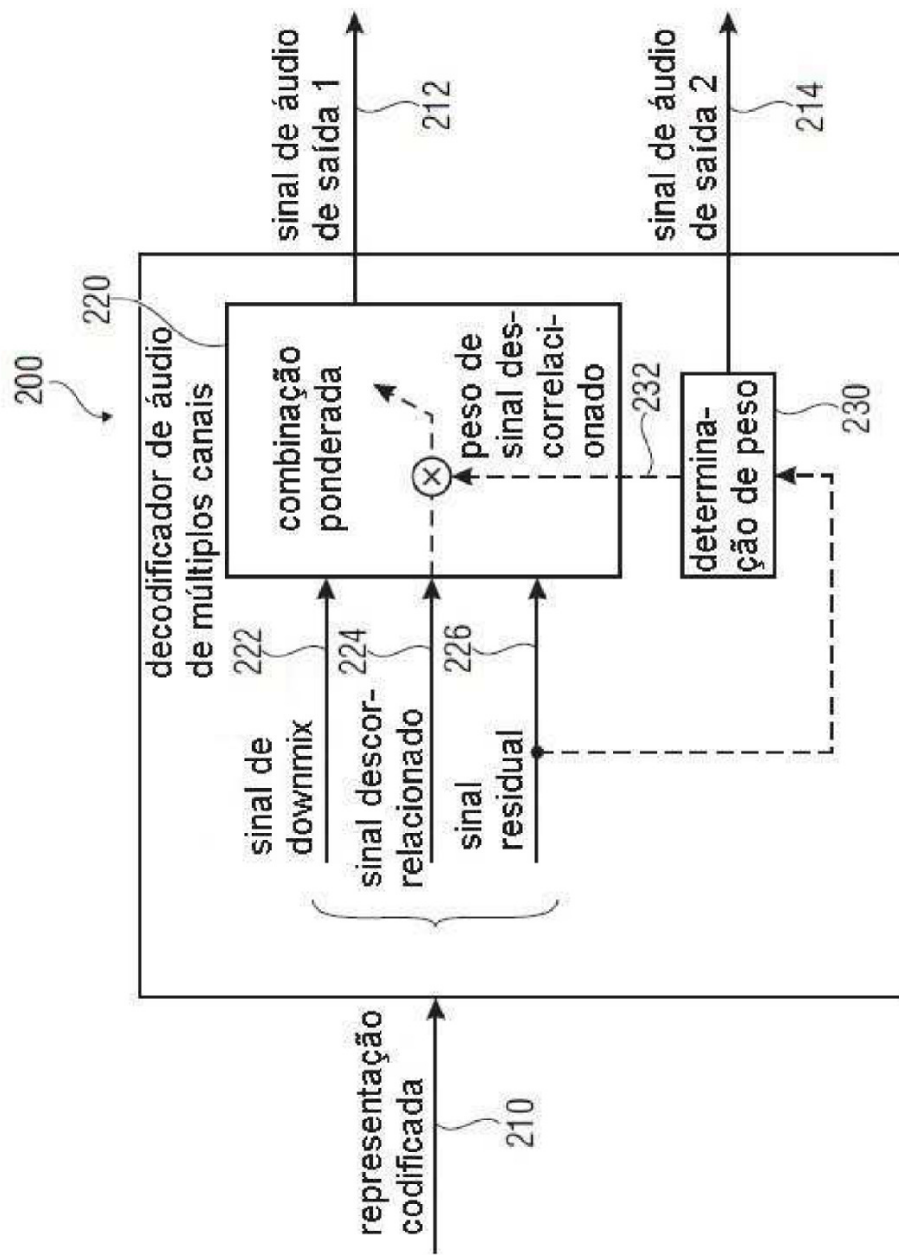


FIG. 2

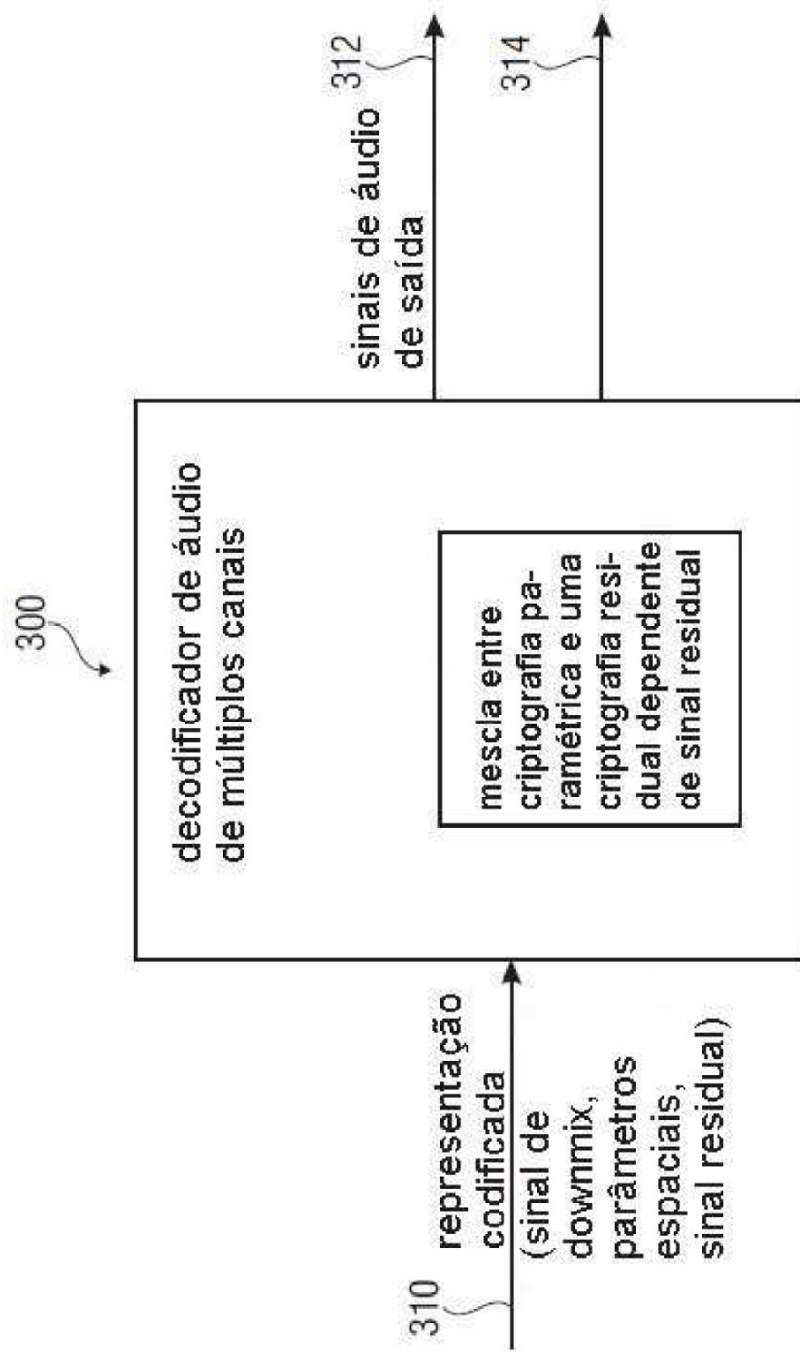


