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

(11) BR 122019026166-0 B1

(22) Data do Depósito: 06/04/2011

(45) Data de Concessão: 05/01/2021

(54) Título: SISTEMA DECODIFICADOR, APARELHO E MÉTODO PARA EMITIR UM SINAL DE ÁUDIO ESTEREOFÔNICO TENDO UM CANAL ESQUERDO E UM CANAL DIREITO E MEIO LEGÍVEL POR COMPUTADOR NÃO TRANSITÓRIO

(51) Int.Cl.: G10L 19/00.

(30) Prioridade Unionista: 09/04/2010 US 61/322,458.

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

(72) Inventor(es): PONTUS CARLSSON; HEIKO PURNHAGEN; LARS VILLEMOES.

(86) Pedido PCT: PCT EP2011055350 de 06/04/2011

(87) Publicação PCT: WO 2011/124608 de 13/10/2011

(85) Data do Início da Fase Nacional: 10/12/2019

(62) Pedido Original do Dividido: BR112012025878-8 - 06/04/2011

(57) Resumo: A presente invenção refere-se a um método e dispositivos para a codificação e decodificação estereofônicas ao utilizar a predição no domínio da frequência. Em uma modalidade, um método de decodificação, para a obtenção de um sinal estereofônico de saída a partir de um sinal estereofônico de entrada codificado pela codificação de predição complexa e que compreende primeiras representações do domínio da frequência de dois canais de entrada, compreende as etapas de upmix de: (i) computação de uma segunda representação do domínio da frequência de um primeiro canal de entrada; e (ii) computação de um canal de saída com base na primeira e segunda representações do domínio da frequência do primeiro canal de entrada, no primeiro representação do domínio da frequência do segundo canal de entrada e em um coeficiente de predição complexa. A upmix pode ser suspensa em resposta aos dados de controle.

Relatório Descritivo da Patente de Invenção para **"SISTEMA DECODIFICADOR, APARELHO E MÉTODO PARA EMITIR UM SINAL DE ÁUDIO ESTEREOFÔNICO TENDO UM CANAL ESQUERDO E UM CANAL DIREITO E MEIO LEGÍVEL POR COMPUTADOR NÃO TRANSITÓRIO".**

Dividido do BR112012025878-8 depositado em 06 de abril de 2011.

Campo Técnico

[0001] A invenção aqui apresentada refere-se de maneira geral à codificação de áudio estereofônica, e mais precisamente às técnicas para a codificação estereofônica ao usar a predição complexa no domínio da frequência.

Antecedentes da Invenção

[0002] A codificação conjunta dos canais esquerdo (L) e direito (R) de um sinal estereofônico permite uma codificação mais eficiente em comparação à codificação independente de L e R. Uma abordagem comum para a codificação estereofônica conjunta é a codificação do meio/lado (M/S). Aqui, um sinal do meio (M) é formado ao adicionar os sinais L e R, por exemplo, o sinal M pode ter a forma

$$M = (L + R)/2.$$

[0003] Além disso, um sinal lateral (S) é formado ao subtrair os dois canais L e R, por exemplo, o sinal S pode ter a forma

$$S = (L - R)/2.$$

[0004] No caso da codificação de M/S, os sinais M e S são codificados em vez dos sinais L e R.

[0005] No padrão AAC (Advanced Audio Coding) (vide o documento de normas ISOP/IEC 13818-7) do MPEG (Moving Picture Experts Group), a codificação estereofônica L/R e a codificação estereofônica M/S podem ser escolhidas em uma maneira que varia com o tempo e varia com a frequência. Desse modo, o codificador estereofônico pode aplicar a codificação L/R para algumas faixas de frequência do sinal

estereofônico, ao passo que a codificação M/S é usado para codificar outras faixas de frequência do sinal estereofônico (que varia com a frequência). Além disso, o codificador pode comutar com o tempo entre as codificações L/R e M/S (que variam com o tempo). Em MPEG AAC, a codificação estereofônica é executada no domínio da frequência, e mais particularmente o domínio de MDCT (transformada discreta de cosseno modificada). Isto permite escolher de maneira adaptável a codificação tanto L/R quanto M/S em uma frequência e também de maneira variável com o tempo.

[0006] A codificação estereofônica paramétrica é uma técnica para codificar eficientemente um sinal de áudio estereofônico como um sinal monoaural mais uma pequena quantidade de informações colaterais para os parâmetros estereofônicos. Ela faz parte do padrão de áudio MPEG-4 (vide o documento de norma ISO/IEC 14496-3). O sinal monoaural pode ser codificado ao usar qualquer codificador de áudio. Os parâmetros estereofônicos podem ser encaixados na parte auxiliar do fluxo de monobits, desse modo obtendo uma compatibilidade para a frente e para trás total. No decodificador, é o sinal monoaural que é decodificado em primeiro lugar, depois do que o sinal estereofônico é reconstruído com o auxílio dos parâmetros estereofônicos. Uma versão descorrelacionada do sinal mono decodificado, que tem uma correlação cruzada nula com o sinal mono, é gerada por meio de um descorrelacionador, por exemplo, um filtro de passagem ampla apropriado que pode incluir uma ou mais linhas de atraso. Essencialmente, o sinal descorrelacionado tem a mesma distribuição de energia espectral e temporal que o sinal mono. O sinal monoaural em conjunto com o sinal descorrelacionado são inseridos no processo de upmix que é controlado pelos parâmetros estereofônicos e que reconstrói o sinal estereofônico. Para mais informações, vide o documento "Low Complexity Parametric Stereo Codificação in MPEG-4", H. Purnhagen, Proc. of the

7th Int. Conference on Digital Audio Effects (DAFx'04), Nápoles, Itália, 5-8 de outubro de 2004, páginas 163-168.

[0007] MPEG Surround (MPS; vide ISO/IEC 23003-1 e o documento "MPEG Surround - The ISO/MPEG Standard for Efficient and Compatible MultiChannel Audio Coding", J. Herre et al., Audio Engineering Convention Paper 7084, 122nd Convention, 5-8 maio de 2007) permite combinar os princípios da codificação estereofônica paramétrica com a codificação residual, substituir o sinal descorrelacionado com um residual transmitido e desse modo melhorar a qualidade de percepção. A codificação residual pode ser obtida mediante a downmix de um sinal de múltiplos canais e, opcionalmente, ao extrair impulsos apropriados espaciais. Durante o processo de downmix, os sinais residuais que representam o sinal de erro são computados e então codificados e transmitidos. Eles podem tomar o lugar dos sinais descorrelacionados no decodificador. Em uma abordagem híbrida, eles podem substituir os sinais descorrelacionados em determinadas faixas de frequência, de preferência em faixas relativamente baixas.

[0008] De acordo com o sistema Unified Speech and Audio Coding (USAC) do MPEG, dois exemplos do qual são mostrados na figura 1, o decodificador compreende um banco de filtros de espelho de quadratura (QMF) de valor complexo localizado a jusante do decodificador de núcleo. A representação de QMF obtida como saída do banco de filtros é complexa - desse modo superamostrada por um fator dois - e pode ser arranjada como um sinal de downmix (ou, de modo equivalente, sinal do meio) M e um sinal residual D, ao qual uma matriz de upmix com entradas complexas é aplicada. Os sinais de L e R (no domínio de QMF) são obtidos como:

$$\begin{bmatrix} L \\ R \end{bmatrix} = g \begin{bmatrix} 1-\alpha & 1 \\ 1+\alpha & -1 \end{bmatrix} \begin{bmatrix} M \\ D \end{bmatrix}$$

[0009] onde g é um fator de ganho de valor real e α é um coeficiente de predição de valor complexo. De preferência, α é escolhido de

maneira tal que a energia do sinal residual D seja minimizada. O fator de ganho pode ser determinado pela normalização, isto é, para assegurar que a potência do sinal da soma seja igual à soma das potências dos sinais da esquerda e da direita. As partes reais e imaginárias de cada um dos sinais de L e R são mutuamente redundantes - em princípio, cada uma delas pode ser computada com base na outra - mas são benéficas para permitir a aplicação subsequente de um decodificador de replicação de faixa espectral (SBR) sem a ocorrência de artefatos de desalinhamento audíveis. O uso de uma representação de sinal superamostrada também pode, por razões similares, ser escolhido com o objetivo de impedir os artefatos conectados com outro processamento com adaptação do tempo ou da frequência (não mostrado), tal como a upmix mono-estereo. A filtragem QMF inversa é a última etapa de processamento no decodificador. Deve-se observar que a representação de QMF com limitação de banda do sinal permite técnicas residuais com limitação de banda e técnicas "de suficiência residual", que podem ser integradas em decodificadores desse tipo.

