



República Federativa do Brasil

Ministério do Desenvolvimento, Indústria,
Comércio e Serviços

Instituto Nacional da Propriedade Industrial



(11) BR 112017006325-5 B1

(22) Data do Depósito: 30/09/2015

(45) Data de Concessão: 26/12/2023

(54) Título: MÉTODO DE DECODIFICAÇÃO E DECODIFICADOR PARA O REALCE DE DIÁLOGO

(51) Int.Cl.: G10L 21/0316; G10L 19/008.

(30) Prioridade Unionista: 02/10/2014 US 62/059,015; 04/03/2015 US 62/128,331.

(73) Titular(es): DOLBY INTERNATIONAL AB.

(72) Inventor(es): JEROEN KOPPENS; PER EKSTRAND.

(86) Pedido PCT: PCT EP2015072578 de 30/09/2015

(87) Publicação PCT: WO 2016/050854 de 07/04/2016

(85) Data do Início da Fase Nacional: 28/03/2017

(57) Resumo: MÉTODO DE DECODIFICAÇÃO E DECODIFICADOR PARA O REALCE DE DIÁLOGO. A presente invenção refere-se a um método para realçar o diálogo em um decodificador de um sistema de áudio. O método compreende a recepção de uma pluralidade de sinais de downmix que consistem em um downmix de uma maior pluralidade de canais; a recepção de parâmetros para o realce de diálogo que é definido com respeito a um subconjunto da pluralidade de canais que é submetida ao downmix em um subconjunto da pluralidade de sinais de downmix; o upmix do subconjunto de sinais de downmix parametricamente a fim de reconstruir o subconjunto da pluralidade de canais com respeito aos quais os parâmetros para o realce de diálogo são definidos; a aplicação do realce de diálogo ao subconjunto da pluralidade de canais com respeito aos quais os parâmetros para o realce de diálogo são definidos ao usar os parâmetros para o realce de diálogo para obter pelo menos um sinal de diálogo realçado; e a sujeição de pelo menos um sinal de diálogo realçado à mixagem para obter versões de diálogo realçadas do subconjunto de sinais de downmix.

Relatório Descritivo da Patente de Invenção para "**MÉTODO DE DECODIFICAÇÃO E DECODIFICADOR PARA O REALCE DE DIÁLOGO**".

Campo Técnico

[0001] A presente invenção refere-se de maneira geral à codificação de áudio. Em particular, ela se refere a métodos e dispositivos para realçar o diálogo em sistemas de áudio baseado em canais.

Antecedentes da Invenção

[0002] O realce de diálogo refere-se ao realce de diálogo com relação a um outro conteúdo de áudio. Isto pode, por exemplo, ser aplicado para permitir que as pessoas com problema de audição acompanhem o diálogo em um filme. Para o conteúdo de áudio baseado em canais, o diálogo está tipicamente presente em vários canais e também é mixado com um outro conteúdo de áudio. Portanto, é uma tarefa não trivial para realçar o diálogo.

[0003] Há vários métodos conhecidos para executar o realce de diálogo em um decodificador. De acordo com alguns destes métodos, o conteúdo de canal total, isto é, a configuração de canal total, é primeiramente decodificado e então os parâmetros de realce de diálogo recebidos são usados para predizer o diálogo com base no conteúdo de canal total. O diálogo predito é usado então para realçar o diálogo nos canais relevantes. No entanto, tais métodos de decodificação se baseiam em um decodificador com capacidade de decodificar a configuração de canal total.

[0004] No entanto, os decodificadores de baixa complexidade não são projetados tipicamente para decodificar a configuração de canal total. Ao invés disto, um decodificador de baixa complexidade pode decodificar e emitir um número menor de canais que representam uma versão de downmix da configuração de canal total. Por conseguinte, a configuração de canal total não fica disponível no decodificador de

baixa complexidade. Uma vez que os parâmetros de realce de diálogo são definidos com respeito aos canais da configuração de canal total (ou pelo menos com respeito a alguns dos canais da configuração de canal total), os métodos de realce de diálogo conhecidos não podem ser aplicados diretamente por um decodificador de baixa complexidade. Em particular, este é o caso, uma vez que os canais com respeito aos quais os parâmetros de realce de diálogo se aplicam ainda podem ser mixados com outros canais.

[0005] Desse modo, há espaço para melhorias que permitam que um decodificador de baixa complexidade aplique o realce de diálogo sem ter que decodificar a configuração de canal total.

Breve Descrição dos Desenhos

[0006] A seguir, as modalidades exemplificadoras serão descritas em mais detalhes e com referência aos desenhos anexos, nos quais:

[0007] Figura 1a é uma ilustração esquemática de uma configuração de canais 7.1+4 que é submetida a downmix em um downmix 5.1 de acordo com um primeiro esquema de downmix.

[0008] A Figura 1b é uma ilustração esquemática de uma configuração de canais 7.1+4 que é submetida a downmix em um downmix 5.1 de acordo com um segundo esquema de downmix.

[0009] A Figura 2 é uma ilustração esquemática de um decodificador da técnica anterior para executar o realce de diálogo em uma configuração de canais totalmente decodificada.

[00010] A Figura 3 é uma ilustração esquemática de realce de diálogo de acordo com um primeiro modo.

[00011] A Figura 4 é uma ilustração esquemática de realce de diálogo de acordo com um segundo modo.

[00012] A Figura 5 é uma ilustração esquemática de um decodificador de acordo com modalidades exemplificadoras.

[00013] A Figura 6 é uma ilustração esquemática de um decodifica-

dor de acordo com modalidades exemplificadoras.

[00014] A Figura 7 é uma ilustração esquemática de um decodificador de acordo com modalidades exemplificadoras.

[00015] A Figura 8 é uma ilustração esquemática de um codificador que corresponde a qualquer um dos decodificadores na Figura 2, na Figura 5, na Figura 6 e na Figura 7.

[00016] A Figura 9 ilustra métodos para computar uma operação de processamento em comum BA composta de duas suboperações A e B, com base nos parâmetros que controlam cada uma das suboperações.

[00017] Todas as figuras são esquemáticas e mostram de maneira geral somente os elementos que são necessários a fim de ilustrar a invenção, ao passo que outros elementos podem ser omitidos ou meramente sugeridos.

Descrição Detalhada

[00018] Em vista do acima exposto, um objetivo consiste na provisão de um decodificador e dos métodos associados que permitem a aplicação de realce de diálogo sem ter que decodificar a configuração de canal total.

Visão Geral

[00019] De acordo com um primeiro aspecto, as modalidades exemplificadoras provêm um método para realçar o diálogo em um decodificador de um sistema de áudio. O método compreende as etapas de:

[00020] receber uma pluralidade de sinais de downmix que consiste em um downmix de uma maior pluralidade de canais;

[00021] receber parâmetros para o realce de diálogo, em que os parâmetros são definidos com respeito a um subconjunto da pluralidade de canais incluindo canais que compreendem o diálogo, em que o subconjunto da pluralidade de canais é submetido a downmix em um

subconjunto da pluralidade de sinais de downmix;

[00022] receber parâmetros de reconstrução que permitem a reconstrução paramétrica dos canais que são submetidos a downmix no subconjunto da pluralidade de sinais de downmix;

[00023] upmixar subconjunto da pluralidade de sinais de downmix baseados parametricamente nos parâmetros de reconstrução a upmix, a fim de reconstruir o subconjunto da pluralidade de canais com respeito aos quais os parâmetros para o realce de diálogo são definidos;

[00024] aplicar o realce de diálogo ao subconjunto da pluralidade de canais com respeito aos quais os parâmetros para o realce de diálogo são definidos ao usar os parâmetros para o realce de diálogo para obter pelo menos um sinal de diálogo realçado; e

[00025] submeter pelo menos um sinal de diálogo realçado à mixagem de modo a prover versões de diálogo realçadas do subconjunto da pluralidade de sinais de downmix.

[00026] Com esse arranjo o decodificador não tem que reconstruir a configuração de canal total a fim de executar o realce de diálogo, reduzindo desse modo a complexidade. Ao invés disto, o decodificador reconstrói os canais que são requeridos para a aplicação de realce de diálogo. Isso inclui, em particular, um subconjunto da pluralidade de canais com respeito aos quais os parâmetros recebidos para o realce de diálogo são definidos. Uma vez que o realce de diálogo tenha sido realizado, isto é, quando pelo menos um sinal de diálogo realçado tiver sido determinado com base nos parâmetros para o realce de diálogo e o subconjunto da pluralidade de canais com respeito aos quais estes parâmetros são definidos, as versões de diálogo realçadas dos sinais de downmix recebidos são determinadas ao submeter o(s) sinal(is) de diálogo realçado a um procedimento de mixagem. Em consequência disto, as versões de diálogo realçadas dos sinais de downmix são produzidas para o playback subsequente pelo sistema de áudio.

[00027] Em modalidades exemplificadoras, uma operação de upmix pode ser completa (reconstrução de todo o conjunto de canais codificados) ou parcial (reconstrução de um subconjunto de canais).

[00028] Tal como usado no presente documento, um sinal de downmix refere-se a um sinal que é uma combinação de um ou mais sinais/canais.

[00029] Tal como usado no presente documento, efetuar o upmix parametricamente refere-se a uma reconstrução de um ou mais sinais/canais de um sinal de downmix por meio de técnicas paramétricas. É enfatizado que as modalidades exemplificadoras divulgadas no presente documento não ficam restringidas ao conteúdo baseado em canal (no sentido de sinais de áudio associados com direções, ângulos e/ou posições no espaço invariáveis ou predefinidos), mas também se estendem ao conteúdo baseado em objeto.

[00030] De acordo com modalidades exemplificadoras, na etapa de upmix do subconjunto da pluralidade de sinais de downmix parametricamente, nenhum sinal descorrelacionado é usado a fim de reconstruir o subconjunto da pluralidade de canais com respeito aos quais os parâmetros para o realce de diálogo são definidos.

[00031] Isso é vantajoso, uma vez que reduz a complexidade computacional ao mesmo tempo que melhora a qualidade das versões realçadas de diálogo resultantes dos sinais de downmix (isto é, a qualidade na saída). Em mais detalhes, as vantagens obtidas ao usar sinais descorrelacionados ao efetuar o upmix são reduzidas pela mixagem subsequente à qual o sinal de diálogo realçado é submetida. Portanto, o uso de sinais descorrelacionados pode ser vantajosamente omitido, desse modo poupando uma complexidade de computação. De fato, o uso de sinais descorrelacionados no upmix pode, em combinação com o resultado de realce de diálogo, resultar em uma qualidade mais baixa, uma vez que pode resultar em uma reverberação do des-

correlacionador no diálogo realçado.

[00032] De acordo com modalidades exemplificadoras, a mixagem é feito de acordo com os parâmetros de mixagem que descrevem uma contribuição de pelo menos um sinal de diálogo realçado para as versões de diálogo realçadas do subconjunto da pluralidade de sinais de downmix. Desse modo, pode haver alguns parâmetros de mixagem que descrevem como mixagem pelo menos um sinal de diálogo realçado a fim de obter versões de diálogo realçadas do subconjunto da pluralidade de sinais de downmix. Por exemplo, os parâmetros de mixagem podem estar na forma de pesos que descrevem quanto de pelo menos um sinal de diálogo realçado deve ser mixado em cada um dos sinais de downmix no subconjunto da pluralidade de sinais de downmix para obter as versões realçadas de diálogo do subconjunto da pluralidade de sinais de downmix. Tais pesos podem, por exemplo, estar na forma de parâmetros de renderização que são indicativos das posições espaciais associadas com pelo menos um sinal de diálogo realçado com relação às posições espaciais associadas com a pluralidade de canais e, portanto, o subconjunto de sinais de downmix correspondente. De acordo com outros exemplos, os parâmetros de mixagem podem indicar se pelo menos um diálogo de sinal realçado deve ou não contribuir, tal como ser incluído em, para uma versão particular da versão de diálogo realçada do subconjunto de sinais de downmix. Por exemplo, um "1" pode indicar que um sinal de diálogo realçado deve ser incluído quando da formação de uma versão particular de diálogo realçado dos sinais de downmix, e um "0" pode indicar que ele não deve ser incluído.

[00033] Na etapa de sujeição de pelo menos um sinal de diálogo realçado à mixagem de modo a obter versões de diálogo realçadas do subconjunto da pluralidade de sinais de downmix, os sinais de diálogo realçado podem ser mixados com outros sinais/canais.

[00034] De acordo com modalidades exemplificadoras, pelo menos um sinal de diálogo realçado é mixado com os canais que são reconstruídos na etapa de upmix, mas que não foram submetidos ao realce de diálogo. Em mais detalhes, a etapa de upmix do subconjunto da pluralidade de sinais de downmix parametricamente pode compreender a reconstrução de pelo menos um canal adicional além da pluralidade de canais com respeito aos quais os parâmetros para o realce de diálogo são definidos, e em que a mixagem compreende a mixagem de pelo menos um canal adicional junto com pelo menos um sinal de diálogo realçado. Por exemplo, todos os canais que são submetidos a downmix no subconjunto da pluralidade de sinais de downmix podem ser reconstruídos e incluídos na mixagem. Em tais modalidades, há tipicamente uma correspondência direta entre cada sinal de diálogo realçado e um canal.

[00035] De acordo com outras modalidades exemplificadoras, pelo menos um sinal de diálogo realçado é mixado com o subconjunto da pluralidade de sinais de downmix. Em mais detalhes, a etapa de upmix do subconjunto da pluralidade de sinais de downmix parametricamente pode compreender a reconstrução de somente o subconjunto da pluralidade de canais com respeito aos quais os parâmetros para o realce de diálogo são definidos, e a etapa de aplicação de realce de diálogo pode compreender a predição e o realce de um componente de diálogo do subconjunto da pluralidade de canais com respeito aos quais os parâmetros para o realce de diálogo são definidos ao usar os parâmetros para o realce de diálogo para obter pelo menos um sinal de diálogo realçado, e a mixagem pode compreender a mixagem de pelo menos um sinal de diálogo realçado com o subconjunto da pluralidade de sinais de downmix. Tais modalidades servem desse modo para prever e realçar o conteúdo do diálogo e a mixagem do mesmo no subconjunto da pluralidade de sinais de downmix.

[00036] De modo geral, deve ser observado que um canal pode compreender o conteúdo do diálogo que é mixado com o conteúdo que não do diálogo. Além disso, o conteúdo do diálogo que corresponde a um diálogo pode ser mixado em vários canais. A predição de um componente de diálogo do subconjunto da pluralidade de canais com respeito aos quais os parâmetros para o realce de diálogo são definidos significa de modo geral que o conteúdo do diálogo é extraído, isto é, separado, dos canais e combinado a fim de reconstruir o diálogo.

[00037] A qualidade do realce de diálogo também pode ser melhorada com a recepção e o uso de um sinal de áudio que representa o diálogo. Por exemplo, o sinal de áudio que representa o diálogo pode ser codificado a uma baixa taxa de bits que causa artefatos bem audíveis quando escutado separadamente. No entanto, quando usado em conjunto com o realce de diálogo paramétrico, isto é, a etapa de aplicação do realce de diálogo ao subconjunto da pluralidade de canais com respeito aos quais os parâmetros para o realce de diálogo são definidos ao usar os parâmetros para o realce de diálogo, o realce de diálogo resultante pode ser melhorado, por exemplo, em termos da qualidade de áudio. Mais particularmente, o método também pode compreender: a recepção de um sinal de áudio que representa o diálogo, em que a etapa de aplicação do realce de diálogo compreende a aplicação do realce de diálogo ao subconjunto da pluralidade de canais com respeito aos quais os parâmetros para o realce de diálogo são definidos ainda ao usar o sinal de áudio que representa o diálogo.