FIG. 3

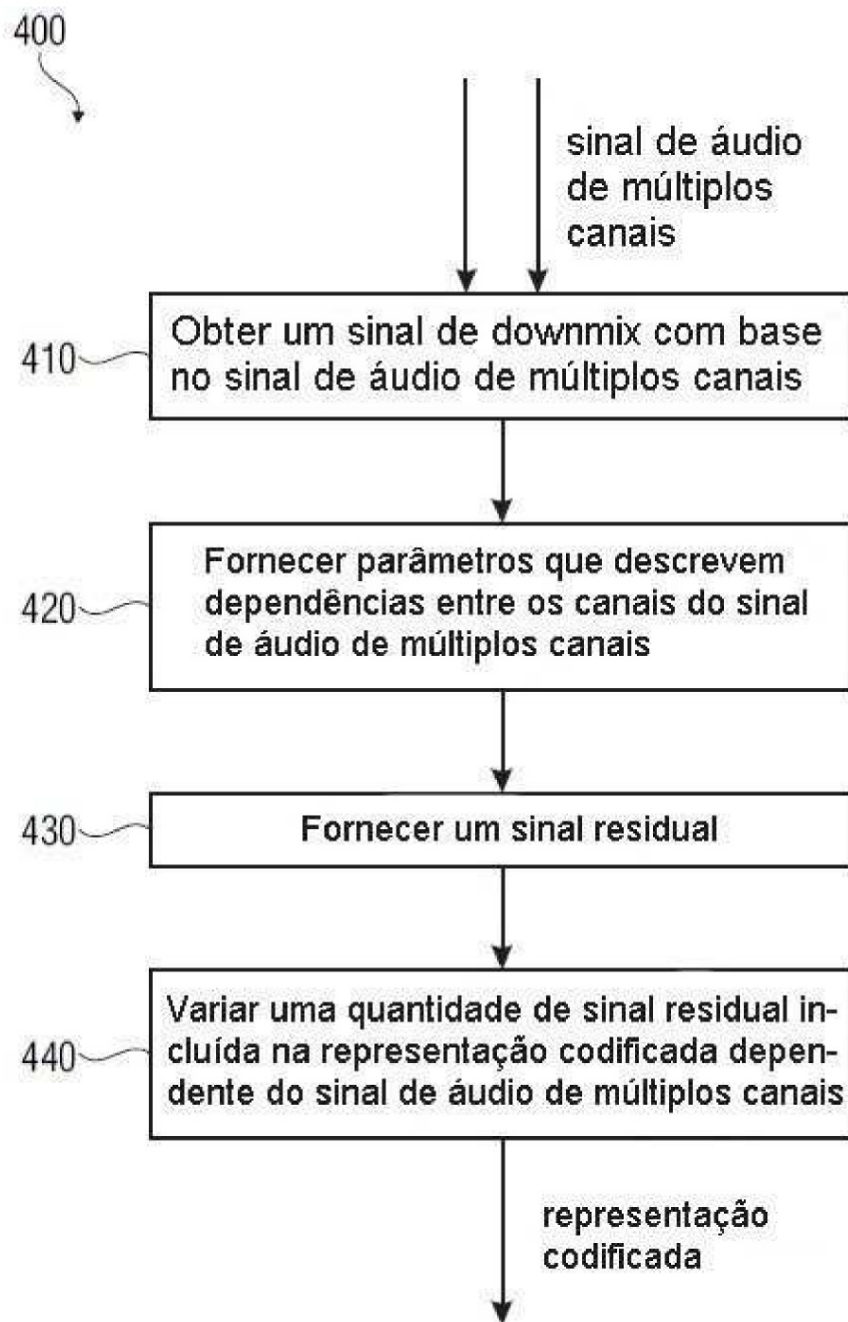


FIG. 4