[00010] A estrutura de codificação acima é bem adequada para taxas de bits baixas, tipicamente abaixo de 80 kb/s, mas não é ideal para taxas de bits mais elevadas com respeito à complexidade computacional. Mais precisamente, a taxas de bits mais elevadas, a ferramenta de SBR não é tipicamente utilizada (uma vez que não deve melhorar a eficiência da codificação). Então, em um decodificador sem um estágio de SBR, somente a matriz de upmix de avaliação complexa motiva a presença do banco de filtros de QMF, que está exigindo computacionalmente e introduz um atraso (a um comprimento de quadro de 1.024 amostras, o banco de filtros de análise/síntese de QMF introduz um atraso de 961 amostras). Isto indica claramente uma necessidade quanto a uma estrutura de codificação mais eficiente.

Descrição Resumida da Invenção

[00011] Um objetivo da presente invenção consiste na provisão de métodos e aparelhos para a codificação estereofônica que são computacionalmente eficientes também na faixa de taxa de bits elevada.

[00012] A invenção atinge esse objetivo ao prover um codificador e um decodificador, métodos de codificação e decodificando e produtos de programa de computador para a codificação e a decodificação, respectivamente, tal como definido pelas concretizações.

[00013] Em um primeiro aspecto, a invenção provê um sistema decodificador para prover um sinal estereofônico pela codificação estereofônica de predição complexa, em que o sistema decodificador compreende:

[00014] uma upmix adaptada para gerar o sinal estereofônico com base nas primeiras representações do domínio da frequência de um sinal de downmix (M) e de um sinal residual (D), em que cada uma das primeiras representações do domínio da frequência compreende os primeiros componentes espectrais que representam o conteúdo espectral do sinal correspondente expresso em um primeiro subespaço de um espaço multidimensional, em que o estágio de upmix compreende:

[00015] um módulo para computar uma segunda representação do domínio da frequência do sinal de downmix com base na sua primeira representação do domínio da frequência, em que a segunda representação do domínio da frequência compreende segundos componentes espectrais que representam o conteúdo espectral do sinal expresso em um segundo subespaço do espaço multidimensional que inclui uma parte do espaço multidimensional não incluído no primeiro subespaço;

[00016] um somador ponderado para computar um sinal lateral (S) com base na primeira e segunda representações do domínio da frequência do sinal de downmix, na primeira representação do domínio da frequência do sinal residual e em um coeficiente de predição complexo (α) codificado no sinal do fluxo de bits; e

[00017] um estágio de soma e diferença para computar o sinal estereofônico com base na primeira representação do domínio da frequência do sinal de downmix e do sinal lateral,

[00018] em que o estágio de upmix também é operável em um modo passante, em que os ditos sinais de downmix e residuais são fornecidos à soma e diferença diretamente.

[00019] Em um segundo aspecto, a invenção provê um sistema codificador para codificar um sinal estereofônico por um sinal do fluxo de bits pela codificação estereofônica de predição complexa, o qual inclui:

[00020] um estimador para estimar um coeficiente de predição complexo;

[00021] um estágio de codificação operável para:

(a) transformar o sinal estereofônico em uma representação do domínio da frequência de uma downmix e um sinal residual, em uma relação determinada pelo valor do coeficiente da predição complexa; e

[00022] um multiplexador para receber a saída a partir do estágio de codificação e do estimador e codificar esta pelo dito sinal do fluxo de bits.

[00023] Em um terceiro e um quarto aspectos da invenção, são providos métodos para codificar um sinal estereofônico em um fluxo de bits e para decodificar um fluxo de bits em pelo menos um sinal estereofônico. As características técnicas de cada método são análogas àquelas do sistema codificador e do sistema decodificador, respectivamente. Em um quinto e um sexto aspectos, a invenção provê ainda um produto de programa de computador que contém instruções para a execução de cada um dos métodos em um computador.

[00024] A invenção se beneficia das vantagens da codificação estereofônica unificada no sistema MPEG USAC. Essas vantagens também são preservadas a taxas de bits mais elevadas, em que o SBR

não é tipicamente utilizado, sem o aumento significativo na complexidade computacional que deve acompanhar uma abordagem à base de QMF. Isto é possível porque a transformação MDCT criticamente amostrada, que é a base do sistema de codificação de transformação MPEG USAC, pode ser usada para a codificação estereofônica de predição complexa tal como provido pela invenção, pelo menos nos casos em que as larguras de faixa de áudio de código dos canais de downmix e residuais são as mesmas e o processo de upmix não inclui a descorrelação. Isto significa que uma transformação QMF adicional não é requerida mais. Uma implementação representativa da codificação estereofônica de predição complexa no domínio de QMF deve aumentar realmente o número de operações por unidade de tempo de maneira significativa em comparação a estéreo L/R ou M/S tradicional. Desse modo, o aparelho de codificação de acordo com a invenção parece ser competitivo a tais taxas de bits, propiciando uma alta qualidade de áudio a um custo computacional moderado.

[00025] Tal como pode ser concluído por um elemento versado na técnica, o fato que o estágio de upmix também é operável em um modo passante permite que o decodificador decodifique de maneira adaptável de acordo com a codificação direta ou conjunta convencional ou a codificação de predição complexa, tal como determinado no lado do codificador. Desse modo, nos casos em que o decodificador não pode aumentar positivamente o nível da qualidade além daquele da codificação estereofônica L/R direta convencional ou da codificação estereofônica M/S conjunta, pode pelo menos garantir que o mesmo nível seja mantido.

[00026] Como uma vantagem em relação ao estéreo codificado com predição à base de QMF, a reconstrução perfeita do sinal é possível (além dos erros de quantização, que podem se tornar arbitrariamente pequenos).

[00027] Desse modo, a invenção provê um aparelho de codificação para a codificação estereofônica à base de transformação pela predição complexa. De preferência, um aparelho de acordo com a invenção não fica limitado à codificação estereofônica de predição complexa, mas também é operável em um regime de codificação estereofônica L/R direta ou de codificação estereofônica M/S conjunta de acordo com os fundamentos da técnica, de modo que seja possível selecionar o método mais apropriado de codificação para uma aplicação particular ou durante um intervalo de tempo particular.

[00028] Uma representação superamostrada (por exemplo, complexa) do sinal, incluindo ambos os ditos primeiro e segundo componentes espectrais, é usada como uma base para a predição complexa de acordo com a invenção, e desse modo, os módulos para computar tal representação superamostrada são arranjados no sistema codificador e no sistema decodificador de acordo com a invenção. Os componentes espectrais referem-se ao primeiro e ao segundo subespaços de um espaço multidimensional, que pode ser o conjunto de funções dependentes do tempo em um intervalo de determinado comprimento (por exemplo, um comprimento de quadro de tempo predefinido) amostrado a uma frequência de amostragem finita. É bem sabido que as funções neste espaço multidimensional particular podem ser aproximadas por uma soma ponderada finita de funções base.

[00029] Com deve ser apreciado por um elemento versado na técnica, um codificador adaptado para cooperar com um decodificador é equipado com módulos equivalentes para prover a representação superamostrada em que a codificação de predição é baseada, de modo a permitir a reprodução fiel do sinal codificado. Tais módulos equivalentes podem ser os módulos idênticos ou similares ou então módulos que têm características de transferência idênticas ou similares. Em particular, os módulos no codificador e no decodificador, respectiva-

mente, podem ser unidades de processamento similares ou dissimilares que executam os respectivos programas de computador respectivos que executam conjuntos equivalentes de operações matemáticas.