[00038] Em algumas modalidades, os parâmetros de mixagem já podem estar disponíveis no decodificador, por exemplo, eles podem ser codificados diretamente. Este deve ser em particular o caso se pelo menos um sinal de diálogo realçado fosse mixado sempre da mesma maneira, por exemplo, se fosse mixado sempre com os mesmos canais reconstruídos. Em outras modalidades o método compreende a

recepção de parâmetros de mixagem para a etapa de sujeição de pelo menos um sinal de diálogo realçado à mixagem. Por exemplo, os parâmetros de mixagem podem fazer parte dos parâmetros de realce de diálogo.

[00039] De acordo com modalidades exemplificadoras, o método compreende a recepção de parâmetros de mixagem que descrevem um esquema de downmix que descreve em que sinal de downmix cada canal da pluralidade de canais é mixado. Por exemplo, se cada sinal de diálogo realçado corresponder a um canal, que é por sua vez mixado com outros canais reconstruídos, a mixagem é realizada de acordo com o esquema de downmix de modo que cada canal seja mixado no sinal de downmix correto.

[00040] O esquema de downmix pode variar com o tempo, isto é, ele pode ser dinâmico, desse modo aumentando a flexibilidade do sistema.

[00041] O método também pode compreender a recepção de dados que identificam o subconjunto da pluralidade de canais com respeito aos quais os parâmetros para o realce de diálogo são definidos. Por exemplo, os dados que identificam o subconjunto da pluralidade de canais com respeito aos quais os parâmetros para o realce de diálogo são definidos podem ser incluídos nos parâmetros para o realce de diálogo. Desta maneira, pode ser sinalizado ao decodificador com respeito a quais canais o realce de diálogo deve ser realizado. Alternativamente, tal informação pode estar disponível no decodificador, por exemplo, ser codificada diretamente, o que significa que os parâmetros para o realce de diálogo são definidos sempre com respeito aos mesmos canais. Em particular, o método também pode incluir a recepção da informação que indica quais sinais dos sinais de diálogo realçado devem ser submetidos à mixagem. Por exemplo, o método de acordo com esta variação pode ser realizado por um sistema de deco-

dificação que opera em um modo particular, em que os sinais de diálogo realçado não são mixados de volta em um conjunto totalmente idêntico de sinais de downmix como era usado para a obtenção de sinais de diálogo realçado. Desta forma, a operação de mixagem pode na prática ser restringida a uma seleção não completa (um ou mais sinais) do subconjunto da pluralidade de sinais de downmix. Os outros sinais de diálogo realçado são adicionados aos sinais de downmix ligeiramente diferentes, tais como os sinais de downmix que foram submetidos a uma conversão de formato. Uma vez que os dados que identificam o subconjunto da pluralidade de canais com respeito aos quais os parâmetros para o realce de diálogo são definidos e o esquema de downmix são conhecidos, é possível encontrar o subconjunto da pluralidade de sinais de downmix em que o subconjunto da pluralidade de canais com respeito aos quais os parâmetros para o realce de diálogo são definidos é submetido ao downmix. Em mais detalhes, os dados que identificam o subconjunto da pluralidade de canais com respeito aos quais os parâmetros para o realce de diálogo são definidos em conjunto com o esquema de downmix podem ser usados para encontrar o subconjunto da pluralidade de sinais de downmix em que o subconjunto da pluralidade de canais com respeito aos quais os parâmetros para o realce de diálogo são definidos é submetido ao downmix.

[00042] As etapas de upmix do subconjunto da pluralidade de sinais de downmix, aplicação de realce de diálogo e mixagem podem ser executadas como operações de matriz definidas pelos parâmetros de reconstrução, pelos parâmetros para o realce de diálogo e pelos parâmetros de mixagem, respectivamente. Isto é vantajoso, uma vez que o método pode ser implementado de uma maneira eficiente com a execução da multiplicação de matriz.

[00043] Além disso, o método também pode compreender a combi-

nação, pela multiplicação de matriz, das operações de matriz que correspondem às etapas de upmix do subconjunto da pluralidade de sinais de downmix, aplicação do realce de diálogo e mixagem, em uma única operação de matriz antes da aplicação ao subconjunto da pluralidade de sinais de downmix. Desse modo, operações de matriz diferentes podem ser combinadas em uma única operação de matriz, desse modo melhorando ainda mais a eficiência e reduzindo a complexidade computacional do método.

[00044] Os parâmetros de realce de diálogo e/ou os parâmetros de reconstrução podem ser dependentes da frequência, permitindo desse modo que os parâmetros difiram entre faixas de frequência diferentes. Desta maneira, o realce de diálogo e a reconstrução podem ser otimizados em faixas de frequência diferentes, melhorando desse modo a qualidade do áudio de saída.

[00045] Em mais detalhes, os parâmetros para o realce de diálogo podem ser definidos com respeito a um primeiro conjunto de faixas de frequência e os parâmetros de reconstrução podem ser definidos com respeito a um segundo conjunto de faixas de frequência, em que o segundo conjunto de faixas de frequência é diferentes do primeiro conjunto de faixas de frequência. Isto pode ser vantajoso na redução da taxa de bits para transmitir os parâmetros para os parâmetros de realce de diálogo e de reconstrução em um fluxo de bits quando, por exemplo, o processo de reconstrução requer parâmetros a uma resolução de frequência maior do que o processo de realce de diálogo e/ou quando, por exemplo, o processo de realce de diálogo é executado em uma largura de faixa menor do que o processo de reconstrução.

[00046] De acordo com modalidades exemplificadoras, os valores (de preferência distintos) dos parâmetros para o realce de diálogo podem ser recebidos repetidamente e associados com um primeiro conjunto de instantes de tempo, em que os respectivos valores se aplicam

exatamente. Na presente invenção, uma afirmação para o efeito a que um valor se aplica, ou é conhecido, "exatamente" em um determinado instante de tempo deve se referir ao fato que o valor foi recebido pelo decodificador, tipicamente em conjunto com uma indicação explícita ou implícita de um instante de tempo onde se aplica. Por outro lado, um valor que é interpolado ou predito por um determinado instante de tempo não se aplica "exatamente" nesse instante de tempo nesse sentido, mas é uma estimativa do lado do decodificador. "Exatamente" não implica que o valor provê a reconstrução exata de um sinal de áudio. Entre instantes de tempo consecutivos no conjunto, um primeiro padrão de interpolação predefinido pode ser prescrito. Um padrão de interpolação, definindo como estimar um valor aproximado de um parâmetro em um instante de tempo localizado entre dois instantes de tempo limítrofes no conjunto em que os valores do parâmetro são conhecidos, pode ser, por exemplo, uma interpolação linear ou constante por partes. Se o instante de tempo da predição for localizado a uma determinada distância afastado de um dos instantes de tempo limítrofes, um padrão de interpolação linear é baseado na suposição que o valor do parâmetro no instante de tempo da predição depende linearmente da dita distância, ao passo que um padrão de interpolação constante por partes assegura que o valor do parâmetro não muda entre cada valor conhecido e o valor seguinte. Também pode haver outros padrões de interpolação possíveis incluindo, por exemplo, padrões que usam polinômios de graus maiores do que um, ranhuras, funções racionais, processos de Gauss, polinômios trigonométricos, ondulações, ou uma combinação destes, para estimar o valor do parâmetro em um determinado instante de tempo de predição. O conjunto de instantes de tempo pode não ser explicitamente transmitido ou declarado, mas ser de preferência inferido do padrão de interpolação, por exemplo, o ponto inicial ou o ponto final de um intervalo de interpolação li-

near, que pode ser implicitamente fixado aos limites de quadro de um algoritmo de processamento de áudio. Os parâmetros de reconstrução podem ser recebidos de uma maneira similar: os valores (de preferência distintos) dos parâmetros de reconstrução podem ser associados com um segundo conjunto de instantes de tempo, e um segundo padrão de interpolação pode ser executado entre instantes de tempo consecutivos.

[00047] O método também pode incluir a seleção de um tipo de parâmetro, em que o tipo consiste em parâmetros para o realce de diálogo ou parâmetros de reconstrução, de uma maneira tal que o conjunto dos instantes de tempo associados com o tipo selecionado inclui pelo menos um instante de predição que é um instante de tempo que está ausente do conjunto associado com o tipo não selecionado. Por exemplo, se o conjunto de instantes de tempo com os quais os parâmetros de reconstrução estão associados incluir um determinado instante de tempo que está ausente do conjunto de instantes de tempo com os quais os parâmetros para o realce de diálogo estão associados, tal instante de tempo será um instante de predição se o tipo selecionado de parâmetros consistir nos parâmetros de reconstrução e o tipo não selecionado de parâmetros consistir nos parâmetros para o realce de diálogo. De uma maneira similar, em uma outra situação, o instante de predição pode ser de preferência encontrado no conjunto de instantes de tempo com os quais os parâmetros para o realce de diálogo estão associados, e os tipos selecionados e não selecionados serão trocados. De preferência, o tipo de parâmetro selecionado é o tipo que tem a maior densidade de instantes de tempo com valores de parâmetro associados; em um exemplo de uso, isto pode reduzir a quantidade total de operações de predição necessárias.

[00048] O valor dos parâmetros do tipo não selecionado, no instante de predição, pode ser predito. A predição pode ser executada ao

usar um método de predição apropriado, tal como a interpolação ou a extrapolação, e em vista do padrão de interpolação predefinido para os tipos de parâmetro.

[00049] O método pode incluir a etapa de computação, com base em pelo menos o valor predito dos parâmetros do tipo não selecionado e um valor recebido dos parâmetros do tipo selecionado, de uma operação de processamento em comum que representa pelo menos o upmix do subconjunto dos sinais de downmix seguido pelo realce o diálogo no instante de predição. Além dos valores dos parâmetros de reconstrução e dos parâmetros para o realce de diálogo, a computação pode ser baseada em outros valores, tais como valores de parâmetros para mixagem, e a operação de processamento em comum também pode representar uma etapa de mixagem de um sinal de diálogo realçado de volta a um sinal de downmix.

[00050] O método pode incluir a etapa de computação, com base em pelo menos um valor (recebido ou predito) dos parâmetros do tipo selecionado e pelo menos um valor (recebido ou predito) dos parâmetros do tipo não selecionado, de maneira tal que pelo menos qualquer um dos valores é um valor recebido, com a operação de processamento em comum em um instante de tempo adjacente no conjunto associado com o tipo selecionado ou não selecionado. O instante de tempo adjacente pode ser mais cedo ou mais tarde do que o instante de predição, e não é essencial requerer que o instante de tempo adjacente seja o vizinho mais próximo em termos de distância.

[00051] No método, as etapas de upmix do subconjunto da pluralidade de sinais de downmix e de aplicação de realce de diálogo podem ser executadas entre o instante de predição e o instante de tempo adjacente por um valor interpolado da operação de processamento em comum computada. Com a interpolação da operação de processamento em comum computada, uma complexidade computacional reduzida

pode ser obtida. Com a não interpolação de ambos os tipos de parâmetros separadamente, e com a não formação de um produto (isto é, uma operação de processamento em comum), em cada ponto de interpolação, menos operações de adição matemática e multiplicação podem ser requeridas para obter um resultado igualmente útil em termos da qualidade de audição percebida.

[00052] De acordo com outras modalidades exemplificadoras, a operação de processamento em comum no instante de tempo adjacente pode ser computada com base em um valor recebido dos parâmetros do tipo selecionado e em um valor predito dos parâmetros do tipo não selecionado. A situação reversa também é possível, onde a operação de processamento em comum no instante de tempo adjacente pode ser computada com base em um valor predito dos parâmetros do tipo selecionado e um valor recebido dos parâmetros do tipo não selecionado. As situações em que um valor do mesmo tipo de parâmetro é um valor recebido no instante de predição e um valor predito no instante de tempo adjacente podem ocorrer se, por exemplo, os instantes de tempo no conjunto com o qual o tipo selecionado do parâmetro é associado estiverem localizados estritamente entre os instantes de tempo no conjunto com o qual o tipo de parâmetro não selecionado é associado.

[00053] De acordo com modalidades exemplificadoras, a operação de processamento em comum no instante de tempo adjacente pode ser computada com base em um valor recebido dos parâmetros do tipo de parâmetro selecionado e em um valor recebido dos parâmetros do tipo de parâmetro não selecionado. Tais situações podem ocorrer, por exemplo, se os valores exatos dos parâmetros de ambos os tipos forem recebidos para limites de quadro, mas também - para o tipo selecionado - para um instante de tempo a meio caminho entre os limites. Então, o instante de tempo adjacente é um instante de tempo as-

sociado com um limite de quadro, e o instante de tempo de predição fica localizado a meio caminho entre limites de quadros.

[00054] De acordo com outras modalidades exemplificadoras, o método também pode incluir a seleção, com base no primeira e no segundo padrões de interpolação, de um padrão de interpolação em comum de acordo com uma regra de seleção predefinida, em que a interpolação das respectivas operações de processamento em comum computadas são de acordo com o padrão de interpolação em comum. A regra de seleção predefinida pode ser definida para o caso em que o primeiro e o segundo padrões de interpolação são iguais, e também pode ser definida para o caso em que os primeiro e segundo padrões de interpolação são diferentes. Como um exemplo, se o primeiro padrão de interpolação for linear (e de preferência, se houver uma relação linear entre parâmetros e propriedades quantitativas da operação de realce de diálogo) e o segundo padrão de interpolação for constante por partes, o padrão de interpolação em comum pode ser selecionado para que seja linear.

[00055] De acordo com modalidades exemplificadoras, a predição do valor dos parâmetros do tipo não selecionado no instante de predição é feita de acordo com o padrão de interpolação para os parâmetros do tipo não selecionado. Isso pode envolver o uso de um valor exato do parâmetro do tipo não selecionado, em um instante de tempo no conjunto associado com o tipo não selecionado que é adjacente ao instante de predição.

[00056] De acordo com modalidades exemplificadoras, a operação de processamento em comum é computada como uma única operação de matriz e aplicada então ao subconjunto da pluralidade de sinais de downmix. De preferência, as etapas de upmix e de aplicação de realce de diálogo são executadas como operações de matriz definidas pelos parâmetros de reconstrução e pelos parâmetros para o realce de diá-

logo. Como um padrão de interpolação em comum, um padrão de interpolação linear pode ser selecionado, e o valor interpolado das respectivas operações de processamento em comum computadas pode ser computado pela interpolação de matriz linear. A interpolação pode ser restringida a tais elementos de matriz que mudam entre o instante de instante de predição e o instante de tempo adjacente, a fim de reduzir a complexidade computacional.

[00057] De acordo com modalidades exemplificadoras, os sinais de downmix recebidos podem ser segmentados em quadros de tempo, e o método pode incluir, na operação em estado constante, uma etapa de recepção de pelo menos um valor dos respectivos tipos de parâmetros que se aplica exatamente em um instante de tempo dentro de cada quadro de tempo. Tal como usado no presente documento, "estado estável" refere-se à operação que não envolve a presença de porções iniciais e finais de, por exemplo, uma música, e a operação que não envolve transientes internos que precisam de subdivisão de quadro.