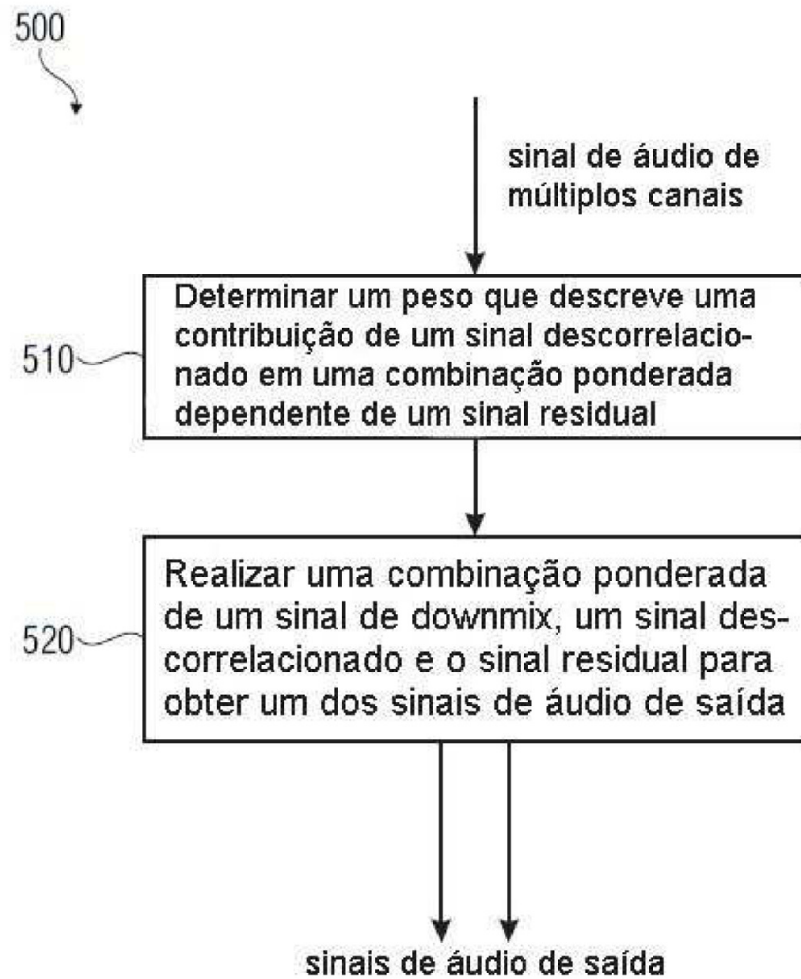


FIG. 5

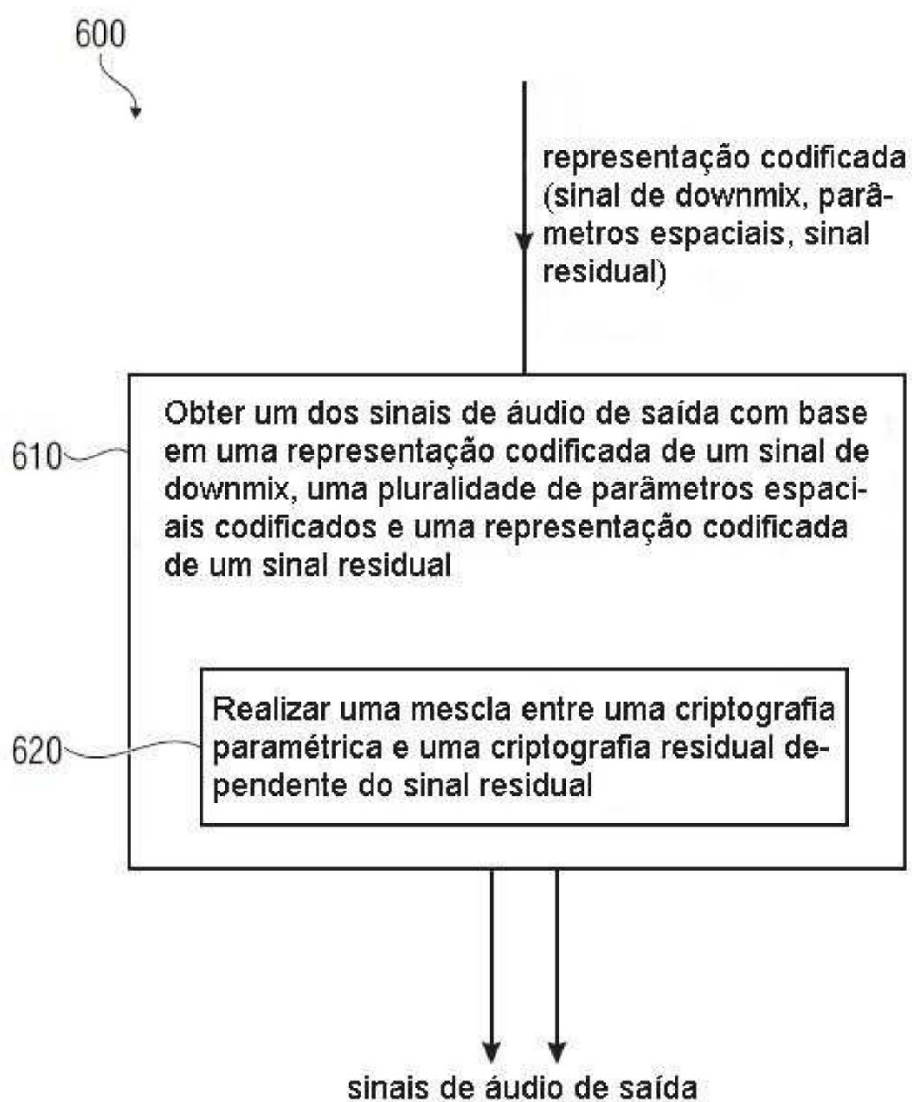