[00030] Em algumas modalidades do sistema decodificador ou do sistema codificador, os primeiros componentes espectrais têm valores reais expressos no primeiro subespaço, e os segundos componentes espectrais têm valores imaginários expressos no segundo subespaço. O primeiro e o segundo componentes espectrais formam conjuntamente uma representação espectral complexa do sinal. O primeiro subespaço pode ser a extensão linear de um primeiro conjunto de funções base, ao passo que o segundo subespaço pode ser a extensão linear de um segundo conjunto de funções base, algumas das quais são linearmente independentes do primeiro conjunto de funções base.

[00031] Em uma modalidade, o módulo para computar a representação complexa é uma transformação de real em imaginário, isto é, um módulo para computar partes imaginárias do espectro de um sinal de tempo discreto com base em uma representação espectral real do sinal. A transformação pode ser baseada em relações matemáticas exatas ou aproximadas, tais como fórmulas de análise harmônica ou relações heurísticas.

[00032] Em algumas modalidades do sistema decodificador ou do sistema codificador, os primeiros componentes espectrais podem ser obtidos por uma transformação de domínio do tempo em frequência, de preferência uma transformação de Fourier, de um sinal de domínio do tempo discreto, tal como por uma transformada discreta de cosseno (DCT), uma transformada discreta de cosseno modificada (MDCT), uma transformada discreta de seno (DST), uma transformada discreta de seno modificada (MDST), uma transformação de Fourier rápida (FFT), um algoritmo de Fourier baseado em fator primo, ou algo do gênero. Nos primeiros quatro casos, os segundos componentes es-

pectrais podem então ser obtidos por DST, MDST, DCT e MDCT, respectivamente. Como é bem sabido, a extensão linear dos cossenos que são periódicos no intervalo unitário forma um subespaço que não é contido inteiramente na extensão linear dos senos periódicos no mesmo intervalo. De preferência, os primeiros componentes espectrais podem ser obtidos por MDCT e os segundos componentes espectrais podem ser obtidos por MDST.

[00033] Em uma modalidade, o sistema decodificador inclui pelo menos um módulo de modelagem de ruído temporal (módulo de TNS, ou filtro de TNS), que é arranjado a montante do estágio de upmix. Falando de maneira genérica, o uso de TNS aumenta a qualidade de áudio percebida para os sinais com componentes do tipo transiente, e isto também se aplica às modalidades do sistema decodificador da invenção que executam TNS. Na codificação estereofônica L/R e M/S convencional, o filtro de TNS pode ser aplicado como uma última etapa de processamento no domínio da frequência, diretamente antes da transformação inversa. No caso da codificação estereofônica de predição complexa, no entanto, é frequentemente vantajoso aplicar o filtro de TNS nos sinais de downmix e residuais, isto é, antes da matriz de upmix. Apresentado de uma maneira distinta, a TNS é aplicada às combinações lineares dos canais esquerdo e direito, o que tem várias vantagens. Em primeiro lugar, pode acarretar uma determinada situação em que a TNS só é benéfica, por exemplo, para o sinal da downmix. Então, a filtragem de TNS pode ser suprimida ou omitida para o sinal residual e, o que pode significar um uso mais econômico da largura de faixa disponível, os coeficientes do filtro de TNS só precisam ser transmitidos para o sinal de downmix. Em segundo lugar, a computação da representação superamostrada do sinal de downmix (por exemplo, os dados de MDST são derivados dos dados de MDCT de modo a formar uma representação complexa do domínio da frequên-

cia), que é necessária na codificação de predição complexa, pode requerer que uma representação do domínio de tempo do sinal de downmix seja computável. Isto significa por sua vez que o sinal de downmix fica de preferência disponível como uma sequência do tempo dos espectros de MDCT obtidos de uma maneira uniforme. Se o filtro de TNS for aplicado no decodificador depois da matriz de upmix, que converte uma representação de downmix/residual em uma representação esquerda/direita, somente uma sequência de espectros MDCT residuais de TNS do sinal de downmix deve estar disponível. Isto deve tornar o cálculo eficiente dos espectros de MDST correspondentes muito desafiador, especialmente se as canais esquerdo e direito estavam usando filtros de TNS com características diferentes.

[00034] É enfatizado que a disponibilidade de uma sequência de tempo de espectros de MDCT não é um critério absoluto a fim de obter uma representação de MDST adequada para servir como uma base para a codificação de predição complexa. Além da evidência experimental, este fato pode ser explicado pela TNS sendo aplicada geralmente somente a frequências mais elevadas, tal como acima alguns quilohertz, de modo que o sinal residual filtrado por TNS corresponda aproximadamente ao sinal residual não filtrado para frequências mais baixas. Desse modo, a invenção pode ser incorporada como um decodificador para a codificação estereofônica de predição complexa, em que os filtros de TNS têm uma colocação diferente do que a montante do estágio de upmix, tal como indicado a seguir.

[00035] Em uma modalidade, o sistema decodificador inclui pelo menos um módulo de TNS adicional localizado a jusante do estágio de upmix. Por meio de um arranjo de seletores, tanto o(s) módulo(s) de TNS a montante do estágio de mixagem quanto o(s) módulo(s) de TNS a jusante do estágio de upmix podem ser ativados seletivamente. Sob determinadas circunstâncias, a computação da representação

complexa do domínio da frequência não requer que uma representação do domínio de tempo do sinal de downmix seja computável. Além disso, tal como indicado acima, o decodificador pode ser seletivamente operável em um modo de codificação direta ou conjunta, não aplicando a codificação de predição complexa, e então pode ser mais apropriado aplicar a localização convencional dos módulos de TNS, isto é, como uma das últimas etapas de processamento no domínio da frequência.

[00036] Em uma modalidade, o sistema decodificador é adaptado para economizar recursos de processamento, e possivelmente energia, através da desativação do módulo para computar uma segunda representação do domínio da frequência do sinal de downmix quando este último não é necessário. Supõe-se que o sinal de downmix esteja dividido em blocos de tempo sucessivos, cada um dos quais é associado com um valor do coeficiente de predição complexo. Esse valor pode ser determinado por uma decisão tomada para cada bloco de tempo por um codificador que coopera com o decodificador. Além disso, nesta modalidade, o módulo para computar uma segunda representação do domínio da frequência do sinal de downmix é adaptado para se desativar no caso em que, para um determinado bloco de tempo, o valor absoluto da parte imaginária do coeficiente de predição complexo for igual a zero ou menor do que uma tolerância predeterminada. A desativação do módulo pode implicar que nenhuma segunda representação do domínio da frequência do sinal de downmix seja computada para esse bloco de tempo. Se a desativação não ocorreu, a segunda representação do domínio da frequência (por exemplo, um conjunto de coeficientes de MDST) deve ser multiplicada por zero ou por um número substancialmente da mesma ordem de magnitude que o épsilon da máquina (unidade de arredondamento) do decodificador ou algum outro valor limite apropriado.

[00037] Em um desenvolvimento adicional da modalidade precedente, a economia de recursos de processamento é conseguida em um subnível do bloco de tempo em que o sinal de downmix é dividido. Por exemplo, tal subnível dentro de um bloco de tempo pode ser uma faixa de frequência, em que o codificador determina um valor do coeficiente de predição complexo para cada faixa de frequência dentro de um bloco de tempo. Similarmente, o módulo para produzir uma segunda representação do domínio da frequência é adaptado para suprimir a sua operação para uma faixa de frequência em um bloco de tempo para o qual o coeficiente de predição complexo é igual a zero ou tem uma magnitude menor do que uma tolerância.

[00038] Em uma modalidade, os primeiros componentes espectrais são coeficientes de transformação arranjados em um ou mais blocos de tempo de coeficientes de transformação, em que cada bloco é gerado pela aplicação de uma transformação a um segmento de tempo de um sinal de domínio do tempo. Além disso, o módulo para computar uma segunda representação do domínio da frequência do sinal de downmix é adaptado para:

• derivar um ou mais primeiros componentes intermediários pelo menos de alguns dos primeiros componentes espectrais;

• formar uma combinação dos ditos um ou mais primeiros componentes espectrais de acordo com pelo menos uma parte de uma ou mais respostas de impulso para obter um ou mais segundos componentes intermediários; e

• derivar os ditos um ou mais segundos componentes espectrais dos ditos um ou mais segundos componentes intermediários.