[00058] De acordo com um segundo aspecto, é provido um produto de programa de computador que compreende um meio que pode ser lido por computador com instruções para executar o método do primeiro aspecto. O meio que pode ser lido por computador pode ser um meio ou um dispositivo que pode ser lido por computador não transitório.

[00059] De acordo com um terceiro aspecto, é provido um decodificador para realçar o diálogo em um sistema de áudio, em que o decodificador compreende:

[00060] um componente de recepção configurado para receber:

[00061] uma pluralidade de sinais de downmix que consistem em um downmix de um maior pluralidade de canais,

[00062] parâmetros para o realce de diálogo, em que os parâmetros são definidos com respeito a um subconjunto da pluralidade de canais

incluindo os canais que compreendem o diálogo, em que o subconjunto da pluralidade de canais é submetida ao downmix em um subconjunto da pluralidade de sinais de downmix, e

[00063] parâmetros de reconstrução que permitem a reconstrução paramétrico dos canais que são submetidos ao downmix no subconjunto da pluralidade de sinais de downmix;

[00064] um componente de upmix configurado para efetuar o upmix do subconjunto da pluralidade dos sinais de downmix parametricamente com base nos parâmetros de reconstrução a fim de reconstruir o subconjunto da pluralidade de canais com respeito aos quais os parâmetros para o realce de diálogo são definidos; e

[00065] um componente de realce de diálogo configurado para aplicar o realce de diálogo ao subconjunto da pluralidade de canais com respeito aos quais os parâmetros para o realce de diálogo são definidos ao usar os parâmetros para o realce de diálogo para obter pelo menos um sinal de diálogo realçado; e

[00066] um componente de mixagem configurado para submeter pelo menos um sinal de diálogo realçado à mixagem de modo a prover versões realçadas de diálogo do subconjunto da pluralidade de sinais de downmix.

[00067] De modo geral, o segundo e o terceiro aspectos podem compreender as mesmas características e vantagens que o primeiro aspecto.

Modalidades Exemplificadoras

[00068] A Figura 1a e a Figura 1b ilustram esquematicamente uma configuração de canais 7.1+4 (que corresponde a uma configuração de alto-falantes 7.1+4) com três canais anteriores L, C, R, dois canais surround LS, RS, dois canais posteriores LB, RB, quatro canais elevados TFL, TFR, TBL, TBR, e um canal de efeitos de baixa frequência LFE. No processo de codificação da configuração de canais 7.1+4, os

canais são tipicamente submetidos ao downmix, isto é, combinados em um número menor de sinais, indicados como sinais de downmix. No processo de downmix, os canais podem ser combinados de maneiras diferentes para formar diferentes configurações de downmix. A Figura 1a ilustra uma primeira configuração de 5.1 downmix 100a com os sinais de downmix l, c, r, ls, rs, lfe. Os círculos na figura indicam quais os canais que são submetidos ao downmix em quais sinais de downmix. A Figura 1b ilustra uma segunda configuração de 5.1 downmix 100b com os sinais de downmix l, c, r, tl, tr, lfe. A segunda configuração de 5.1 downmix 100b é diferente da primeira configuração de 5.1 downmix 100a, uma vez que os canais são combinados de uma maneira diferente. Por exemplo, na primeira configuração de downmix 100a, os canais L e TFL são submetidos ao downmix no sinal de downmix l, ao passo que na segunda configuração de downmix 100b os canais L, LS, LB são submetidos ao downmix no sinal de downmix l. A configuração de downmix é indicada às vezes no presente documento como um esquema de downmix que descreve quais os canais que são submetidos ao downmix em quais sinais de downmix. A configuração de downmix ou o esquema de downmix podem ser dinâmicos, uma vez que podem variar entre quadros de tempo de um sistema de codificação de áudio. Por exemplo, o primeiro esquema de downmix 100a pode ser usado em alguns quadros de tempo, ao passo que o segundo esquema de downmix 100b pode ser usado em outros quadros de tempo. No caso em que o esquema de downmix varia dinamicamente, o codificador pode enviar dados ao decodificador indicando qual o esquema de downmix que foi usado quando da codificação dos canais.

[00069] A Figura 2 ilustra um decodificador 200 da técnica anterior para o realce de diálogo. O decodificador compreende três componentes principais, um componente de recepção 202, um componente de

upmix, ou de reconstrução 204, e um componente de realce de diálogo (DE) 206. O decodificador 200 é do tipo que recebe uma pluralidade de sinais de downmix 212, reconstrói a configuração de canal total 218 com base nos sinais de downmix 212 recebidos, executa o realce de diálogo com respeito à configuração de canal total 218, ou pelo menos um subconjunto deste, e emite uma configuração de canais de diálogo realçado 220.

[00070] Em mais detalhes, o componente de recepção 202 é configurado para receber um fluxo de dados 210 (indicado às vezes como um fluxo de bits) de um codificador. O fluxo de dados 210 pode compreender tipos diferentes de dados, e o componente de recepção 202 pode decodificar o fluxo de dados 210 recebido em tipos diferentes de dados. Neste caso, o fluxo de dados compreende uma pluralidade de sinais de downmix 212, os parâmetros de reconstrução 214 e os parâmetros para o realce de diálogo 216.

[00071] O componente de upmix 204 reconstrói então a configuração de canal total com base na pluralidade de sinais de downmix 212 e nos parâmetros de reconstrução 214. Em outras palavras, o componente de upmix 204 reconstrói todos os canais 218 que foram submetidos ao downmix nos sinais de downmix 212. Por exemplo, o componente de upmix 204 pode reconstruir a configuração de canal total parametricamente com base nos parâmetros de reconstrução 214.

[00072] No exemplo ilustrado, os sinais de downmix 212 correspondem aos sinais de downmix de uma das configurações de 5.1 downmix das Figuras 1a e 1b, e os canais 218 correspondem aos canais da configuração de canais 7.1+4 das Figuras 1a e 1b. No entanto, os princípios do decodificador 200 devem naturalmente aplicar a outras configurações de canais/configurações de downmix.

[00073] Os canais reconstruídos 218, ou pelo menos um subconjunto dos canais reconstruídos 218, são então submetidos ao realce de

diálogo pelo componente de realce de diálogo 206. Por exemplo, o componente de realce de diálogo 206 pode executar uma operação de matriz nos canais reconstruídos 218, ou em pelo menos um subconjunto dos canais reconstruídos 218, a fim de emitir os canais de diálogo realçado. Tal operação de matriz é definida tipicamente pelos parâmetros de realce de diálogo 216.

[00074] A título de exemplo, o componente de realce de diálogo 206 pode submeter os canais C, L, R ao realce de diálogo a fim de obter os canais de diálogo realçado C_{DE} , L_{DE} , R_{DE} , ao passo que os outros canais são apenas atravessados tal como indicado pelas linhas tracejadas na Figura 2. Em tal situação, os parâmetros de realce de diálogo são definidos apenas com respeito aos canais C, L, R, isto é, com respeito a um subconjunto da pluralidade de canais 218. Por exemplo, os parâmetros de realce de diálogo 216 podem definir uma matriz 3x3 que pode ser aplicada aos canais C, L, R.

$$\begin{bmatrix} C_{DE} \\ L_{DE} \\ R_{DE} \end{bmatrix} = \begin{bmatrix} m_{11} & m_{12} & m_{13} \\ m_{21} & m_{22} & m_{23} \\ m_{31} & m_{32} & m_{33} \end{bmatrix}$$

[00075] Alternativamente, os canais não envolvidos no realce de diálogo podem ser atravessados por meio da matriz de realce de diálogo com 1 nas posições diagonais correspondentes e com 0 em todos os outros elementos nas fileiras e colunas correspondentes.

$$\begin{bmatrix} C_{DE} \\ L_{DE} \\ TFL \\ R_{DE} \\ TFR \\ LS \\ TBL \\ LB \\ RS \\ TBR \\ RB \\ LFE \end{bmatrix} = \begin{bmatrix} m_{11} & m_{12} & 0 & m_{13} & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ m_{21} & m_{22} & 0 & m_{23} & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 1 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ m_{31} & m_{32} & 0 & m_{33} & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 1 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 1 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 1 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 1 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 1 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 1 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 1 & 1 \end{bmatrix} \cdot \begin{bmatrix} C \\ L \\ TFL \\ R \\ TFR \\ LS \\ TBL \\ LB \\ RS \\ TBR \\ RB \\ LFE \end{bmatrix}$$

[00076] O componente de realce de diálogo 206 pode realizar o realce de diálogo de acordo com modos diferentes. Um primeiro modo, indicado no presente documento como realce paramétrico independente de canal, é ilustrado na Figura 3. O realce de diálogo é realizado com respeito a pelo menos um subconjunto de canais reconstruídos 218, tipicamente os canais que compreendem o diálogo, neste caso os canais L, R, C. Os parâmetros para o realce de diálogo 216 compreendem um conjunto de parâmetros para cada um dos canais a serem realçados. No exemplo ilustrado, os conjuntos de parâmetros são fornecidos pelos parâmetros p_1 , p_2 , p_3 que corresponde aos canais L, R, C, respectivamente. Em princípio, os parâmetros transmitidos neste modo representam a contribuição relativa do diálogo à energia da mixagem, para uma parcela da frequência de tempo em um canal. Além disso, há um fator de ganho g envolvido no processo de realce de diálogo. O fator de ganho g pode ser expresso como:

$$g = 10^{\frac{G}{20}} - 1$$

onde G é um ganho de realce de diálogo expresso em dB. O ganho de realce de diálogo G pode ser inserido, por exemplo, por um usuário e, portanto, não é incluído tipicamente no fluxo de dados 210 da Figura 2.

[00077] Quando no modo de realce paramétrico independente de canal, o componente de realce de diálogo 206 multiplica cada canal por seu parâmetro p_i correspondente e pelo fator de ganho g , e adiciona então o resultado ao canal, de modo a produzir os canais de diálogo realçado 220, neste caso L_{DE} , R_{DE} , C_{DE} . Ao usar a notação de matriz, isso pode ser escrito como:

$$X_e = (I + \text{diag}(p) \cdot g) \cdot X$$

onde X é uma matriz que tem os canais 218 (L, R, C) como fileiras, X_e é uma matriz que tem os canais de diálogo realçado 220 como fileiras, p é um vetor de fileira com entradas que correspondem aos parâmetros de realce de diálogo p_1 , p_2 , p_3 , para cada canal, e $\text{diag}(p)$ é uma

matriz diagonal que tem as entradas de p na diagonal.

[00078] Um segundo modo de realce de diálogo, indicado no presente documento como predição de diálogo de múltiplos canais, é ilustrado na Figura 4. Nesse modo, o componente de realce de diálogo 206 combina múltiplos canais 218 em uma combinação linear para predizer um sinal de diálogo 419. Além da adição coerente da presença do diálogo nos múltiplos canais, esta abordagem pode se beneficiar da subtração do ruído de fundo em um canal que compreende o diálogo ao usar um outro canal sem diálogo. Para esta finalidade, os parâmetros de realce de diálogo 216 compreendem um parâmetro para cada canal 218 que define o coeficiente do canal correspondente quando da formação da combinação linear. No exemplo ilustrado, os parâmetros de realce de diálogo 216 compreendem os parâmetros p_1 , p_2 , p_3 que correspondem aos canais L, R, C, respectivamente. Tipicamente, os algoritmos de otimização de erro de quadrado médio mínimo (MMSE) podem ser usados para gerar os parâmetros de predição no lado do codificador.

[00079] O componente de realce de diálogo 206 pode então realçar, isto é, ganhar, o sinal de diálogo predito 419 mediante a aplicação de um fator de ganho g , e adicionar o sinal de diálogo realçado aos canais 218, a fim de produzir os canais de diálogo realçado 220. Para adicionar o sinal de diálogo realçado aos canais corretos na posição espacial correta (ou então não irá realçar o diálogo com o ganho previsto) o acompanhamento do movimento entre os três canais é transmitido pelos coeficientes de renderização, neste caso r_1 , r_2 , r_3 . Sob a restrição que os coeficientes de renderização são de preservação de energia, isto é,

$$r_1^2 + r_2^2 + r_3^2 = 1$$

o terceiro coeficiente de renderização r_3 pode ser determinado a partir dos dois primeiros coeficientes de maneira tal que

$$r_3 = \sqrt{1 - r_1^2 - r_2^2}.$$

[00080] Ao usar a notação de matriz, o realce de diálogo realizado pelo componente de realce de diálogo 206 quando no modo de predição de diálogo de múltiplos canais pode ser escrito como:

$$X_e = (I + g \cdot H \cdot P) \cdot X$$

ou

$$X_e = \begin{bmatrix} 1 + g \cdot r_1 \cdot p_1 & g \cdot r_1 \cdot p_2 & g \cdot r_1 \cdot p_3 \\ g \cdot r_2 \cdot p_1 & 1 + g \cdot r_2 \cdot p_2 & g \cdot r_2 \cdot p_3 \\ g \cdot r_3 \cdot p_1 & g \cdot r_3 \cdot p_2 & 1 + g \cdot r_3 \cdot p_3 \end{bmatrix} \cdot X$$

onde I é a matriz de identidade, X é uma matriz que tem os canais 218 (L, R, C) como fileiras, X_e é uma matriz que tem os canais de diálogo realçado 220 como fileiras, P é um vetor de fileira com entradas que correspondem aos parâmetros de realce de diálogo p_1 , p_2 , p_3 para cada canal, H é um vetor de coluna que tem os coeficientes de renderização r_1 , r_2 , r_3 como entradas, e g é o fator de ganho com

$$g = 10^{\frac{G}{20}} - 1.$$

[00081] De acordo com um terceiro modo, indicado no presente documento como híbrido paramétrico de forma de onda, o componente de realce de diálogo 206 pode combinar qualquer um dentre o primeiro e o segundo modos com a transmissão de um sinal de áudio adicional (um sinal de forma de onda) que representa o diálogo. Este último é codificado tipicamente a uma taxa de bits baixa que causa artefatos bem audíveis quando escutado separadamente. Dependendo das propriedades do sinal dos canais 218 e do diálogo, e da taxa de bits atribuída à codificação de sinal de forma de onda de diálogo, o codificador também determina um parâmetro de combinação, que indica como as contribuições de ganho devem ser divididas entre a contribuição paramétrica (do primeiro ou do segundo modo) e o sinal de áudio adicional que representa o diálogo.

[00082] Em combinação com o segundo modo, o realce de diálogo

do terceiro modo pode ser escrito como:

$$X_g = H \cdot g_1 \cdot d_c + (I + H \cdot g_2 \cdot P) \cdot X$$

ou

$$X_g = \begin{bmatrix} 1 + g_2 \cdot r_1 \cdot p_1 & g_2 \cdot r_1 \cdot p_2 & g_2 \cdot r_1 \cdot p_3 & g_1 \cdot r_1 \\ g_2 \cdot r_2 \cdot p_1 & 1 + g_2 \cdot r_2 \cdot p_2 & g_2 \cdot r_2 \cdot p_3 & g_1 \cdot r_2 \\ g_2 \cdot r_3 \cdot p_1 & g_2 \cdot r_3 \cdot p_2 & 1 + g_2 \cdot r_3 \cdot p_3 & g_1 \cdot r_3 \end{bmatrix} \cdot \begin{bmatrix} X \\ d_c \end{bmatrix}$$

onde d_c é o sinal de áudio adicional que representa o diálogo, com

$$g_1 = \alpha_c \cdot \left(10^{\frac{G}{20}} - 1\right),$$

$$g_2 = (1 - \alpha_c) \cdot \left(10^{\frac{G}{20}} - 1\right).$$

[00083] Para a combinação com o realce independente de canal (o primeiro modo), um sinal de áudio que representa o diálogo é recebido

$$D_c = \begin{pmatrix} d_{c,1} \\ d_{c,2} \\ d_{c,3} \end{pmatrix}$$

para cada canal 218. Ao escrever

o realce de diálogo pode ser escrito como:

$$X_g = g_1 \cdot D_c + (I + \text{diag}(p) \cdot g_2) \cdot X.$$

[00084] A Figura 5 ilustra um decodificador 500 de acordo com as modalidades exemplificadoras. O decodificador 500 é do tipo que decodifica uma pluralidade de sinais de downmix, que é um downmix de uma maior pluralidade de canais, para o playback subsequente. Em outras palavras, o decodificador 500 é diferente do decodificador da Figura 2, uma vez que não é configurado para reconstruir a configuração de canal total.