FIG. 6



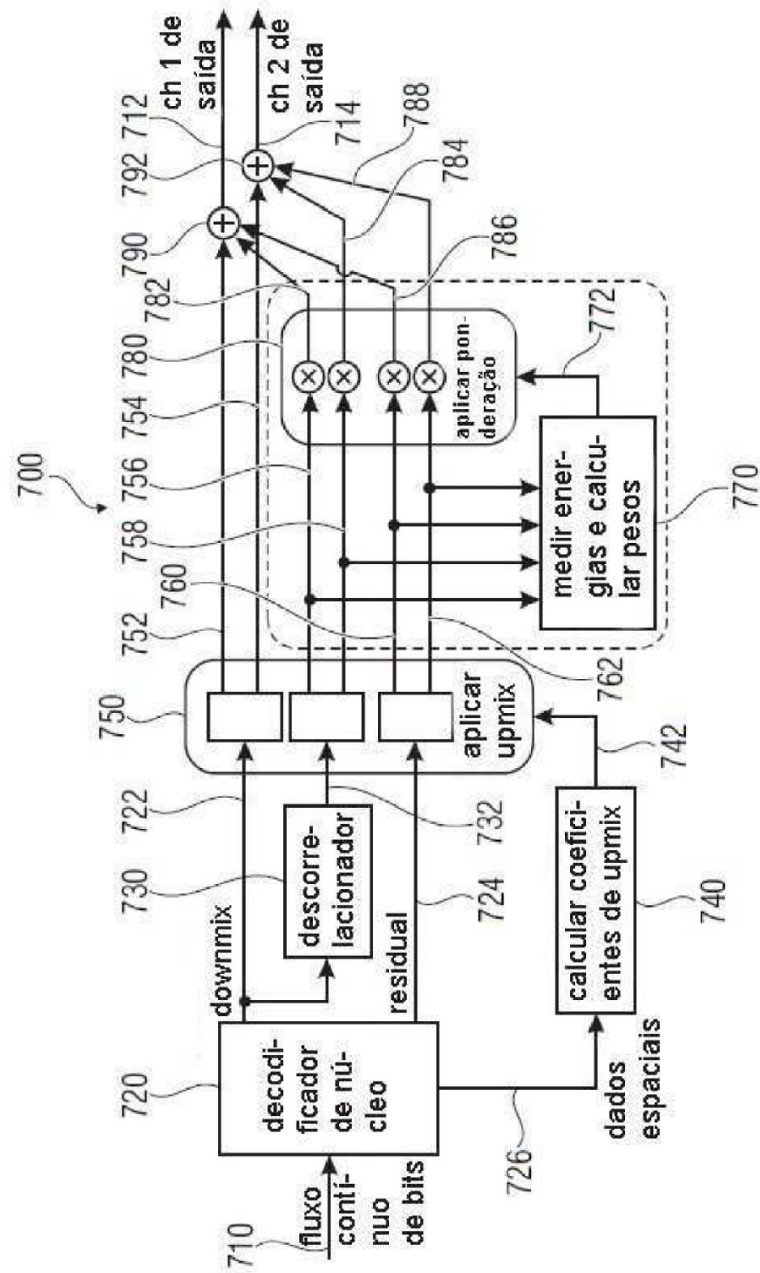