[00039] Este procedimento resulta em uma computação da segunda representação do domínio da frequência diretamente da primeira representação do domínio da frequência, tal como descrito em mais detalhes no documento de patente US 6.980.933 B2, principalmente

nas colunas 8-28 e particularmente na equação 41 na mesma. Tal como é concluído por um elemento versado na técnica, a computação não é executada através do domínio de tempo, ao contrário, por exemplo, da transformação inversa (por exemplo, IMDCT) seguida por uma transformação diferente (por exemplo, MDST).

[00040] Para uma implementação exemplificadora da codificação estereofônica de predição complexa de acordo com a invenção, foi estimado que a complexidade computacional aumenta apenas ligeiramente (significativamente menos do que o aumento causado pela codificação estereofônica de predição complexa no domínio de QMF) em comparação à codificação estereofônica de L/R ou M/S tradicional. Uma modalidade deste tipo que inclui a computação exata dos segundos componentes espectrais introduz um atraso que é tipicamente somente uma pequena porcentagem mais longo do que aquele introduzido por uma implementação à base de QMF (supondo que o comprimento do bloco de tempo é de 1.024 amostras e comparando com o atraso do banco de filtros de análise/síntese de QMF híbrido, que é de 961 amostras).

[00041] Apropriadamente, em pelo menos algumas das modalidades precedentes, as respostas de impulso são adaptadas para a transformação por meio da qual a primeira representação do domínio da frequência pode ser obtida, e mais precisamente de acordo com as suas características da resposta de frequência.

[00042] Em algumas modalidades, a primeira representação do domínio da frequência do sinal de downmix é obtida por uma transformação que esteja sendo aplicada em relação a uma ou mais funções de janela de análise (ou funções de corte, por exemplo, janela retangular, janela de seno, janela derivada de Kaiser-Bessel, etc.), um objetivo da qual é a obtenção de uma segmentação temporal sem introduzir uma quantidade prejudicial de ruído ou mudar o espectro de uma ma-

neira indesejável. Possivelmente, tais funções de janela são parcialmente sobrepostas. Então, de preferência, as características da resposta de frequência da transformação são dependentes das características das ditas uma ou mais funções de janela de análise.

[00043] Ainda com referência às modalidades que caracterizam a computação da segunda representação do domínio da frequência dentro do domínio da frequência, é possível diminuir a carga computacional envolvida ao usar uma segunda representação aproximada do domínio da frequência. Tal aproximação pode ser obtida ao não exigir a informação completa na qual se deve basear a computação. Pelos ensinamentos do documento de patente US 6.980.933 B2, por exemplo, os dados do primeiro domínio da frequência de três blocos do tempo são requeridos para o cálculo exato da segunda representação do domínio da frequência do sinal de downmix em um bloco, ou seja, um bloco contemporâneo do bloco de saída, um bloco precedente e um bloco subsequente. Para a finalidade de codificação de predição complexa de acordo com a presente invenção, as aproximações apropriadas podem ser obtidas ao omitir – ou ao substituir por zero - os dados que emanam do bloco subsequente (por meio do que a operação do módulo pode se tornar causal, isto é, não contribui com um atraso) e/ou do bloco precedente, de modo que a segunda representação do domínio da frequência só é baseada nos dados de um ou dois blocos de tempo. Deve-se observar que muito embora a omissão dos dados de entrada possa implicar no reescalonamento da segunda representação do domínio da frequência - no sentido que, por exemplo, ela não representa mais a mesma potência – ainda pode ser usada como uma base para a codificação de predição complexa contanto que seja computada de uma maneira equivalente em ambas extremidades do codificador e do decodificador, tal como observado acima. Certamente, um possível reescalonamento desse tipo será compensado por uma mu-

dança correspondente do valor do coeficiente de predição.

[00044] Ainda um outro método aproximado para computar um componente espectral que faz parte da segunda representação do domínio da frequência do sinal de downmix pode incluir a combinação de pelo menos dois componentes da primeira representação do domínio da frequência. Estes últimos componentes podem ser adjacentes com respeito ao tempo e/ou à frequência. Como uma alternativa, eles podem ser combinados pela filtragem de resposta de impulso finita (FIR), com relativamente poucas derivações. Por exemplo, em um sistema que aplica um tamanho de bloco de tempo de 1024, tais filtros de FIR podem incluir 2, 3, 4, etc., derivações. As descrições de métodos de computação aproximada dessa natureza podem ser encontradas, por exemplo, no documento de patente US 2005/0197831 A1. Se uma função de janela que dá pesos relativamente menores à vizinhança de cada limite de bloco de tempo for usada, por exemplo, uma função não retangular, pode ser expediente basear os segundos componentes espectrais apenas em um bloco de tempo nas combinações de primeiros componentes espectrais no mesmo bloco de tempo, implicando que não é a mesma quantidade de informação que está disponível para os componentes mais externos. O erro de aproximação possivelmente introduzido por tal prática é até alguma extensão suprimido ou ocultado pela forma da função de janela.

[00045] Em uma modalidade de um decodificador, que é projetado para emitir um sinal estereofônico do domínio do tempo, é incluída uma possibilidade de comutar entre a codificação estereofônica direta ou conjunta e a codificação de predição complexa. Isto é obtido por meio da provisão de:

• um comutador que seja seletivamente operável tanto como um estágio passante (que não modifica os sinais) quanto como uma transformação de soma e diferença;

• um estágio de transformação inversa para executar uma transformação de frequência em tempo; e

• um arranjo seletor para alimentar o estágio de transformação inversa com um sinal codificado diretamente (ou conjuntamente) ou então com um sinal codificado pela predição complexa.

[00046] Tal como é concluído por um elemento versado na técnica, tal flexibilidade por parte do decodificador propicia ao codificador uma liberdade para escolher entre a codificação direta ou conjunta convencional e a codificação de predição complexa. Desse modo, nos casos em que o nível de qualidade da codificação estereofônica de L/R direta convencional ou da codificação estereofônica de M/S conjunta não pode ser ultrapassado, esta modalidade pode pelo menos garantir que o mesmo nível seja mantido.

[00047] Um outro grupo de modalidade do sistema decodificador efetua a computação dos segundos componentes espectrais na segunda representação do domínio da frequência através do domínio de tempo. Mais precisamente, um inverso da transformação por meio da qual os primeiros componentes espectrais foram obtidos (ou podem ser obtidos) é aplicado e seguido por uma transformação diferente que tem como saída os segundos componentes espectrais. Em particular, uma MDCT inversa pode ser seguida por uma MDST. A fim de reduzir o número de transformações e de transformações inversas, a saída de MDCT inversa pode, em tal modalidade, ser alimentado à MDST e aos terminais de saída (possivelmente precedida por etapas de processamento adicionais) do sistema de decodificação.

[00048] Para uma implementação exemplificadora da codificação estereofônica de predição complexa de acordo com a invenção, foi estimado que a complexidade computacional aumenta apenas ligeiramente (ainda significativamente menos do que o aumento causado pela codificação estereofônica de predição complexa no domínio de

QMF) em comparação a estéreo de L/R ou de M/S tradicional.

[00049] Como um desenvolvimento adicional da modalidade mencionada no parágrafo precedente, o estágio de upmix pode compreender um outro estágio de transformação inversa para processar o sinal lateral. Então, o estágio da soma e diferença é provido com uma representação do domínio do tempo do sinal lateral, gerada pelo dito estágio de transformação adicional, e uma representação do domínio do tempo do sinal de downmix, gerada pelo estágio de transformação inversa já mencionado. Deve-se recordar que, vantajosamente do ponto de vista da complexidade computacional, este último sinal é provido a ambos o estágio da soma e diferença e o dito estágio de transformação diferente (por exemplo, MDST) indicados acima.