[00085] O decodificador 500 compreende um componente de recepção 502, e um bloco de realce de diálogo 503 que compreende um componente de upmix 504, um componente de realce de diálogo 506 e um componente de mixagem 508.

[00086] Tal como explicado com referência à Figura 2, o componente de recepção 502 recebe um fluxo de dados 510 e decodifica os mesmos em seus componentes, neste caso uma pluralidade de sinais

de downmix 512 que consistem em um downmix de uma pluralidade maior de canais (de acordo com as Figuras 1a e 1b), dos parâmetros de reconstrução 514 e dos parâmetros para o realce de diálogo 516. Em alguns casos, o fluxo de dados 510 também compreende os dados indicativos dos parâmetros de mixagem 522. Por exemplo, os parâmetros de mixagem podem fazer parte dos parâmetros para o realce de diálogo. Em outros casos, os parâmetros de mixagem 522 já estão disponíveis no decodificador 500, por exemplo, eles podem ser codificados diretamente no decodificador 500. Em outros casos, os parâmetros de mixagem 522 estão disponíveis para múltiplos conjuntos de parâmetros de mixagem e os dados no fluxo de dados 510 fornecem uma indicação de qual conjunto desses múltiplos conjuntos de parâmetros de mixagem é usado.

[00087] Os parâmetros para o realce de diálogo 516 são definidos tipicamente com respeito a um subconjunto da pluralidade de canais. Os dados que identificam o subconjunto da pluralidade de canais com respeito aos quais os parâmetros para o realce de diálogo são definidos podem ser incluídos no fluxo de dados 510 recebido, por exemplo, como parte dos parâmetros para o realce de diálogo 516. Alternativamente, o subconjunto da pluralidade de canais com respeito aos quais os parâmetros para o realce de diálogo são definidos pode ser codificado diretamente no decodificador 500. Por exemplo, com respeito à Figura 1a, os parâmetros para o realce de diálogo 516 podem ser definidos com respeito aos canais L, TFL que são submetidos ao downmix no sinal de downmix l, ao canal C que é compreendido no sinal de downmix c, e aos canais R, TFR que são submetidos ao downmix no sinal de downmix r. Para finalidades de ilustração, supõe-se que o diálogo só está presente nos canais L, C e R. Deve ser observado que os parâmetros para o realce de diálogo 516 podem ser definidos com respeito aos canais que compreendem o diálogo, tais como os canais

L, C, R, mas também podem ser definidos com respeito aos canais que não compreendem nenhum diálogo, tais como os canais TFL, TFR neste exemplo. Dessa maneira, o ruído de fundo em um canal que compreende o diálogo pode, por exemplo, ser subtraído ao usar um outro canal sem diálogo.

[00088] O subconjunto de canais com respeito aos quais os parâmetros para o realce de diálogo 516 são definidos é submetido ao downmix em um subconjunto 512a da pluralidade de sinais de downmix 512. No exemplo ilustrado, o subconjunto de sinais de downmix 512a compreende os sinais de downmix c, l e r. Esse subconjunto de sinais de downmix 512a é inserido no bloco de realce de diálogo 503. O subconjunto de sinais de downmix 512a correspondente pode, por exemplo, ser encontrado com base no conhecimento do subconjunto da pluralidade de canais com respeito aos quais os parâmetros para o realce de diálogo são definidos e no esquema de downmix.

[00089] O componente de upmix 514 usa técnicas paramétricas tal como é conhecido no estado da técnica para a reconstrução dos canais que são submetidos ao downmix no subconjunto de sinais de downmix 512a. A reconstrução é baseada nos parâmetros de reconstrução 514. Em particular, o componente de upmix 504 reconstrói o subconjunto da pluralidade de canais com respeito aos quais os parâmetros para o realce de diálogo 516 são definidos. Em algumas modalidades, o componente de upmix 504 reconstrói somente o subconjunto da pluralidade de canais com respeito aos quais os parâmetros para o realce 516 do diálogo são definidos. Tais modalidades exemplificadoras serão descritas com referência à Figura 7. Em outras modalidades, o componente de upmix 504 reconstrói pelo menos um canal além do subconjunto da pluralidade de canais com respeito aos quais os parâmetros para o realce de diálogo 516 são definidos. Tais modalidades exemplificadoras serão descritas com referência à Figura 6.

[00090] Os parâmetros de reconstrução podem não ser somente variáveis com o tempo, mas também podem ser dependentes da frequência. Por exemplo, os parâmetros de reconstrução podem assumir valores diferentes para faixas de frequência diferentes. Isso irá melhorar de modo geral a qualidade dos canais reconstruídos.

[00091] Tal como é conhecido no estado da técnica, o upmix paramétrico pode incluir em geral a formação de sinais descorrelacionados a partir dos sinais de entrada que são submetidos ao upmix, e a reconstrução paramétrica de sinais com base nos sinais de entrada e nos sinais descorrelacionados. Vide, por exemplo, o livro "Spatial Audio Processing: MPEG Surround and Other Applications" da autoria de Jeroen Breebaart e Christof Faller, ISBN: 978-9-470-03350-0. No entanto, o componente de upmix 504 executa de preferência um upmix paramétrico sem usar nenhum de tais sinais descorrelacionados. As vantagens obtidas com o uso de sinais descorrelacionados são neste caso reduzidas pelo downmix subsequente executado no componente de mixagem 508. Portanto, o uso de sinais descorrelacionados pode ser vantajosamente omitido pelo componente de upmix 504, conservando desse modo a complexidade de computação. De fato, o uso de sinais descorrelacionados no upmix em combinação com o realce de diálogo resulta em uma qualidade mais pobre, uma vez que pode resultar em uma reverberação do descorrelacionador no diálogo.

[00092] O componente de realce de diálogo 506 aplica então o realce de diálogo ao subconjunto da pluralidade de canais com respeito aos quais os parâmetros para o realce 516 do diálogo são definidos para produzir pelo menos um sinal de diálogo realçado. Em algumas modalidades, o sinal de diálogo realçado corresponde às versões de diálogo realçado do subconjunto da pluralidade de canais com o respeito aos quais os parâmetros para o realce 516 do diálogo são definidos. Isso será explicado em mais detalhes a seguir com referência à

Figura 6. Em outras modalidades, o sinal de diálogo realçado corresponde a um componente de diálogo predito e realçado do subconjunto da pluralidade de canais com respeito aos quais os parâmetros para o realce de diálogo 516 são definidos. Isto será explicado em mais detalhes a seguir com referência à Figura 7.

[00093] De modo similar aos parâmetros de reconstrução, os parâmetros para o realce de diálogo podem variar com o tempo, bem como com a frequência. Em mais detalhes, os parâmetros para o realce de diálogo podem assumir valores diferentes para faixas de frequência diferentes. O conjunto de faixas de frequência com respeito ao qual os parâmetros de reconstrução são definidos pode diferir do conjunto de faixas de frequência com respeito ao qual os parâmetros de realce de diálogo são definidos.

[00094] O componente de mixagem 508 executa então a mixagem com base em pelo menos um sinal de diálogo realçado de modo a obter as versões de diálogo realçado 520 do subconjunto de sinais de downmix 512a. No exemplo ilustrado, as versões de diálogo realçado 520 do subconjunto de sinais de downmix 512a são fornecidas por c_{DE} , l_{DE} , r_{DE} que correspondem aos sinais de downmix c , l , r , respectivamente.

[00095] A mixagem pode ser feita de acordo com os parâmetros de mixagem 522 que descrevem uma contribuição de pelo menos um sinal de diálogo realçado para as versões de diálogo realçado 520 do subconjunto de sinais de downmix 512a. Em algumas modalidades, vide a Figura 6, pelo menos um sinal de diálogo realçado é mixado em conjunto com os canais que foram reconstruídos pelo componente de upmix 504. Nesses casos, os parâmetros de mixagem 522 podem corresponder a um esquema de downmix, vide as Figuras 1a e 1b, descrevendo em qual dos sinais de downmix de diálogo realçado 520 cada canal deve ser mixado. Em outras modalidades, vide a Figura 7,

pelo menos um dos sinais de diálogo realçado é mixado em conjunto com o subconjunto de sinais de downmix 512a. Em tal caso, os parâmetros de mixagem 522 podem corresponder aos fatores de ponderação que descrevem como pelo menos um sinal de diálogo realçado deve ser ponderado no subconjunto de sinais de downmix 512a.

[00096] A operação de upmix executada pelo componente de upmix 504, a operação de realce de diálogo executada pelo componente de realce de diálogo 506 e a operação de mixagem executada pelo componente de mixagem 508 são tipicamente operações lineares, cada uma das quais pode ser definida por uma operação de matriz, isto é, por um produto de vetor de matriz. Isso é pelo menos verdadeiro se os sinais do descorrelacionador forem omitidos na operação de upmix. Em particular, a matriz associada com a operação de upmix (U) é definida por/pode ser derivada dos parâmetros de reconstrução 514. A este respeito, deve ser observado que o uso de sinais do descorrelacionador na operação de upmix ainda é possível, mas que a criação dos sinais descorrelacionados não é então parte da operação de matriz para o upmix. A operação de upmix com os descorrelacionadores pode ser vista como uma abordagem de dois estágios. Em um primeiro estágio, os sinais de downmix de entrada são alimentados em uma matriz do pré-descorrelacionador, e cada um dos sinais de saída depois da aplicação da matriz do pré-descorrelacionador é alimentado em um descorrelacionador. Em um segundo estágio, os sinais de downmix de entrada e os sinais de saída dos descorrelacionadores são alimentados na matriz de upmix, onde os coeficientes da matriz de upmix que correspondem aos sinais de downmix de entrada formam o que é indicado como "matriz de upmix seca" e os coeficientes que correspondem aos sinais de saída dos descorrelacionadores formam o que é indicado como "matriz de upmix úmida". Cada submatriz mapeia para a configuração de canal de upmix. Quando os sinais do descorre-

lacionador não são usados, a matriz associada com a operação de upmix só é configurada para a operação nos sinais de entrada 512a, e as colunas relacionadas aos sinais descorrelacionados (a matriz de upmix úmida) não são incluídas na matriz. Em outras palavras, a matriz de upmix neste caso corresponde à matriz de upmix seca. No entanto, tal como observado acima, o uso de sinais do descorrelacionador irão neste caso resultar tipicamente em uma qualidade mais pobre.

[00097] A matriz associada com a operação de realce de diálogo (M) é definida/pode ser derivada dos parâmetros para realce de diálogo 516, e a matriz associada com a operação de mixagem é definida/pode ser derivada dos parâmetros de mixagem 522.

[00098] Uma vez que a operação de upmix, a operação de realce de diálogo e a operação de mixagem são todas operações lineares, as matrizes correspondentes podem ser combinadas, pela multiplicação de matriz, em uma única matriz E (então, $X_{DE} = E \cdot X$ com $E = C \cdot M \cdot U$). Neste caso, X é um vetor de coluna dos sinais de downmix 512a, e X_{DE} é um vetor de coluna dos sinais de downmix de diálogo realçado 520. Desse modo, o bloco de realce de diálogo completo 503 pode corresponder a uma única operação de matriz que é aplicada ao subconjunto do produto dos sinais de downmix 512a a fim de produzir as versões de diálogo realçado 520 do subconjunto de sinais de downmix 512a. Por conseguinte, os métodos descritos no presente documento podem ser implementados de uma maneira muito eficiente.

[00099] A Figura 6 ilustra um decodificador 600 que corresponde a uma modalidade exemplificadora do decodificador 500 da Figura 5. O decodificador 600 compreende um componente de recepção 602, um componente de upmix 604, um componente de realce de diálogo 606 e um componente de mixagem 608.

[000100] De modo similar ao decodificador 500 da Figura 5, o componente de recepção 602 recebe um fluxo de dados 610 e decodifica o

mesmo em uma pluralidade de sinais de downmix 612, parâmetros de reconstrução 614 e parâmetros para o realce de diálogo 616.

[000101] O componente de upmix 604 recebe um subconjunto 612a (que corresponde ao subconjunto 512a) da pluralidade de sinais de downmix 612. Para cada um dos sinais de downmix no subconjunto 612a, o componente de upmix 604 reconstrói todos os canais que foram submetidos ao downmix no sinal de downmix ($X_u = U \cdot X$). Isso inclui os canais 618a com respeito aos quais os parâmetros para o realce de diálogo são definidos, e os canais 618b que não devem ser envolvidos no realce de diálogo. Com respeito à Figura 1b, os canais 618a com respeito aos quais os parâmetros para o realce de diálogo são definidos podem, por exemplo, corresponder aos canais L, LS, C, R, RS, e os canais 618b que não devem ser envolvidos no realce de diálogo podem corresponder aos canais LB, RB.

[000102] Os canais 618a com respeito aos quais os parâmetros para o realce de diálogo são definidos (X'_u) são então submetidos ao realce de diálogo pelo componente de realce de diálogo 606, ao passo que os canais 618b que não devem ser envolvidos no realce de diálogo são desviados do componente de realce de diálogo 606.

[000103] O componente de realce de diálogo 606 pode aplicar qualquer uma dentre o primeiro, o segundo e o terceiro modos de realce de diálogo descritos acima. No caso em que o terceiro modo é aplicado, o fluxo de dados 610 pode, tal como explicado acima, compreender um sinal de áudio que representa o diálogo (isto é, uma forma de onda codificada que representa o diálogo) a ser aplicado no realce de diálogo em conjunto com o subconjunto 618a da pluralidade de canais com respeito aos quais os parâmetros para o realce de diálogo são

definidos $\left(X_s = M \cdot \begin{bmatrix} X'_u \\ D_c \end{bmatrix} \right)$.

[000104] Em consequência disto, o componente de realce de diálogo 606 emite os sinais de diálogo realçados 619, que neste caso corres-

pondem às versões de diálogo realçado do subconjunto 618a dos canais com respeito aos quais os parâmetros para o realce de diálogo são definidos. A título de exemplo, os sinais de diálogo realçados 619 podem corresponder às versões de diálogo realçado dos canais L, LS, C, R, RS da Figura 1b.