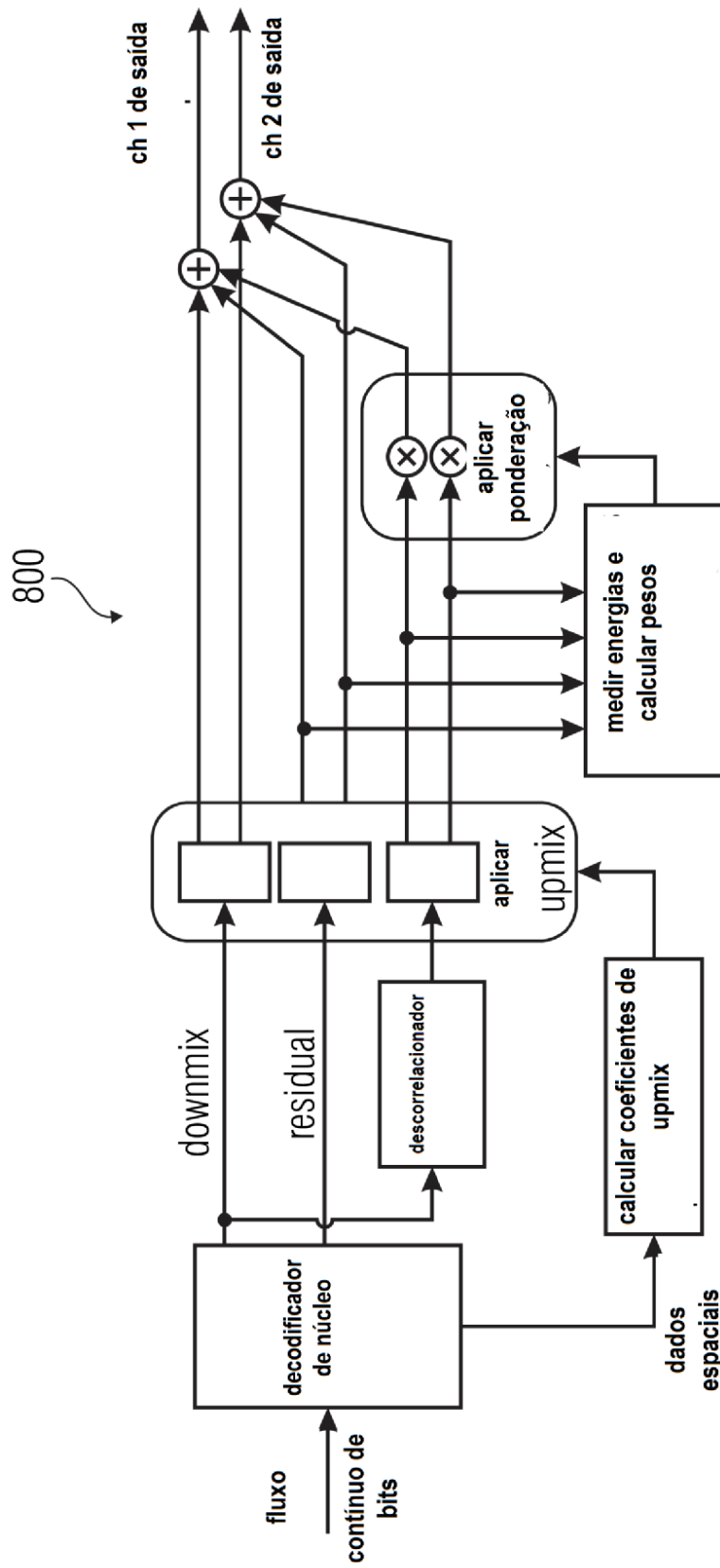


DIAGRAMA DE FLUXO DE DECODIFICADOR

FIG. 7



REPRESENTAÇÃO ESQUEMÁTICA DE DECODIFICADOR RESIDUAL HÍBRIDO

FIG 8