[00050] Em uma modalidade, um decodificador projetado para emitir um sinal estereofônico de domínio do tempo inclui uma possibilidade de comutar entre a codificação estereofônica de L/R direta ou a codificação estereofônica de M/S conjunta e a codificação estereofônica de predição complexa. Isto é conseguido mediante a provisão de:

• um comutador operável como um estágio passante ou como um estágio de soma e diferença;

• um estágio de transformação inversa adicional para computar uma representação do domínio de tempo do sinal lateral;

• um arranjo de seletor para conectar os estágios de transformação inversa em um estágio de soma e diferença adicional conectado a um ponto a montante do estágio de upmix e a jusante do comutador (de preferência quando o comutador tiver sido acionado para funcionar como um filtro de passagem, tal como pode ser o caso na decodificação de um sinal estereofônico gerado pela codificação de predição complexa) ou uma combinação de um sinal de downmix do comutador e de um sinal lateral do somador ponderado (de preferência quando o comutador tiver sido acionado para funcionar como um está-

gio de soma e diferença, tal como pode ser o caso na decodificação de um sinal estereofônico diretamente codificado).

[00051] Tal como é concluído por um elemento versado na técnica, isto propicia ao codificador uma liberdade para escolher entre a codificação direta ou conjunta convencional e a codificação de predição complexa, o que significa que um nível de qualidade pelo menos equivalente àquele da codificação estereofônica direto ou conjunta pode ser garantido.

[00052] Em uma modalidade, o sistema codificador de acordo com o segundo aspecto da invenção pode compreender um estimador para estimar o coeficiente de predição complexo com o objetivo de reduzir ou minimizar a potência do sinal médio ou a potência do sinal residual. A minimização pode ocorrer por um intervalo de tempo, de preferência um segmento de tempo ou um bloco de tempo ou um quadro de tempo do sinal a ser codificado. O quadrado da amplitude pode ser tomado como uma medida da potência de sinal momentânea, e uma integral por um intervalo de tempo da amplitude quadrada (forma de onda) pode ser tomada como uma medida da potência média do sinal nesse intervalo. Apropriadamente, o coeficiente de predição complexo é determinado em uma base de bloco de tempo e faixa de frequência, isto é, o seu valor é ajustado de uma maneira tal que reduz a potência média (isto é, a energia total) do sinal residual nesse bloco de tempo e faixa de frequência. Em particular, os módulos para estimar parâmetros de codificação estereofônica paramétricos tais como IID, ICC e IPD ou similares, podem prover a saída na qual o coeficiente de predição complexo pode ser computado de acordo com as relações matemáticas conhecidas por um elemento versado na técnica.

[00053] Em uma modalidade, o estágio de codificação do sistema codificador também é operável para, funcionar como um estágio passante de modo a permitir a codificação estereofônica direta. Ao seleci-

onar a codificação estereofônica direta nas situações em que se espera que propicie uma qualidade maior, o sistema codificador pode garantir que o sinal estereofônico codificado tenha pelo menos a mesma qualidade que aquela na codificação direta. Similarmente, nas situações em que o esforço computacional maior incorrido pela codificação de predição complexa não é motivado por um aumento significativo na qualidade, uma opção de economia de recursos computacionais fica desse modo prontamente disponível ao sistema codificador. A decisão entre a codificação conjunta, direta, de predição real e de predição complexa no codificador é em geral baseada em uma base racional de otimização da taxa/distorção.

[00054] Em uma modalidade, o sistema codificador pode compreender um módulo para computar uma segunda representação do domínio da frequência diretamente (isto é, sem aplicar uma transformação inversa no domínio de tempo e sem usar os dados de domínio do tempo do sinal) com base nos primeiros componentes espectrais. Com relação às modalidades correspondentes do sistema decodificador descrito acima, este módulo pode ter uma estrutura análoga, ou seja, compreender as operações de processamento análogas, mas em uma ordem diferente, de modo que o codificador seja adaptado para emitir os dados apropriados como uma entrada no lado do decodificador. Para as finalidades de ilustração desta modalidade, supõe-se que o sinal estereofônico a ser codificado compreende os canais do meio e laterais, ou foi transformado nessa estrutura, e o estágio de codificação é adaptado para receber uma primeira representação do domínio da frequência. O estágio de codificação compreende um módulo para computar uma segunda representação do domínio da frequência do canal do meio. (A primeira e a segunda representações do domínio da frequência aqui mencionadas são tal como definido acima; em particular as primeiras representações do domínio da frequência podem ser representações de MDCT e

a segunda representação do domínio da frequência pode ser uma representação de MDST). O estágio de codificação compreende ainda um somador ponderado para computar um sinal residual como uma combinação linear formada a partir do sinal lateral e as duas representações do domínio da frequência do sinal do meio ponderado pelas partes reais e imaginárias, respectivamente, do coeficiente de predição complexo. O sinal do meio, ou apropriadamente a sua primeira representação do domínio da frequência, pode ser usado diretamente como um sinal de downmix. Nesta modalidade, além disso, o estimador determina o valor do coeficiente de predição complexo com o objetivo de minimizar a potência ou a potência média do sinal residual. A operação final (otimização) pode ser efetuada tanto pelo controle de retroalimentação, em que o estimador pode receber o sinal residual obtido pelos valores do coeficiente de predição atual a serem ajustados ainda mais caso necessário, ou, de uma maneira de alimentação de avanço, pelas computações efetuadas diretamente nos canais esquerdo/direito de um sinal estereofônico original ou nos canais do meio/lateral. O método de alimentação de avanço é o preferido, por meio do qual o coeficiente de predição complexo é determinado diretamente (em particular, de uma maneira não iterativa ou sem retroalimentação) com base na primeira e segunda representações do domínio da frequência do sinal do meio e na primeira representação do domínio da frequência do sinal lateral. Deve-se observar que a determinação do coeficiente de predição complexo pode ser seguida por uma decisão se deve ser aplicada a codificação direta, conjunta, de predição real ou de predição complexa, em que a qualidade resultante (de preferência a qualidade de percepção, levando em consideração, por exemplo, os efeitos do sinal na máscara) de cada opção disponível é considerada; desse modo as indicações acima não devem ser interpretadas para o efeito que não existe nenhum mecanismo de alimentação no codificador.

[00055] Em uma modalidade, o sistema codificador compreende módulos para computar uma segunda representação do domínio da frequência (ou downmix) do sinal do meio através do domínio de tempo. Deve ser compreendido que os detalhes da implementação relacionados a esta modalidade, pelo menos no que se refere à computação da segunda representação do domínio da frequência, são similares ou podem ser trabalhados analogamente para modalidades correspondentes do decodificador. Nesta modalidade, o estágio de codificação compreende:

• um estágio de soma e diferença para converter o sinal estereofônico em uma forma que compreende os canais do meio e laterais;

• um estágio de transformação para prover uma representação do domínio da frequência do canal lateral e uma representação do domínio da frequência de valor complexo (e desse modo superamostrada) do canal do meio; e

• um somador ponderado para computar um sinal residual, em que o coeficiente de predição complexo é usado como uma ponderação.

[00056] Aqui, o estimador pode receber o sinal residual e determinar, possivelmente em uma forma de controle de retroalimentação, o coeficiente de predição complexo de modo a reduzir ou minimizar a potência ou a média do sinal residual. De preferência, no entanto, o estimador recebe o sinal estereofônico a ser codificado e determina o coeficiente da predição com base no mesmo. É vantajoso do ponto de vista de economia computacional usar uma representação do domínio da frequência criticamente provada do canal lateral, uma vez que este último não será sujeitado à multiplicação por um número complexo nesta modalidade. Apropriadamente, o estágio de transformação pode compreender um estágio de MDCT e um estágio de MDST arranjados em paralelo, ambos os quais têm a representação

do domínio do tempo do canal do meio como entrada. Desse modo, uma representação do domínio da frequência do canal do meio e uma representação do domínio da frequência criticamente amostrada do canal lateral é produzida.

[00057] Deve-se observar que os métodos e o aparelho apresentados nesta seção podem ser aplicados, depois de modificações apropriadas dentro das habilidades de um elemento versado na técnica, incluindo a experimentação rotineira, para a codificação dos sinais que têm mais de dois canais. As modificações em tal operabilidade de múltiplos canais podem prosseguir, por exemplo, ao longo das linhas das seções 4 e 5 no documento da autoria de J. Herre et al. citado acima.