[000105] O componente de mixagem 608 mistura então os sinais de diálogo realçados 619 junto com os canais 618b que não foram envol-

vidos no realce de diálogo $\left(X_{DE} = C \cdot \begin{bmatrix} X'_s \\ X'_u \end{bmatrix} \right)$ a fim de produzir as versões de diálogo realçado 620 do subconjunto de sinais de downmix 612a. O componente de mixagem 608 efetua a mixagem de acordo com o esquema de downmix atual, tal como o esquema de downmix ilustrado na Figura 1b. Neste caso, os parâmetros de mixagem 622 correspondem desse modo a um esquema de downmix que descreve em qual sinal de downmix 620 cada canal 619, 618b deve ser mixado. O esquema de downmix pode ser estático e, portanto, conhecido pelo decodificador 600, o que significa que o mesmo esquema de downmix se aplica sempre, ou o esquema de downmix pode ser dinâmico, o que significa que pode variar de quadro a quadro, ou pode ser um dentre vários esquemas conhecidos no decodificador. Neste último caso, uma indicação a respeito do esquema de downmix é incluída no fluxo de dados 610.

[000106] Na Figura 6, o decodificador é equipado com um componente de reembaralhamento 630 opcional. O componente de reembaralhamento 630 pode ser usado para a conversão entre esquemas de downmix diferentes, por exemplo, para converter do esquema 100b no esquema 100a. Deve ser observado que o componente de reembaralhamento 630 deixa tipicamente os sinais c e lfe inalterados, isto é, ele age como um componente passante com respeito a esses sinais. O componente de reembaralhamento 630 pode receber e operar (não mostrado) com base em vários parâmetros tais como, por exemplo, os

parâmetros de reconstrução 614 e os parâmetros para o realce de diálogo 616.

[000107] A Figura 7 ilustra um decodificador 700 que corresponde a uma modalidade exemplificadora do decodificador 500 da Figura 5. O decodificador 700 compreende um componente de recepção 702, um componente de upmix 704, um componente de realce de diálogo 706 e um componente de mixagem 708.

[000108] De modo similar ao decodificador 500 da Figura 5, o componente de recepção 702 recebe um fluxo de dados 710 e decodifica o mesmo em uma pluralidade de sinais de downmix 712, parâmetros de reconstrução 714 e parâmetros para o realce de diálogo 716.

[000109] O componente de upmix 704 recebe um subconjunto 712a (que corresponde ao subconjunto 512a) da pluralidade de sinais de downmix 712. Ao contrário da modalidade descrita com respeito à Figura 6, o componente de upmix 704 reconstrói somente o subconjunto 718a da pluralidade de canais com respeito aos quais os parâmetros para o realce 716 do diálogo são definidos. Com respeito à Figura 1b, os canais 718a com respeito aos quais os parâmetros para o realce de diálogo são definidos podem, por exemplo, corresponder aos canais C, L, LS, R, RS.

[000110] O componente de realce de diálogo 706 executa então o realce de diálogo nos canais 718a com o respeito a que os parâmetros para o realce de diálogo são definidos. Neste caso, o componente de realce de diálogo 706 prossegue para predizer um componente de diálogo com base nos canais 718a mediante a formação de uma combinação linear dos canais 718a, de acordo com um segundo modo de realce de diálogo. Os coeficientes usados quando da formação da combinação linear, denotados por p_1 a p_5 na Figura 7, são incluídos nos parâmetros para o realce de diálogo 716. O componente de diálogo predito é então realçado pela multiplicação de um fator de ganho g

para produzir um sinal de diálogo realçado 719. O fator de ganho g pode ser expresso como:

$$g = 10^{\frac{G}{20}} - 1$$

[000111] onde G é um ganho de realce de diálogo expresso em dB. O ganho de realce de diálogo G pode, por exemplo, ser inserido por um usuário e, portanto, não é incluído tipicamente no fluxo de dados 710. Deve ser observado que, no caso em que há vários componentes de diálogo, o procedimento de predição e realce acima pode ser aplicado uma vez por componente de diálogo.

[000112] O sinal de diálogo realçado predito 719 (ou seja, os componentes de diálogo preditos e realçados) é mixado então no subconjunto 712a de sinais de downmix a fim de produzir as versões de diálogo realçado 720 do subconjunto 712a de sinais de downmix $(X_{DE} = C \cdot \begin{bmatrix} X_d \\ X \end{bmatrix})$. A mixagem é efetuada de acordo com os parâmetros de mixagem 722 que descrevem uma contribuição do sinal de diálogo realçado 719 às versões de diálogo realçado 720 do subconjunto de sinais de downmix. Os parâmetros de mixagem são incluídos tipicamente no fluxo de dados 710. Neste caso, os parâmetros de mixagem 722 correspondem aos fatores de ponderação r_1 , r_2 , r_3 que descrevem como pelo menos um sinal de diálogo realçado 719 deve ser ponderado no subconjunto de sinais de downmix 712a:

$$X_{DE} = X + \begin{bmatrix} r_1 \\ r_2 \\ r_3 \end{bmatrix} \cdot X_d = \begin{bmatrix} r_1 & 1 & 0 & 0 \\ r_2 & 0 & 1 & 0 \\ r_3 & 0 & 0 & 1 \end{bmatrix} \cdot \begin{bmatrix} X_d \\ X_1 \\ X_2 \\ X_3 \end{bmatrix}$$

[000113] Em mais detalhes, os fatores ponderados podem corresponder aos coeficientes de renderização que descrevem o acompanhamento em movimento de pelo menos um sinal de diálogo realçado 719 com respeito ao subconjunto de sinais de downmix 712a, de maneira tal que o sinal de diálogo realçado 719 é adicionado aos sinais de downmix 712a nas posições espaciais corretas.

[000114] Os coeficientes de renderização (os parâmetros de mixagem 722) no fluxo de dados 710 podem corresponder aos canais sujeitos ao upmix 718a. No exemplo ilustrado, há cinco canais sujeitos ao upmix 718a e, desse modo, pode haver cinco coeficientes de renderização correspondentes, rc_1, rc_2, \dots, rc_5 , digamos. Os valores de r_1, r_2, r_3 (que correspondem aos sinais de downmix 712a) podem então ser calculados a partir de rc_1, rc_2, \dots, rc_5 , em combinação com o esquema de downmix. Quando múltiplos dos canais 718a correspondem ao mesmo sinal de downmix 712a, os coeficientes de renderização de diálogo podem ser somados. Por exemplo, no exemplo ilustrado, temos que $r_1 = rc_1, r_2 = rc_2 + rc_3$ e $r_3 = rc_4 + rc_5$. Esta pode também ser uma soma ponderada no caso em que o downmix dos canais foi feito ao usar coeficientes de downmix.

[000115] Deve ser observado que, também neste caso, o componente de realce de diálogo 706 pode empregar um sinal de áudio adicionalmente recebido que representa o diálogo. Em tal caso, o sinal de diálogo realçado predito 719 pode ser ponderado em conjunto com o sinal de áudio que representa o diálogo antes de ser inserido no componente de mixagem 708 ($X_d = (1 - \alpha_c) \cdot M_d \cdot X'_u + \alpha_c \cdot g \cdot D_c$). A ponderação apropriada é fornecida por um parâmetro de mixagem α_c incluído nos parâmetros para o realce de diálogo 716. O parâmetro de mixagem α_c indica como as contribuições de ganho devem ser divididas entre o componente predito 719 (tal como descrito acima) e o sinal de áudio adicional que representa o diálogo D_c . Isto é análogo ao que foi descrito com respeito ao terceiro modo de realce de diálogo quando combinado com o segundo modo de realce de diálogo.

[000116] Na Figura 7, o decodificador é equipado com um componente de reembaralhamento 730 opcional. O componente de reembaralhamento 730 pode ser usado para a conversão entre esquemas de downmix diferentes, por exemplo, para converter o esquema 100b no

esquema 100a. Deve ser observado que o componente de reembaralhamento 730 deixa tipicamente os sinais c e lfe inalterados, isto é, ele age como um componente passante com respeito a esses sinais. O componente de reembaralhamento 730 pode receber e operar (não mostrado) com base em vários parâmetros tais como, por exemplo, os parâmetros de reconstrução 714 e os parâmetros para o realce de diálogo 716.

[000117] O texto acima foi explicado principalmente com respeito a uma configuração de canais 7.1+4 e um downmix 5.1. No entanto, deve ser compreendido que os princípios dos decodificadores e dos métodos de decodificação descritos no presente documento se aplicam igualmente bem a outras configurações de canal e de downmix.

[000118] A Figura 8 é uma ilustração de um codificador 800 que pode ser usado para codificar uma pluralidade de canais 818, alguns dos quais incluem o diálogo, a fim de produzir um fluxo de dados 810 para a transmissão a um decodificador. O codificador 800 pode ser usado com qualquer um dos decodificadores 200, 500, 600, 700. O codificador 800 compreende um componente de downmix 805, um componente de codificação de realce de diálogo 806, um componente de codificação paramétrico 804 e um componente de transmissão 802.

[000119] O codificador 800 recebe uma pluralidade de canais 818, por exemplo, aqueles das configurações 100a, 100b mostrados nas Figuras 1a e 1b.

[000120] O componente de downmix 805 efetua o downmix da pluralidade de canais 818 em uma pluralidade de sinais de downmix 812, os quais são alimentados então no componente de transmissão 802 para a inclusão no fluxo de dados 810. A pluralidade de canais 818 pode, por exemplo, ser submetida ao downmix de acordo com um esquema de downmix, tal como aquele ilustrado na Figura 1a ou na Figura 1b.

[000121] A pluralidade de canais 818 e os sinais de downmix 812 são inseridos no componente de codificação paramétrico 804. Com base nos seus sinais de entrada, o componente de codificação paramétrico 804 calcula os parâmetros de reconstrução 814 que permitem a reconstrução dos canais 818 dos sinais de downmix 812. Os parâmetros de reconstrução 814 podem, por exemplo, ser calculados ao usar algoritmos de otimização de erro de quadrado médio mínimo (MMSE) tal como é sabido no estado da técnica. Os parâmetros de reconstrução 814 são alimentados então no componente de transmissão 802 para a inclusão no fluxo de dados 810.

[000122] O componente de codificação de realce de diálogo 806 calcula os parâmetros para o realce de diálogo 818 com base em um ou mais canais da pluralidade de canais 818 e um ou mais sinais de diálogo 813. Os sinais de diálogo 813 representam o diálogo puro. De modo marcante, o diálogo já é mixado em um ou mais dos canais 818. Nos canais 818 pode haver desse modo um ou mais componentes de diálogo que correspondem aos sinais de diálogo 813. Tipicamente, o componente de codificação de realce de diálogo 806 calcula os parâmetros para o realce de diálogo 816 ao usar algoritmos de otimização de erro de quadrado médio mínimo (MMSE). Tais algoritmos podem prover os parâmetros que permitem a predição dos sinais de diálogo 813 de alguns canais da pluralidade de canais 818. Os parâmetros para o realce de diálogo 816 podem desse modo ser definidos com respeito a um subconjunto da pluralidade de canais 818, ou seja, aqueles dos quais os sinais de diálogo 813 podem ser preditos. Os parâmetros para a predição de diálogo 816 são alimentados no componente de transmissão 802 para a inclusão no fluxo de dados 810.

[000123] Em conclusão, o fluxo de dados 810 compreende desse modo pelo menos a pluralidade de sinais de downmix 812, os parâmetros de reconstrução 814 e os parâmetros para o realce de diálogo

816.

[000124] Durante a operação normal do decodificador, os valores dos parâmetros de tipos diferentes (tais como os parâmetros para o realce de diálogo, ou os parâmetros de reconstrução) são recebidos repetidamente pelo decodificador a determinadas taxas. Se as taxas em que os valores de parâmetro diferentes são recebidos forem mais baixas do que a taxa na qual a saída do decodificador deve ser calculada, os valores dos parâmetros podem ter que ser interpolados. Se o valor de um parâmetro genérico p for conhecido, nos pontos t_1 e t_2 no tempo, como igual a $p(t_1)$ e $p(t_2)$, respectivamente, o valor $p(t)$ do parâmetro em um tempo intermediário $t_1 \leq t < t_2$ pode ser calculado ao usar esquemas de interpolação diferentes. Um exemplo de tal esquema, indicado no presente documento como um padrão de interpolação linear, pode calcular o valor intermediário ao usar a interpolação linear, por exemplo, $p(t) = p(t_1) + [p(t_2) - p(t_1)](t - t_1)/(t_2 - t_1)$. Um outro padrão, indicado no presente documento como um padrão de interpolação constante por partes, pode incluir de preferência a manutenção de um valor de parâmetro fixado em um dos valores conhecidos durante todo o intervalo de tempo, por exemplo, $p_t = p(t_1)$, ou uma combinação dos valores conhecidos, tal como, por exemplo, o valor médio. A informação sobre qual esquema de interpolação deve ser usado para um certo tipo de parâmetro durante um determinado intervalo de tempo pode ser construída no decodificador, ou fornecida ao decodificador de maneiras diferentes, tal como junto com os próprios parâmetros ou como informação adicional contida no sinal recebido.

[000125] Em um exemplo ilustrativo, um decodificador recebe valores de parâmetro para um primeiro tipo e um segundo tipo de parâmetro. Os valores recebidos de cada tipo de parâmetro são exatamente aplicáveis em um primeiro ($T1 = \{t_{11}, t_{12}, t_{13} \dots\}$) e um segundo ($T2 = \{t_{21}, t_{22}, t_{23} \dots\}$) conjuntos de instantes de tempo, respectivamente, e

o decodificador também tem acesso à informação sobre como os valores de cada tipo de parâmetro devem ser interpolados no caso que um valor tem que ser estimado em um instante de tempo não presente no conjunto correspondente. Os valores de parâmetro controlam as propriedades quantitativas de operações matemáticas nos sinais, em que as operações podem, por exemplo, ser representadas como matrizes. No exemplo a seguir, supõe-se que a operação controlada pelo primeiro tipo de parâmetro esteja representada por uma primeira matriz A, a operação controlada pelo segundo tipo do parâmetro esteja representada por uma segunda matriz B, e os termos "operação" e "matriz" podem ser usados intercambiavelmente no exemplo. Em um instante de tempo em que um valor de saída do decodificador tem que ser calculado, uma operação de processamento em comum que corresponde à composição de ambas as operações deve ser computada. Se também for suposto que a matriz A é a operação de upmix (controlada pelos parâmetros de reconstrução) e que a matriz B é a operação de aplicação de realce de diálogo (controlada pelos parâmetros para o realce de diálogo) então, conseqüentemente, a operação de processamento em comum de upmix seguida pelo realce de diálogo é representada pelo produto de matriz BA.

[000126] Os métodos de computação da operação de processamento em comum são ilustrados nas Figuras 9a a 9e, onde o tempo segue ao longo do eixo horizontal e as marcas de ticagem no eixo indicam os instantes de tempo em que uma operação de processamento em comum deve ser computada (instantes de tempo de saída). Nas figuras, os triângulos correspondem à matriz A (que representa a operação de upmix), os círculos correspondem à matriz B (que representa a operação de aplicação de realce de diálogo) e os quadrados correspondem à matriz de operação em comum BA (que representa a operação em comum de upmix seguida pelo realce de diálogo). Os triângulos e os

círculos cheios indicam que a respectiva matriz é conhecida exatamente (isto é, que os parâmetros, que controlam a operação que a matriz representa, são conhecidos exatamente) no instante de tempo correspondente, ao passo que os triângulos e os círculos vazios indicam que o valor da respectiva matriz é predito ou interpolado (ao usar, por exemplo, qualquer um dos padrões de interpolação esboçados acima). Um quadrado cheio indica que a matriz de operação em comum BA foi computada, no instante de tempo correspondente, por exemplo, por um produto da matriz das matrizes A e B, e um quadrado vazio indica que o valor de BA foi interpolado a partir de um instante de tempo anterior. Além disso, as setas tracejadas indicam entre quais instantes de tempo uma interpolação é executada. Finalmente, uma linha horizontal contínua que conecta os instantes de tempo indica que se supõe que o valor de uma matriz seja ser constante por partes nesse intervalo.

[000127] Um método de computação de uma operação de processamento em comum BA, não fazendo uso da presente invenção, é ilustrado na Figura 9a. Os valores recebidos para as operações A e B se aplicam exatamente nos instantes de tempo t_{11} , t_{21} e t_{12} , t_{22} , respectivamente, e para computar a matriz de operação de processamento em comum em cada instante de tempo de saída o método interpola cada matriz individualmente. Para completar cada etapa de antemão no tempo, a matriz que representa a operação de processamento em comum é computada como um produto dos valores preditos de A e B. Aqui, supõe-se que cada matriz deve ser interpolada ao usar um padrão de interpolação linear. Se a matriz A tiver N fileiras e N' colunas, e a matriz B tiver M fileiras e M' colunas, cada etapa de antemão no tempo deve requerer $O(MN'N)$ operações de multiplicação por faixa de parâmetro (a fim de executar a multiplicação de matriz requerida para computar a matriz de processamento em comum BA). Uma densidade

elevada de instantes de tempo de saída, e/ou um grande número de faixas de parâmetros, portanto, riscos (devido à complexidade computacional relativamente alta de uma operação de multiplicação em comparação com uma operação de adição), colocando uma grande demanda nos recursos computacionais. Para reduzir a complexidade computacional, o método alternativo ilustrado na Figura 9b pode ser usado. Com a computação da operação de processamento em comum (por exemplo, ao executar uma multiplicação de matriz) somente então os instantes onde os valores de parâmetro mudam (isto é, onde os valores recebidos são exatamente aplicáveis, em t_{11} , t_{21} e t_{12} , t_{22}), a matriz de operação de processamento em comum BA pode ser interpolada diretamente em vez de interpolar as matrizes A e B separadamente. Ao fazer isso, se as operações forem representadas por matrizes, cada etapa de antemão no tempo (entre os instantes de tempo em que os valores de parâmetro exatos mudam) então irá requerer somente $O(NM)$ operações (para a adição de matriz) por faixa de parâmetro, e a complexidade computacional reduzida irá requerer menos demanda nos recursos computacionais. Além disso, se as matrizes A e B forem tais que $N' > N \times M / (N + M)$, a matriz que representa a operação de processamento em comum BA irá ter menos elementos do que aqueles encontrados nas matrizes individuais A e B combinadas. O método de interpolação da matriz BA diretamente, no entanto, irá requerer que A e B sejam conhecidos nos mesmos instantes de tempo. Quando os instantes de tempo para os quais A é definido são (pelo menos parcialmente) diferentes dos instantes de tempo para os quais B é definido, um método de interpolação incrementado é requerido. Tal método incrementado, de acordo com modalidades exemplificadoras da presente invenção, é ilustrado nas Figuras 9c a 9e. Em conexão com a discussão das Figuras 9a a 9e, supõe-se, para fins de simplificação, que a matriz de operação de processamento em comum BA

seja computada como um produto das matrizes individuais A e B, cada uma das quais foi gerada com na base em valores de parâmetros (recebidos ou preditos/interpolados). Em outras situações, pode ser igualmente ou mais vantajoso computar a operação representada pela matriz BA diretamente a partir dos valores de parâmetro, sem passar através de uma representação como dois fatores de matriz. Em combinação com qualquer uma das técnicas ilustradas com referência às Figuras 9c a 9e, cada uma dessas abordagens se enquadra dentro do âmbito da presente invenção.

[000128] Na Figura 9c, é ilustrada uma situação na qual o conjunto T1 de instantes de tempo para o parâmetro que corresponde à matriz A inclui um valor de tempo t_{12} não presente no conjunto T2 (instantes de tempo para o parâmetro que corresponde à matriz B). Ambas as matrizes devem ser interpoladas ao usar um padrão de interpolação linear, e o método identifica a predição $t_p = t_{12}$ imediato onde o valor da matriz B deve ser predito (ao usar, por exemplo, a interpolação). Depois que o valor tiver sido encontrado, o valor da matriz de operação de processamento em comum BA em t_p pode ser computado ao multiplicar A e B. Para continuar, o método calcula o valor de BA em um instante de tempo adjacente $t_a = t_{11}$, e interpola então BA entre t_a e t_p . O método também pode computar, caso desejado, o valor de BA em um outro instante de tempo adjacente $t_a = t_{13}$, e interpolar BA de t_p para t_a . Mesmo que uma multiplicação de matriz adicional (a $t_p = t_{12}$) seja requerida, o método permite a interpolação da matriz de operação de processamento em comum BA diretamente, ainda reduzindo a complexidade computacional em comparação ao método, por exemplo, na Figura 9a. Tal como indicado acima, a operação de processamento em comum pode ser alternativamente computada diretamente a partir de valores de parâmetros (recebidos ou preditos/interpolados) e não como um produto explícito de duas matrizes que dependem por

sua vez dos respectivos valores de parâmetros.

[000129] No caso precedente, somente o tipo do parâmetro que corresponde a A tinha instantes de tempo que não estavam incluídos entre os instantes do tipo de parâmetro que corresponde a B. Na Figura 9d, é ilustrada uma situação diferente em que o instante de tempo t_{12} está ausente do conjunto T2, e onde o instante de tempo t_{22} está ausente no conjunto T1. Se um valor de BA tiver que ser computado em um instante de tempo intermediário t' entre t_{12} e t_{22} , o método pode prever o valor de B em $t_p = t_{12}$ e o valor de A em $t_a = t_{22}$. Depois de ter computado a matriz de operação processamento em comum BA em ambos os tempos, a BA pode ser interpolada para encontrar o seu valor em t' . De modo geral, o método executa somente multiplicações de matriz em instantes de tempo onde os valores de parâmetro mudam (isto é, nos instantes de tempo nos conjuntos T1 e T2 onde os valores recebidos são aplicáveis exatamente). Nesse ínterim, a interpolação da operação de processamento em comum requer somente adições de matriz que têm menos complexidade computacional do que as suas contrapartes de multiplicação.

[000130] Nos exemplos acima, foi suposto que todos os padrões de interpolação são lineares. Um método para a interpolação quando os parâmetros devem ser inicialmente interpolados ao usar esquemas diferentes também é ilustrado na Figura 9e. Na figura, os valores de parâmetros que correspondem à matriz A são mantidos de modo a serem constante por partes até o instante de tempo t_{12} , onde os valores mudam abruptamente. Se os valores de parâmetros forem recebidos em uma base por quadros, cada quadro pode conter uma sinalização indicando um instante de tempo em que um valor recebido se aplica exatamente. No exemplo, o parâmetro que corresponde a B só recebeu os valores aplicáveis exatamente em t_{21} e t_{22} , e o método pode prever primeiramente o valor de B no instante de tempo t_p imediata-

mente antes de t_{12} . Depois de computar a matriz de operação de processamento em comum BA em t_p , e $t_a = t_{11}$, a matriz BA pode ser interpolada entre t_a e t_p . O método pode então prever o valor de B em um novo instante de predição $t_p = t_{12}$, computar os valores de BA em t_p e $t_a = t_{22}$, e interpolar a BA diretamente entre t_p e t_a . Mais uma vez, a operação de processamento em comum BA foi interpolada através do intervalo, e o seu valor foi encontrado em todos os instantes de tempo de saída. Em comparação com a situação precedente, tal como ilustrado na Figura 9a, onde A e B podem ter sido interpolados individualmente, e BA computada ao multiplicar A e B em cada instante de tempo de saída, um número reduzido de multiplicações de matriz é necessário e a complexidade computacional é reduzida.

Equivalentes, extensões, alternativas e generalidades

[000131] Outras modalidades da presente invenção tornar-se-ão aparentes a um elemento versado no estado da técnica depois de ter estudado a descrição acima. Embora a presente descrição e os desenhos divulguem modalidades e exemplos, a invenção não é restringida a esses exemplos específicos. Numerosas modificações e variações podem ser feitas sem desviar do âmbito da presente invenção, o qual é definido pelas concretizações. Todos os sinais de referência que aparecem nas concretizações não devem ser compreendidos como limitadores do âmbito das mesmas.

[000132] Além disso, as variações das modalidades divulgadas podem ser compreendidas e efetuadas pelo elemento versado no estado da técnica na prática da invenção, a partir de um estudo dos desenhos, da descrição e das concretizações. Nas concretizações, a palavra "compreende" não exclui outros elementos ou etapas, e o artigo indefinido "um" ou "uma" não exclui uma pluralidade. O mero fato que determinadas medidas são recitadas em concretizações mutuamente diferentes não indica que uma combinação dessas medidas não pode

ser usada com vantagem.

[000133] Os sistemas e os métodos divulgados acima podem ser implementados como software, firmware, hardware ou uma combinação destes. Em uma implementação de hardware, a divisão das tarefas entre as unidades funcionais indicadas na descrição acima não corresponde necessariamente à divisão em unidades físicas; por outro lado, um componente físico pode ter múltiplas funcionalidades, e uma tarefa pode ser realizada por vários componentes físicos em cooperação. Determinados componentes ou todos os componentes podem ser implementados como software executado por um processador ou um microprocessador de sinal digital, ou ser implementados como hardware ou como um circuito integrado específico de aplicativo. Tal software pode ser distribuído em meios que podem ser lidos por computador, que podem compreender meios de armazenamento em computador (ou meios não transitórios) e meios de comunicação (ou meios transitórios). Tal como é bem sabido por um elemento versado no estado da técnica, o termo meios de armazenamento em computador inclui meios voláteis e não voláteis, removíveis e irremovíveis implementados em qualquer método ou tecnologia para o armazenamento de informações tais como instruções que podem ser lidas por computador, estruturas de dados, módulos de programa ou outros dados. Os meios de armazenamento em computador incluem, mas sem ficar a eles limitados, RAM, ROM, EEPROM, memória flash ou outras tecnologias de memória, CD-ROM, discos versáteis digitais (DVD) ou outro armazenamento em disco óptico, cassetes magnéticos, fitas magnéticas, armazenamento em disco magnético ou outros dispositivos de armazenamento magnético, ou qualquer outro meio que pode ser usado para armazenar a informação desejada e que pode ser acessado por um computador. Além disso, é bem sabido por um elemento versado no estado da técnica que meios de comunicação incorporam tipicamente

instruções que podem ser lidas por computador, estruturas de dados, módulos de programa ou outros dados em um sinal de dado modulado, tal como uma onda portadora ou um outro mecanismo de transporte, e inclui quaisquer meios de entrega de informações.

REIVINDICAÇÕES

1. Método para realçar diálogo em um decodificador de um sistema de áudio, em que o método é **caracterizado pelo fato de que** compreende as etapas de:

receber uma pluralidade de sinais de downmix (512, 612, 712) que consiste em um downmix de uma pluralidade maior de canais;

receber parâmetros para realce de diálogo (516, 616, 716), em que os parâmetros são definidos com respeito a um subconjunto da pluralidade de canais incluindo canais que compreendem diálogo, em que o subconjunto da pluralidade de canais é submetido a um downmix em um subconjunto (512a, 612a, 712a) da pluralidade de sinais de downmix (512, 612, 712);

receber parâmetros de reconstrução (514, 614, 714) que permitem reconstrução paramétrica de canais que são submetidos ao downmix no subconjunto (512a, 612a, 712a) da pluralidade de sinais de downmix (512, 612, 712);

upmixar somente o subconjunto (512a, 612a, 712a) da pluralidade de sinais de downmix (512, 612, 712) parametricamente com base nos parâmetros de reconstrução (514, 614, 714) a fim de reconstruir apenas um subconjunto (618a, 618b, 718a) da pluralidade de canais incluindo o subconjunto (618a, 718a) da pluralidade de canais com respeito aos quais os parâmetros para realce de diálogo (516, 616, 716) são definidos;

aplicar realce de diálogo ao subconjunto (618a, 718a) da pluralidade de canais com respeito aos quais os parâmetros para realce de diálogo (516, 616, 716) são definidos ao usar os parâmetros para realce de diálogo (516, 616, 716) para prover pelo menos um sinal de diálogo realçado (619, 719); e

prover versões de diálogo realçado (520, 620, 720) do subconjunto (512a, 612a, 712a) da pluralidade de sinais de downmix (512,

612, 712) por mixagem do pelo menos um sinal de diálogo realçado (619, 719) com pelo menos um outro do subconjunto da pluralidade de sinais de downmix (618b, 712a).

2. Método, de acordo com a reivindicação 1, **caracterizado pelo fato de que**, na etapa de upmixar somente o subconjunto (512a, 612a, 712a) da pluralidade de sinais de downmix (512, 612, 712) parametricamente, nenhum sinal descorrelacionado é usado a fim de reconstruir apenas um subconjunto (618a, 618b, 718a) da pluralidade de canais incluindo o subconjunto (618a, 718a) da pluralidade de canais com respeito aos quais os parâmetros para realce de diálogo (516, 616, 716) são definidos.

3. Método, de acordo com a reivindicação 1, **caracterizado pelo fato de que** a mixagem é efetuada de acordo com parâmetros de mixagem (522, 622, 722) que descrevem uma contribuição do pelo menos um sinal de diálogo realçado (619, 719) para as versões de diálogo realçadas (620, 720) do subconjunto (512a, 612a, 712a) da pluralidade de sinais de downmix (512, 612, 712).

4. Método, de acordo com qualquer uma das reivindicações 1 a 3, **caracterizado pelo fato de que** a etapa de upmixar somente o subconjunto (512a, 712a) da pluralidade de sinais de downmix (512, 612, 712) parametricamente compreende reconstruir somente o subconjunto (718a) da pluralidade de canais com respeito aos quais os parâmetros para realce de diálogo (516, 716) são definidos,

em que a etapa de aplicar realce de diálogo compreende predição e realce de um componente de diálogo do subconjunto da pluralidade de canais com respeito aos quais os parâmetros para realce de diálogo são definidos ao usar os parâmetros para realce de diálogo (516, 716) para fornecer o pelo menos um sinal de diálogo realçado (719), e

em que a mixagem compreende mixagem do pelo menos um

signal de diálogo realçado (719) com o subconjunto (712a) da pluralidade de sinais de downmix (512, 612, 712); e/ou

compreendendo ainda receber um sinal de áudio que representa diálogo, em que a etapa de aplicar realce de diálogo compreende aplicar realce de diálogo ao subconjunto (618a, 718a) da pluralidade de canais com respeito aos quais os parâmetros para realce de diálogo (516, 616, 716) são definidos ao usar ainda o sinal de áudio que representa diálogo.

5. Método, de acordo com qualquer uma das reivindicações 1 a 4, **caracterizado pelo fato de que** ainda compreende receber parâmetros de mixagem (522, 622, 722) para mixar o pelo menos um sinal de diálogo realçado (619, 719) com pelo menos um outro sinal do subconjunto da pluralidade de sinais de downmix (618b, 712a).

6. Método, de acordo com a reivindicação 3 ou 5, **caracterizado pelo fato de que** as etapas de upmixar somente o subconjunto (512a, 612a, 712a) da pluralidade de sinais de downmix (512, 612, 712), aplicar realce de diálogo e mixar são executadas como operações de matriz definidas pelos parâmetros de reconstrução (514, 614, 714), pelos parâmetros para realce de diálogo (516, 616, 716) e pelos parâmetros de mixagem (522, 622, 722), respectivamente, e opcionalmente, compreendendo ainda a combinação, por multiplicação de matriz, das operações de matriz que correspondem às etapas de upmixar somente o subconjunto (512a, 612a, 712a) da pluralidade de sinais de downmix (512, 612, 712), aplicar de realce de diálogo e mixar, em uma única operação de matriz antes da aplicação ao subconjunto (512a, 612a, 712a) da pluralidade de sinais de downmix (512, 612, 712).