[00058] As características de duas ou mais modalidades esboçadas acima podem ser combinadas, a menos que sejam claramente complementares, em outras modalidades. O fato que duas características são recitadas em uma concretização diferente não impossibilita o fato que elas podem ser combinadas com vantagem. Do mesmo modo, modalidades adicionais também podem ser providas com a omissão de determinadas características que não são necessárias ou não essenciais para a finalidade desejada. Como um exemplo, o sistema de decodificação de acordo com a invenção pode ser incorporado sem um estágio de desquantização nos casos em que o sinal codificado a ser processado não é quantizado nem já se encontra disponível em uma forma apropriada para processamento pelo estágio de upmix.

Breve Descrição dos Desenhos

[00059] A invenção será ilustrada agora ainda pelas modalidades descritas na seção a seguira, sendo feita referência aos desenhos em anexo, nos quais:

[00060] a figura 1 consiste em dois diagramas de blocos generalizados que mostram decodificadores à base de QMF de acordo com os fundamentos da técnica;

[00061] a figura 2 é um diagrama de blocos generalizado de um sistema estereofônico à base de MDCT com predição complexa, de acordo com uma modalidade da presente invenção, em que a representação complexa de um canal do sinal a ser decodificado é computada no domínio da frequência;

[00062] a figura 3 é um diagrama de blocos generalizado de um sistema decodificador estereofônico à base de MDCT com predição complexa, de acordo com uma modalidade da presente invenção, em que a representação complexa de um canal do sinal a ser decodificado é computada no domínio do tempo;

[00063] a figura 4 mostra uma modalidade alternativa do sistema decodificador da figura 2, em que a posição do estágio de TNS ativo é selecionável;

[00064] a figura 5 compreende diagramas de blocos generalizados que mostram sistemas codificadores estereofônicos à base de MDCT com predição complexa, de acordo com modalidade de um outro aspecto da presente invenção;

[00065] a figura 6 é um diagrama de blocos generalizado de um codificador estereofônico à base de MDCT com predição complexa, de acordo com uma modalidade da invenção, em que uma representação complexa de um canal do sinal a ser codificado é computada com base na sua representação do domínio do tempo;

[00066] a figura 7 mostra uma modalidade alternativa do sistema codificador da figura 6, a que também é operável em um modo de codificação de L/R direta;

[00067] a figura 8 é um diagrama de blocos generalizado de um sistema estereofônico à base de MDCT, de acordo com uma modalidade da invenção, em que uma representação complexa de um canal do sinal a ser codificado é computada com base em uma sua primeira representação do domínio da frequência, em que o sistema decodifica-

dor também é operável em um modo de codificação de L/R direta;

[00068] a figura 9 mostra uma modalidade alternativa do sistema codificador da figura 7, a qual inclui ainda um estágio de TNS arranjado a jusante do estágio de codificação;

[00069] a figura 10 mostra modalidades alternativas da parte etiquetada com um A nas figuras 2 e 8;

[00070] a figura 11 mostra uma modalidade alternativa do sistema codificador da figura 8, a qual inclui ainda dois dispositivos modificadores do domínio da frequência respectivamente arranjados a jusante e a montante do estágio de codificação;

[00071] a figura 12 é uma apresentação gráfica dos resultados dos testes de audição a 96 kb/s de seis indivíduos que mostram opções diferentes de alternância de complexidade-qualidade diferente para a computação ou a aproximação do espectro de MDST, em que os pontos de dados etiquetados com um "+" referem-se à referência oculta, "x" refere-se a uma âncora com limite de faixa de 3,5 kHz, "*" refere-se a estéreo tradicional da USAC (M/S ou L/R), "D" refere-se à codificação estereofônica unificada de domínio de MDCT pela predição complexa com a parte imaginária do coeficiente de predição desabilitada (isto é, a predição de valor real, não requerendo nenhuma MDST), "■" refere-se à codificação estereofônica unificada de domínio de MDCT pela predição complexa ao usar um quadro de MDCT atual para computar uma aproximação da MDST, "○" refere-se à codificação estereofônica unificada de domínio de MDCT pela predição complexa ao usar quadros de MDCT atuais e anteriores para computar uma aproximação da MDST, e "●" refere-se à codificação estereofônica unificada de domínio de MDCT por predição complexa ao usar quadros de MDCT atuais, anteriores e próximos para computar a MDST;

[00072] a figura 13 apresenta os dados da figura 12, no entanto como contagens diferenciais em relação à codificação estereofônica

unificada de domínio de MDCT pela predição complexa ao usar um quadro de MDCT atual para computar uma aproximação da MDST;

[00073] a figura 14 compreende os diagramas de blocos generalizados que mostram três modalidades de um sistema decodificador de acordo com as modalidades da invenção;

[00074] a figura 15 é um fluxograma que mostra um método de decodificação de acordo com uma modalidade da invenção; e

[00075] a figura 16 é um fluxograma que mostra um método de codificação de acordo com uma modalidade da invenção.

Descrição Detalhada das Modalidades

I. Sistemas Decodificadores

[00076] A Figura 2 mostra, na forma de um diagrama de blocos generalizado, um sistema da decodificação para decodificar um fluxo de bits que compreende pelo menos um valor de um coeficiente de predição complexo $\alpha = \alpha_R + i\alpha_I$, e uma representação de MDCT de um sinal estereofônico que tem a downmix M e os canais D residuais. As partes real e imaginária α_R, α_I do coeficiente de predição podem ter sido quantizadas e/ou codificadas em conjunto. De preferência, no entanto, as partes real e imaginária são quantizadas independente e uniformemente, tipicamente com um tamanho de etapa de 0,1 (número sem dimensão). A resolução da faixa de frequência usada para o coeficiente de predição complexo não é necessariamente a mesma que a resolução para as faixas de fatores de escala (sfb; isto é, um grupo de linhas de MDCT que estão usando o mesmo tamanho de etapa de quantização de MDCT e a mesma escala de quantização) de acordo com o padrão MPEG. Em particular, a definição de faixa de frequência para o coeficiente de predição pode ser uma que é psicoacusticamente justificada, tal como a escala de Bark. Um demultiplexador 201 é adaptado para extrair essas representações de MDCT e o coeficiente de predição (parte da Informação de Controle tal como in-

dicado na figura) do fluxo de bits que é provido ao mesmo. Certamente, mais informação de controle do que meramente o coeficiente de predição complexo pode ser codificada no fluxo de bits, por exemplo, instruções se o fluxo de bits deve ser decodificada no modo de predição ou de não predição, informações de TNS, etc. As informações de TNS podem incluir valores dos parâmetros de TNS a ser aplicados pelos filtros de TNS (síntese) do sistema decodificador. Se conjuntos idênticos de parâmetros de TNS tiverem de ser usados para vários filtros de TNS, tal como para ambos os canais, é econômico receber essas informações na forma de um bit que indica tal identidade dos conjuntos de parâmetros ao invés de receber os dois conjuntos de parâmetros independentemente. Também podem ser incluídas informações se deve ser aplicado TNS antes ou depois do estágio de upmix, tal como apropriado com base, por exemplo, em uma avaliação psicoacústica das duas opções disponíveis. Além disso, então a informação de controle pode indicar larguras de faixa individualmente limitadas para a downmix e sinais residuais. Para cada canal, as faixas de frequência acima de um limite da largura de faixa não serão decodificadas, mas serão ajustadas em zero. Em determinados casos, as faixas de frequência mais elevadas têm um teor de energia tão baixo que são já são quantizadas até zero. A prática normal (de acordo com o parâmetro max_sfb no padrão MPEG) consiste em usar a mesma limitação da largura de faixa para a downmix e sinais residuais. No entanto, o sinal residual, até uma extensão maior do que o sinal de downmix, tem o seu teor de energia localizado em faixas de frequência mais baixa. Portanto, ao colocar um limite superior da largura de faixa dedicada no sinal residual, uma redução da taxa de bits é possível sem nenhuma perda significativa da qualidade. Por exemplo, isto pode ser regulado por dois parâmetros max_sfb independentes codificados no fluxo de bits, um para o sinal de downmix e um para o sinal residual.