7. Método, de acordo com qualquer uma das reivindicações 1 a 6, **caracterizado pelo fato de que** os parâmetros de realce de diálogo (516, 616, 716) e os parâmetros de reconstrução (514, 614, 714) são dependentes da frequência e, opcionalmente, em que os

parâmetros para realce de diálogo (516, 616, 716) são definidos com respeito a um primeiro conjunto de faixas de frequência e os parâmetros de reconstrução (514, 614, 714) são definidos com respeito a um segundo conjunto de faixas de frequência, em que o segundo conjunto de faixas de frequência é diferente do primeiro conjunto de faixas de frequência.

8. Método, de acordo com qualquer uma das reivindicações 1 a 7, **caracterizado pelo fato de que:**

valores dos parâmetros para realce de diálogo (516, 616, 716) são recebidos repetidamente e associados com um primeiro conjunto de instantes de tempo ($T1 = \{t_{11}, t_{12}, t_{13}, \dots\}$), nos quais respectivos valores se aplicam exatamente, em que um primeiro padrão de interpolação predefinido (I1) deve ser executado entre instantes de tempo consecutivos; e

valores dos parâmetros de reconstrução (514, 614, 714) são recebidos repetidamente e associados com um segundo conjunto de instantes de tempo ($T2 = \{t_{21}, t_{22}, t_{23}, \dots\}$), nos quais respectivos valores se aplicam exatamente, em que um segundo padrão de interpolação predefinido (I2) deve ser executado entre instantes de tempo consecutivos,

em que o método ainda compreende:

selecionar um tipo de parâmetro que consiste em parâmetros para realce de diálogo (516, 616, 716) ou em parâmetros de reconstrução (514, 614, 714) e de uma maneira tal que o conjunto de instantes de tempo associado com o tipo selecionado compreende pelo menos um instante de predição que é um instante de tempo (t_p) que é ausente do conjunto associado com o tipo não selecionado;

predizer um valor dos parâmetros do tipo não selecionado no instante de predição (t_p);

computar, com base em pelo menos o valor predito dos

parâmetros do tipo não selecionado e um valor recebido dos parâmetros do tipo selecionado, uma operação de processamento em comum que representa pelo menos upmixar somente o subconjunto dos sinais de downmix (512a, 612a, 712a) seguido por realce de diálogo no instante de predição (t_p); e

computar, com base em pelo menos um valor dos parâmetros do tipo selecionado e um valor dos parâmetros do tipo não selecionado, em que pelo menos qualquer um deles é um valor recebido, a dita operação de processamento em comum em um instante de tempo adjacente (t_a) no conjunto associado com o tipo selecionado ou o tipo não selecionado,

em que as ditas etapas de upmixar somente o subconjunto (512a, 612a, 712a) da pluralidade de sinais de downmix (512, 612, 712) e aplicar realce de diálogo são executadas entre o instante de predição (t_p) e o instante de tempo adjacente (t_a) por meio de um valor interpolado da operação de processamento em comum computada e, opcionalmente, em que o tipo selecionado de parâmetros consiste nos parâmetros de reconstrução (514, 614, 714).

9. Método, de acordo com a reivindicação 8, **caracterizado pelo fato de que** a dita operação de processamento em comum no instante de tempo adjacente (t_a) é computada com base em um valor recebido dos parâmetros do tipo selecionado e um valor recebido dos parâmetros do tipo não selecionado.

10. Método, de acordo com a reivindicação 8 ou 9, **caracterizado pelo fato de que**

ainda compreende selecionar, com base nos primeiro e segundo padrões de interpolação, um padrão de interpolação em comum (I3) de acordo com uma regra de seleção predefinida,

em que a dita interpolação das respectivas operações de processamento em comum computadas é de acordo com o padrão de

interpolação em comum e, opcionalmente, em que a regra de seleção predefinida é definida para o caso em que os primeiro e segundo padrões de interpolação são diferentes e, em que, em resposta ao primeiro padrão de interpolação (I1) ser linear e o segundo padrão de interpolação (I2) ser constante por partes, interpolação linear é selecionada como o padrão de interpolação em comum.

11. Método, de acordo com qualquer uma das reivindicações 8 a 10, **caracterizado pelo fato de que** a predição do valor dos parâmetros do tipo não selecionado no instante de predição (t_p) é feita de acordo com o padrão de interpolação para os parâmetros do tipo não selecionado.

12. Método de acordo com qualquer uma das reivindicações 8 a 11, **caracterizado pelo fato de que** a operação de processamento em comum é computada como uma única operação de matriz antes de ser aplicada ao subconjunto da pluralidade de sinais de downmix (512a, 612a, 712a) e, opcionalmente, em que:

a interpolação linear é selecionada como o padrão de interpolação em comum; e

o valor interpolado das respectivas operações de processamento em comum computadas é computado por interpolação de matriz linear.

13. Método, de acordo com qualquer uma das reivindicações 1 a 12, **caracterizado pelo fato de que** a mixagem do pelo menos um sinal de diálogo realçado (619, 719) com pelo menos um outro sinal (618b, 712a) é restringida a uma seleção não completa da pluralidade de sinais de downmix (512, 612, 712).

14. Decodificador (500, 600, 700) para realçar diálogo em um sistema de áudio, em que o decodificador (500, 600, 700) é **caracterizado pelo fato de que** compreende:

um componente de recepção (502, 602, 702) configurado

para receber:

uma pluralidade de sinais de downmix (512, 612, 712) que consistem em um downmix de um pluralidade maior de canais, parâmetros para realce de diálogo (516, 616, 716), em que os parâmetros são definidos com respeito a um subconjunto da pluralidade de canais incluindo canais que compreendem diálogo, em que o subconjunto da pluralidade de canais é submetido a downmix em um subconjunto (512a, 612a, 712a) da pluralidade de sinais de downmix (512, 612, 712), e

parâmetros de reconstrução (514, 614, 714) que permitem reconstrução paramétrica de canais que são submetidos ao downmix no subconjunto (512a, 612a, 712a) da pluralidade de sinais de downmix (512, 612, 712);

um componente de upmix (504, 604, 704) configurado para efetuar upmix somente do subconjunto (512a, 612a, 712a) da pluralidade de sinais de downmix (512, 612, 712) parametricamente com base nos parâmetros de reconstrução (514, 614, 714) a fim de reconstruir apenas um subconjunto (618a, 618b, 718a) da pluralidade de canais incluindo o subconjunto (618a, 718a) da pluralidade de canais com respeito aos quais os parâmetros para realce de diálogo (516, 616, 716) são definidos; e

um componente de realce de diálogo (506, 606, 706) configurado para aplicar realce de diálogo ao subconjunto (618a, 718a) da pluralidade de canais com respeito aos quais os parâmetros para realce de diálogo (516, 616, 716) são definidos ao usar os parâmetros para realce de diálogo (516, 616, 716) para prover pelo menos um sinal de diálogo realçado (619, 719); e

um componente de mixagem (508, 608, 708) configurado para prover versões de diálogo realçado (520, 620, 720) do subconjunto (512a, 612a, 712a) da pluralidade de sinais de downmix (512, 612, 712)

por mixagem do pelo menos um sinal de diálogo realçado (619, 719) com pelo menos um outro sinal do subconjunto da pluralidade de sinais de downmix (618b, 712a).