[00077] Nesta modalidade, a representação de MDCT do sinal estereofônico é segmentada em quadros de tempo sucessivos (ou blocos de tempo) que compreendem um número fixo de pontos de dados (por exemplo, 1.024 pontos), um de vários números fixos de pontos de dados (por exemplo, 128 ou 1.024 pontos) ou um número variável de pontos. Tal como é sabido pelos elementos versados na técnica, a MDCT é amostrada de maneira crítica. A saída do sistema de decodificação, indicada na parte à direita do desenho, é um sinal estereofônico de domínio do tempo que tem canais esquerdo L e direito R. Os módulos de desquantização 202 são adaptados para manipular o fluxo de bits inserida no sistema de decodificação ou, onde apropriado, dois fluxos de bits obtidos depois da demultiplexação de um fluxo de bits original e que corresponde a cada uma dentre a downmix e os canais residuais. Os sinais de canais desquantizados são providos a um conjunto de comutação 203 operável em um modo passante ou em um modo de soma e diferença que correspondem às respectivas matrizes de transformação

$$\begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix}$$

e

$$\frac{1}{2}\begin{bmatrix} 1 & 1 \\ 1 & -1 \end{bmatrix}.$$

[00078] Tal como será discutido ainda no parágrafo a seguir, o sistema decodificador inclui um segundo conjunto de comutação 205. Ambos os conjuntos de comutação 203, 205, tal como a maior parte dos outros comutadores e conjuntos de comutação nesta modalidade e nas modalidades a serem descritas, são operáveis em uma forma seletiva de frequência. Isto permite a decodificação de uma grande variedade de modos de decodificação, por exemplo, a decodificação L/R dependente da frequência ou a decodificação L/R ou M/S, tal como conhecido na técnica relacionada.

[00079] Supondo por ora que o conjunto de comutação 203 está no modo passante, os sinais de canais desquantizados são passados, nesta modalidade, através dos respectivos filtros de TNS 204. Os filtros de TNS 204 não são essenciais à operação do sistema de decodificação e podem ser substituídos por elementos passantes. Depois disto, o sinal é provido ao segundo conjunto de comutação 205 que tem a mesma funcionalidade que o conjunto de comutação 203 localizado a montante. Com os sinais de entrada tal como descrito previamente e com o segundo conjunto de comutação 205 ajustado em seu modo passante, a saída do segundo conjunto de comutação 205 é o sinal do canal de downmix e o sinal do canal residual. O sinal de downmix, representado ainda por seus espectros de MDCT sucessivos no tempo, é provido a uma transformação de real em imaginário 206 adaptada para cálculo, com base nos mesmos, os espectros de MDST do sinal de downmix. Nesta modalidade, um quadro de MDST é baseado em três quadros de MDCT, um quadro precedente, um quadro atual (ou contemporâneo) e um quadro subsequente. É indicado simbolicamente (Z^{-1}, Z) que o lado da entrada da transformação de real em imaginário 206 compreende componentes de atraso.

[00080] A representação de MDST do sinal de downmix obtido da transformação de real em imaginário 206 é ponderada pela parte imaginária α_I do coeficiente de predição e adicionada à representação de MDCT do sinal de downmix ponderado pela parte real α_R do coeficiente de predição e à representação de MDCT do sinal residual. As duas adições e multiplicações são executadas pelos multiplicadores e adicionadores 210, 211, formando em conjunto (funcionalmente) um adicionador ponderado, os quais são alimentados com o valor do coeficiente de predição complexo codificado no fluxo de bits inicialmente recebida pelo sistema decodificador. O coeficiente de predição complexo pode ser determinado uma vez para todo quadro de tempo. Também

pode ser determinado mais frequentemente, tal como uma vez para cada faixa de frequência dentro de um quadro, em que as faixas de frequência são uma divisória psicoacusticamente motivada. Também pode ser determinado menos normalmente, tal como será discutido a seguir com relação aos sistemas codificadores de acordo com a invenção. A transformação de real em imaginário 206 é sincronizada com o adicionador ponderado de uma maneira tal que um quadro de MDST atual do sinal do canal de downmix é combinado com os quadros contemporâneos de uma MDCT de cada um dentre o sinal do canal de downmix e do sinal de canal residual. A soma desses três sinais é um sinal lateral S = Re{αM} + D. Nessa expressão, M inclui as representações de MDCT e MDST do sinal de downmix, ou seja, M = M_{MDCT} - iM_{MDST}, ao passo que D = D_{MDCT} tem um valor real. Desse modo, é obtido um sinal estereofônico que tem um canal de downmix e um canal lateral, a partir do qual uma transformação de soma e diferença 207 restaura os canais esquerdo e direito tal como segue:

$$\begin{cases} L = M + S \\ R = M - S \end{cases}$$

[00081] Esses sinais são representados no domínio de MDCT. A última etapa do sistema de codificação consiste na aplicação de uma MDCT inversa 209 a cada um dos canais, por meio do que uma representação do domínio do tempo do sinal estereofônico esquerdo/direito é obtida.

[00082] Uma implementação possível da transformação de real em imaginário 206 é descrita ainda na patente US 6.980.933 B2 do requerente, tal como indicado acima. Pela fórmula 41 na mesma, a transformação pode ser expressa como um filtro de resposta de impulso finito, por exemplo, para pontos uniformes,

$$S(2v) = \frac{2}{N}\sum_{p=0}^{N-1}[(-1)^{p+1}X_I(p) + X_{III}(p)]h_{I,III}(2v-p) + \frac{4}{N}\sum_{l=0}^{N-1}X_{II}(2l+1)h_{II}(2v-2l-1),$$

[00083] onde S(2v) é o 2v-ésimo ponto de dados de MDST, X_I, X_{II}, X_{III} são os dados de MDCT de cada um dos quadros, e N é o comprimento do quadro. Além disso, $h_{I,III}$, h_{II} são respostas de impulsos dependendo da função da janela aplicada ,e portanto, são determinados para cada escolha da função da janela, tal como retangular, sinusoidal e derivada de Kaiser-Bessel, e para cada comprimento do quadro. A complexidade dessa computação pode ser reduzida ao omitir as respostas de impulsos que têm um teor de energia relativamente menor e contribuem relativamente menos para os dados de MDST. Como uma alternativa ou uma extensão para esta simplificação, as próprias respostas de impulsos podem ser encurtadas, por exemplo, do comprimento total N do quadro a um número menor de pontos. Como um exemplo, o comprimento da resposta de impulso pode ser diminuído de 1.024 pontos (derivações) para 10 pontos. O truncamento mais extremo que ainda pode ser considerado como significativo é

$$S(v) = X_{II}(v+1) - X_{II}(v-1).$$

[00084] Outras abordagens diretas podem ser encontradas no documento US de patente 2005/0197831 A1.

[00085] Também é possível reduzir a quantidade de dados de entrada nos quais a computação é baseada. Para ilustrar isto, a transformação de real em imaginário 206 e suas conexões a montante, que são indicadas como uma parte denotada por "A" no desenho, são substituídas por uma variante simplificada, em que duas de tais A' e A" são mostradas na figura 10. A variante A' provê uma representação imaginária aproximada do sinal. Aqui, a computação de MDST só leva em consideração o quadro atual e o quadro precedente. Com referência à fórmula acima neste parágrafo, isto pode ser efetuado ao ajustar $X_{III}(p) = 0$ para p = 0, ... N - 1 (o índice III denota o quadro de tempo subsequente). Devido ao fato que a variante A' não requer o espectro de MDCT do quadro subsequente como entrada, o cálculo de MDST

não incorre em nenhum atraso de tempo. Claramente, essa abordagem reduz um pouco a precisão do sinal de MDST obtido, mas também pode implicar que a energia desse sinal fica reduzida; este último fato pode ser inteiramente compensado por um aumento em α_I como consequência da natureza da codificação da predição.