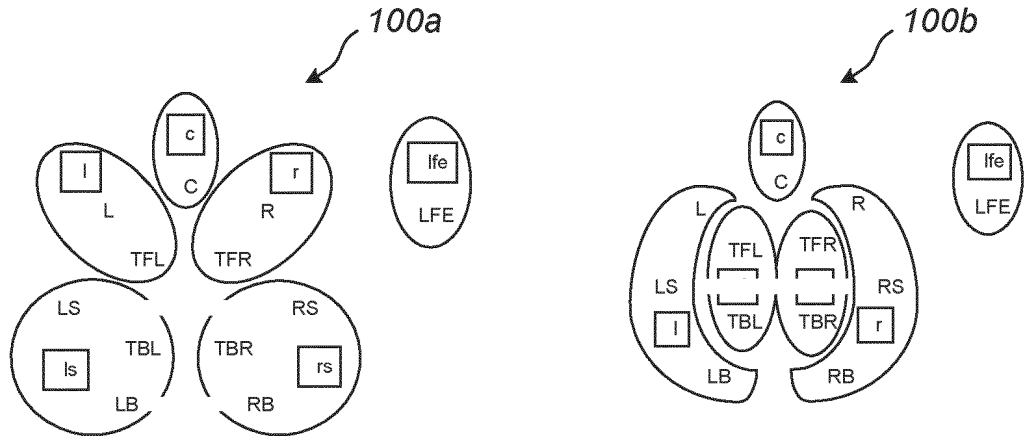


Fig. 1a

Fig. 1b

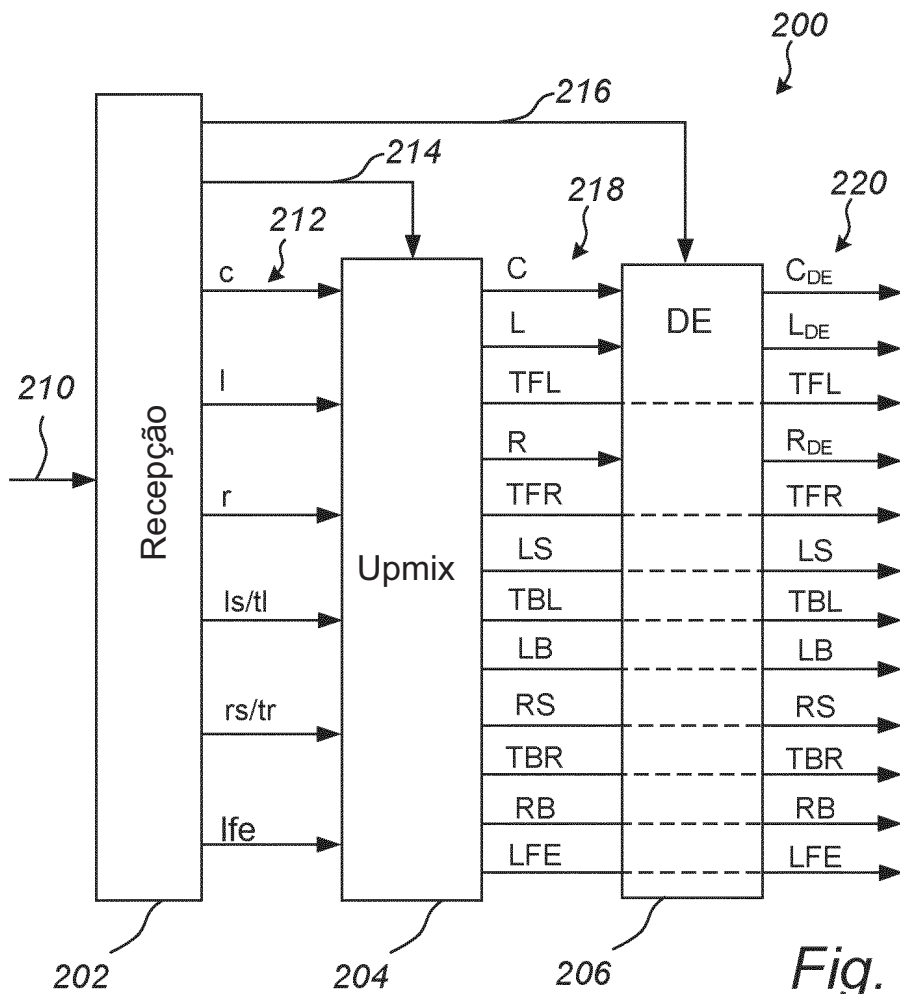


Fig. 2

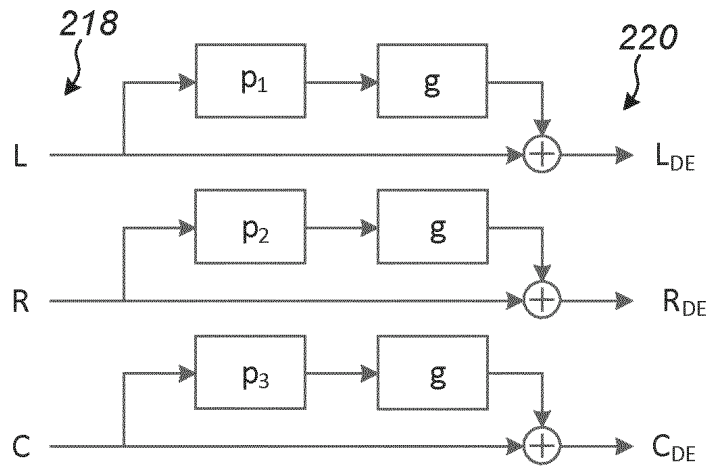


Fig. 3

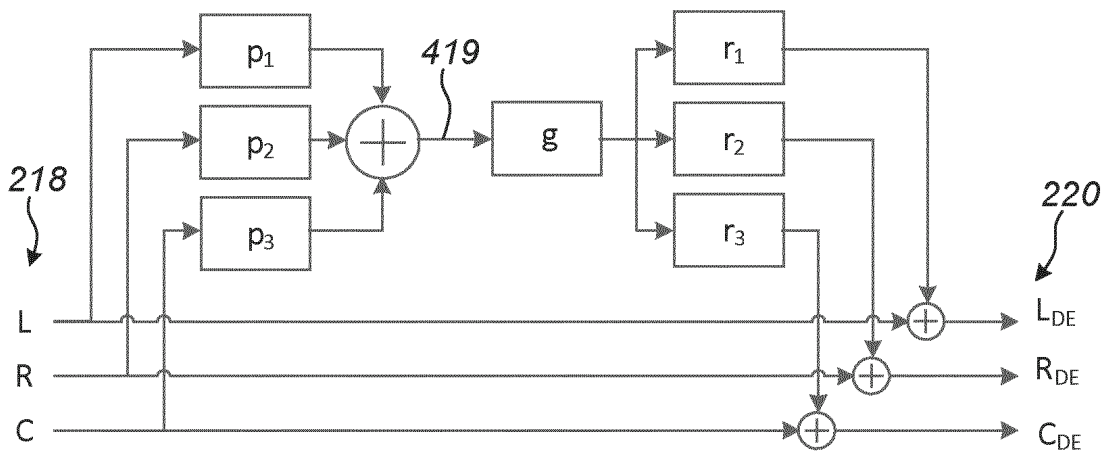


Fig. 4

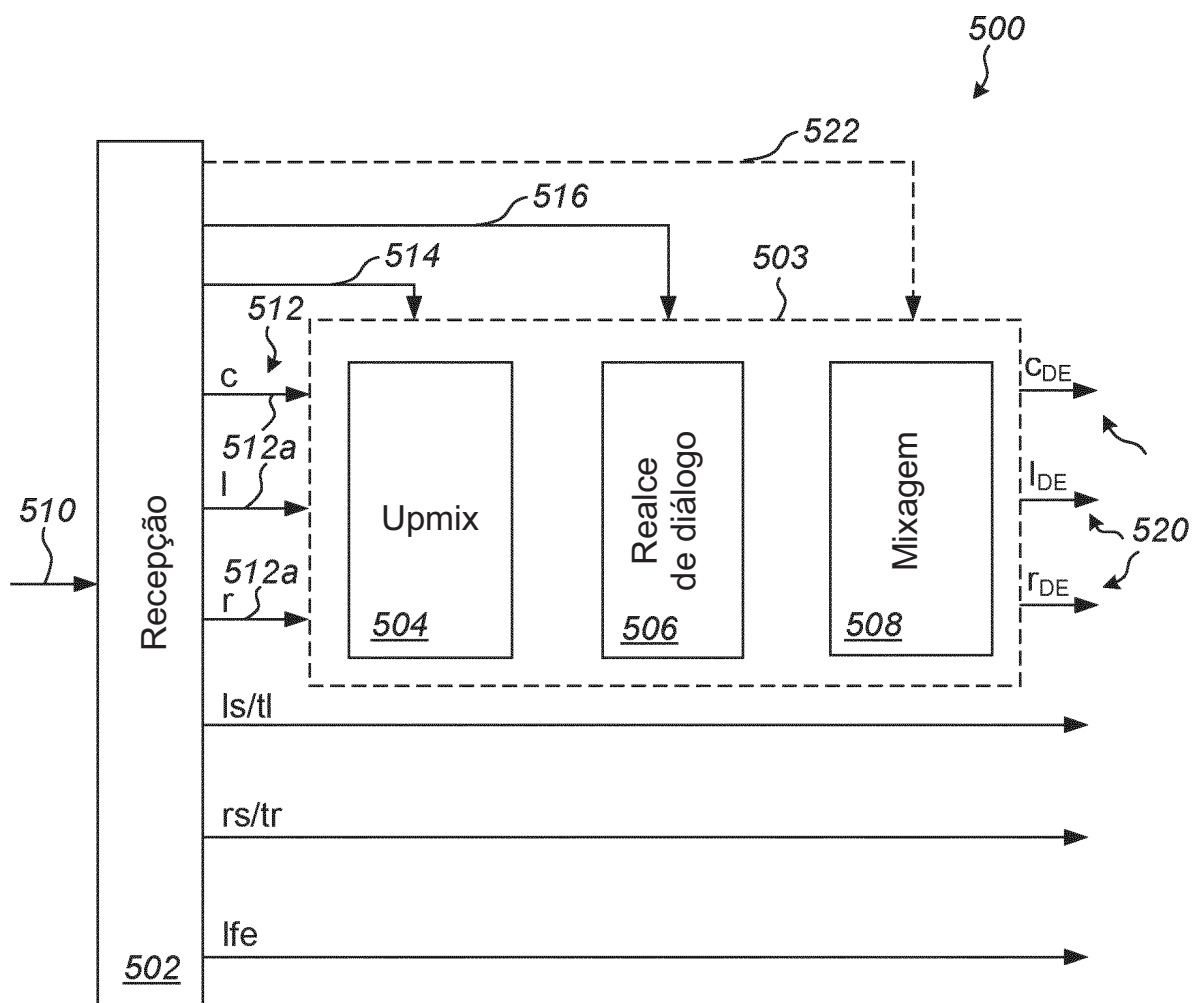


Fig. 5

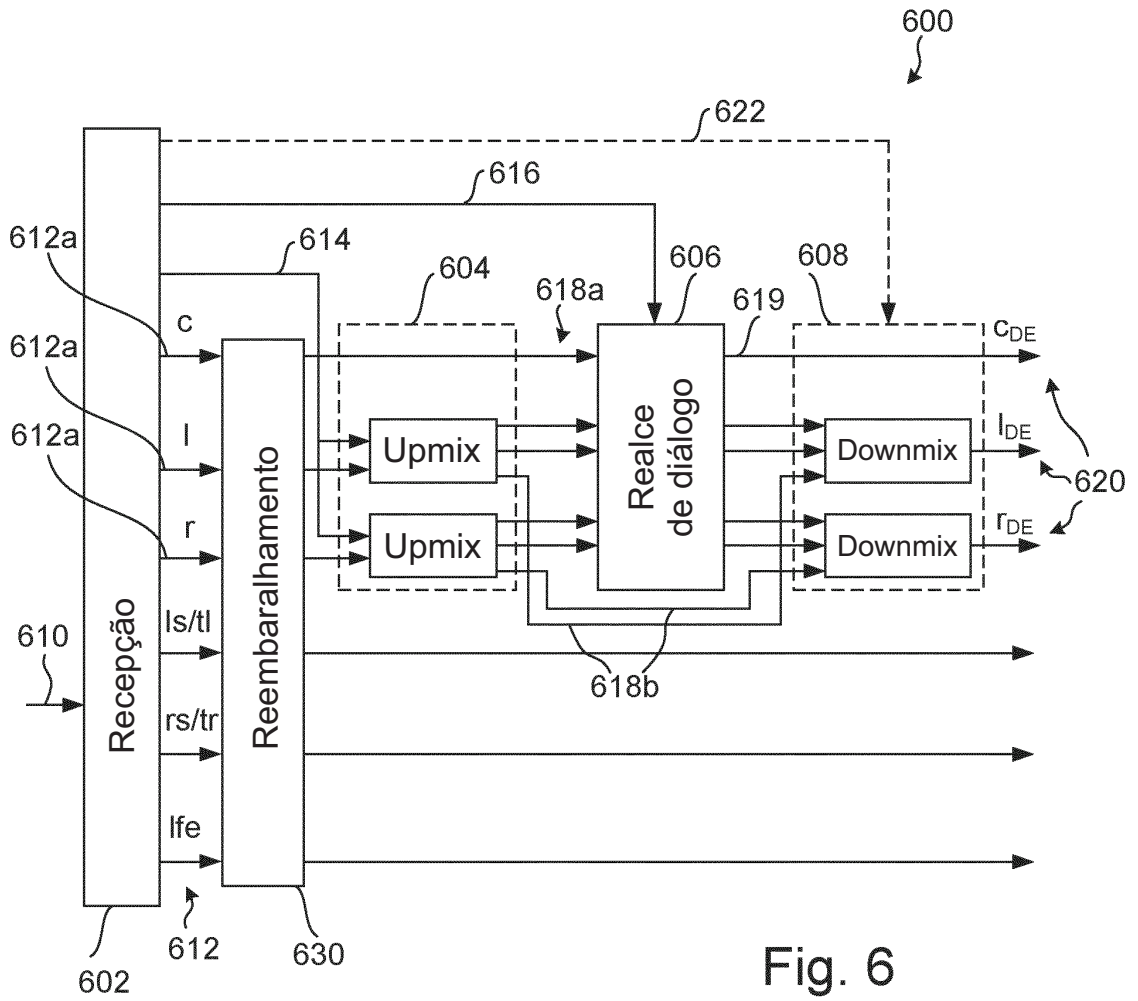


Fig. 6

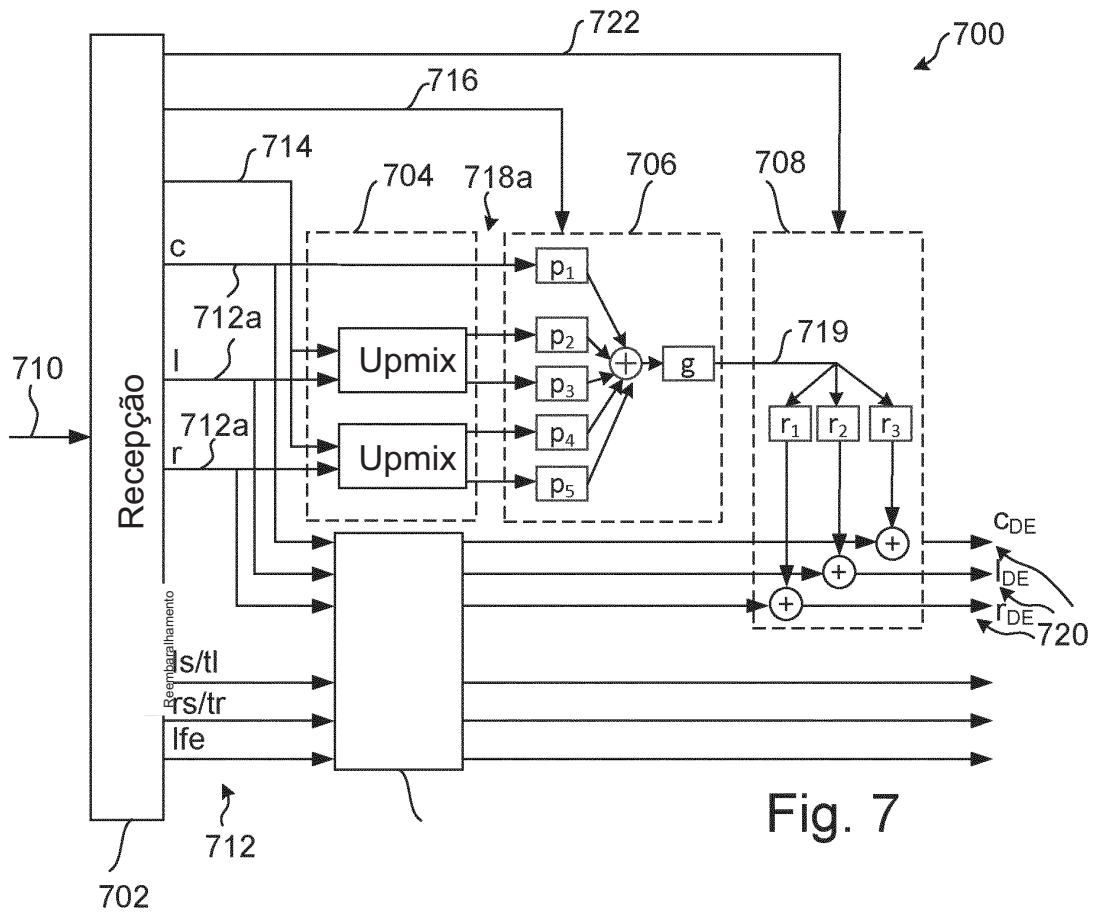


Fig. 7

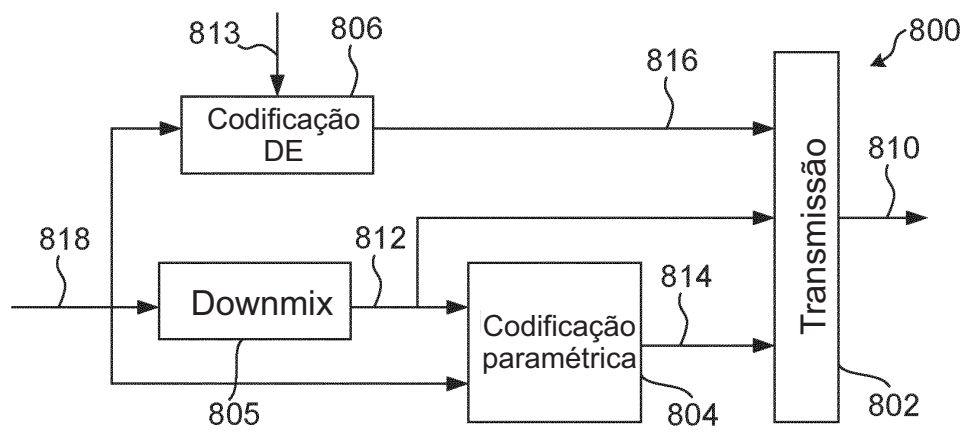


Fig. 8

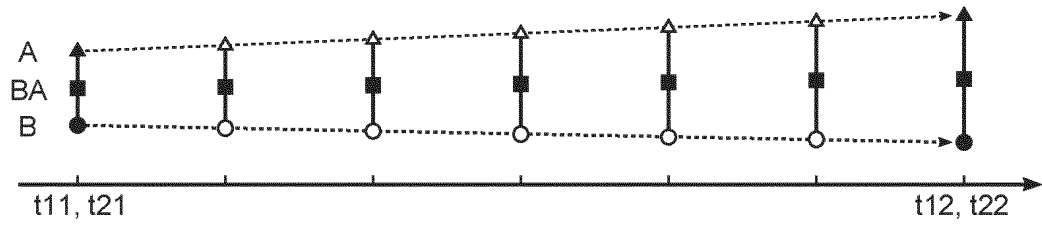


Fig. 9a

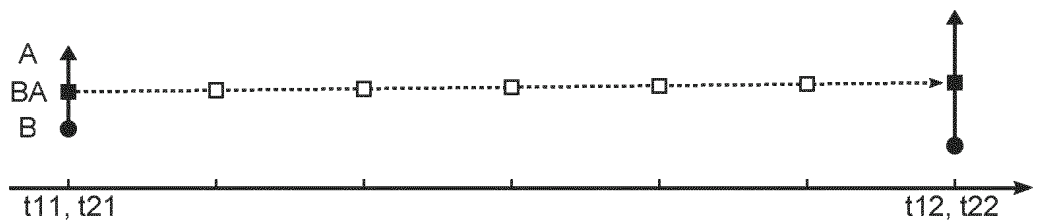


Fig. 9b

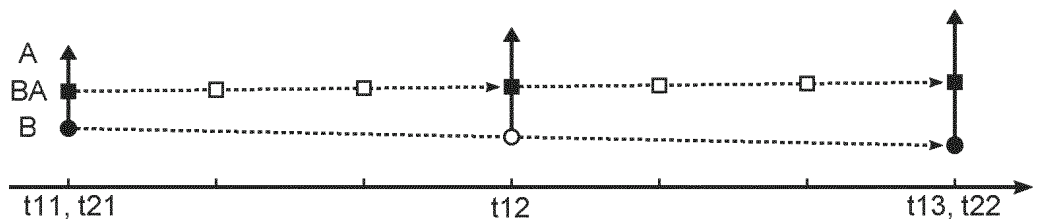


Fig. 9c

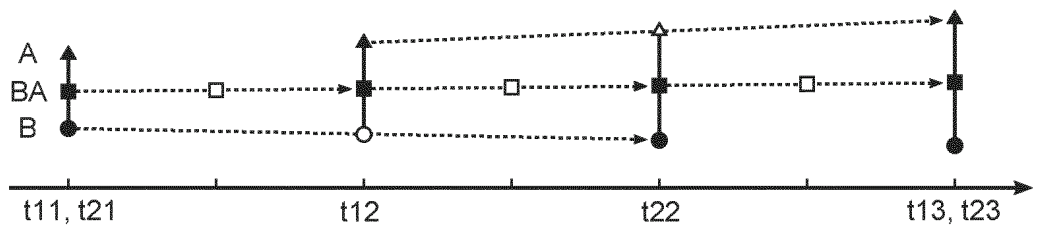


Fig. 9d

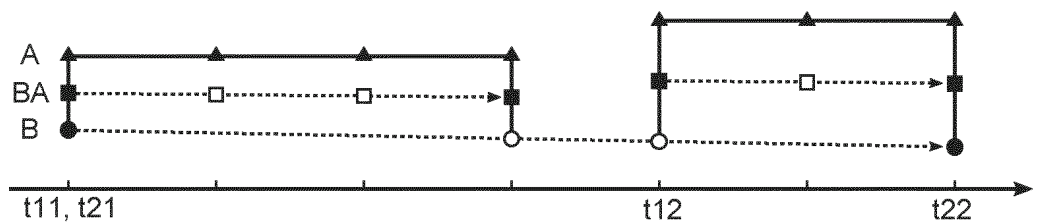


Fig. 9e