[00086] Também é mostrada na figura 10 a variante A", que usa como entrada somente os dados de MDCT para o quadro de tempo atual. A variante A" produz de maneira sustentável uma representação de MDST menos precisa do que a variante A'. Por outro lado, opera a um atraso igual a zero, exatamente tal como a variante A', e tem uma complexidade computacional mais baixa. Tal como já foi mencionado, as propriedades de codificação da forma de onda não são afetadas contanto que a mesma aproximação seja usada no sistema codificador e no sistema decodificador.

[00087] Deve ser anotado, independente do fato se a variante A, A' ou A" ou qualquer outro desenvolvimento das mesmas forem usados, que somente as partes do espectro da MDST precisam ser computadas para que a parte imaginária do coeficiente de predição complexo seja diferente de zero, $\alpha I \neq 0$. Em circunstâncias práticas, isto será tomado como significando que o valor absoluto $|\alpha_I|$ da parte imaginária do coeficiente é maior do que um valor limite predeterminado, o qual pode estar relacionado ao arredondamento da unidade do hardware usado. No caso em que a parte imaginária do coeficiente é igual a zero para todas as faixas de frequência dentro de um quadro de tempo, não há nenhuma necessidade de computar quaisquer dados de MDST para esse quadro. Desse modo, apropriadamente, a transformação de real em imaginário 206 é adaptada para responder às ocorrências de valores de $|\alpha_I|$ muito pequenos ao não gerar saída de MDST, por meio do que os recursos de computação podem ser economizados. Nas modalidades em que mais quadros do que o atual são usados para

produzir um quadro de dados de MDST, no entanto, quaisquer unidades a montante da transformação 206 devem continuar operando apropriadamente mesmo que nenhum espectro de MDST seja necessário – em particular, o segundo conjunto de comutação 205 deve manter os espectros de MDCT de avanço - de modo que dados de entrada suficientes já estejam disponíveis para a transformação de real em imaginário 206 já quando ocorrer o próximo quadro de tempo associado com um coeficiente de predição diferente de zero; este pode ser naturalmente ser o próximo bloco de tempo.

[00088] Retornando à figura 2, a função do sistema de decodificação foi descrita sob a suposição de ambos os conjuntos de comutação 203, 205 são ajustados em seus respectivos modos passantes. Tal como será discutido agora, o sistema decodificador também pode decodificar os sinais que não são codificados para predição. Para este uso, o segundo conjunto de comutação 205 será ajustado no seu modo de soma e diferença e apropriadamente, tal como indicado no desenho, um arranjo de seletor 208 será ajustado em sua posição inferior, assegurando desse modo que os sinais sejam alimentados diretamente à transformação inversa 209 de um ponto de origem entre os filtros de TNS 204 e o segundo conjunto de comutação 205. Para assegurar a decodificação correta, o sinal tem apropriadamente a forma de L/R no ponto de origem. Portanto, para assegurar que a transformação de real em imaginário seja provida com o sinal do meio correto (isto é, de downmix) todas as vezes (ao invés de, por exemplo, intermitentemente por um sinal esquerdo), o segundo conjunto de comutação 205 é ajustado de preferência no seu modo de soma e diferença durante a decodificação de um sinal estereofônico não codificado para predição. Tal como indicado acima, a codificação da predição pode ser substituída pela codificação direta ou conjunta convencional para determinados quadros com base, por exemplo, em uma decisão da taxa

de dados/qualidade de áudio. O resultado de tal decisão pode ser comunicado do codificador ao decodificador de várias maneiras, por exemplo, pelo valor de um bit de indicador dedicado em cada quadro, ou pela ausência ou presença de um valor do coeficiente de predição. Tendo sido estabelecidos estes fatos, o papel do primeiro conjunto de comutação 203 pode ser facilmente desempenhado. Certamente, no modo de codificação de não predição, o sistema decodificador pode processar sinais de acordo com a codificação estereofônica (L/R) direta ou a codificação conjunta (M/S), e, com a operação do primeiro conjunto de comutação 203 em modo passante ou em modo de soma e diferença, é possível assegurar que o ponto de origem seja sempre provido com um sinal diretamente codificado. Claramente, o conjunto de comutação 203, quando funcionando como estágio de soma e diferença, irá converter um sinal de entrada na forma M/S em um sinal de saída (fornecido aos filtros de TNS 204 opcionais) na forma L/R.

[00089] O sistema decodificador recebe um sinal se um quadro de tempo particular deve ser decodificado pelo sistema decodificador no modo de codificação de predição ou no modo de codificação sem predição. O modo sem predição pode ser sinalizado pelo valor de um bit de indicador dedicado em cada quadro ou pela ausência (ou valor zero) do coeficiente de predição. O modo de predição pode ser comunicado de maneira análoga. Uma implementação particularmente vantajosa, que permite o recuo sem qualquer extrapolação faz uso de um quarto valor reservado do campo de dois bits ms_mask_present (vide MPEG-2 AAC, documento ISO/IEC 13818-7), que é transmitido por quadro de tempo e definido tal como segue:

Tabela 1: Definição de ms_mask_present em USAC	
Valor	Significado
00	codificação de L/R para todas as faixas de frequência
01	um bit de sinalização por faixa é usado para indicar L/R ou M/S
10	codificação de M/S para todas as faixas de frequência
11	reservado

[00090] Com a redefinição do valor 11 para que signifique "codificação de predição complexa", o decodificador pode ser operado em todos os modos de legado, em particular a codificação de M/S e de L/R, sem nenhuma penalidade da taxa de bits e ainda é capaz de receber um sinal que indica o modo de codificação de predição complexa para os quadros correspondentes.

[00091] A Figura 4 mostra um sistema decodificador da mesma estrutura geral que aquela mostrada na figura 2, mas incluindo, no entanto, pelo menos duas estruturas diferentes. Em primeiro lugar, o sistema da figura 4 inclui os comutadores 404, 411 permitindo a aplicação de alguma etapa de processamento que envolve a modificação do domínio de frequência do estágio de upmix e/ou a jusante. Isso é conseguido, por um lado, por um primeiro conjunto de modificadores domínio da frequência 403 (desenhados como filtros de síntese TNS nessa figura) providos em conjunto com o primeiro comutador 404 a jusante dos módulos de desquantização 401 e um primeiro conjunto de comutação 402, mas a montante de um segundo conjunto de comutação 405 arranjado imediatamente a montante do estágio de upmix 406, 407, 408, 409. Por outro lado, o sistema decodificador inclui um segundo conjunto de modificadores do domínio da frequência 410 provido em conjunto com um segundo comutador 411 a jusante do estágio de upmix 406, 407, 408, 409, mas a montante de um estágio de transformação inversa 412. Vantajosamente, tal como indicado no desenho, cada modificador do domínio da frequência é arranjado em paralelo com uma linha passante que é conectada a montante ao lado de entrada do modificador do domínio da frequência e é conectada a jusante do comutador associado. Em virtude dessa estrutura, o modificador do domínio da frequência é provido com os dados de sinal todas as vezes, permitindo o processamento no domínio da frequência com base em mais quadros de tempo do que apenas aquele atual. A deci-

são quanto ao fato se devem ser aplicados os primeiros 403 ou os segundos conjuntos de modificadores do domínio da frequência 410 pode ser tomada pelo codificador (e conduzida no fluxo de bits), ou pode ser baseada no fato se a codificação da predição é aplicada, ou pode ser baseada em algum outro critério considerado como apropriado em circunstâncias práticas. Como um exemplo, se o modificador do domínio da frequência consistir em filtros de TNS, então o primeiro conjunto 403 pode ser usado de maneira vantajosa para alguns tipos de sinais, ao passo que o segundo conjunto 410 pode ser usado de maneira vantajosa para outros tipos de sinais. Se o resultado dessa seleção for codificado no fluxo de bits, então o sistema decodificador irá ativar o respectivo conjunto de filtros de TNS de maneira correspondente.

[00092] Para facilitar a compreensão do sistema decodificador mostrado na figura 4, observa-se explicitamente que a decodificação de um sinal codificado diretamente (L/R) ocorre quando $\alpha = 0$ (implicando que pseudo-L/R e L/R são idênticos e que os canais laterais e residuais não diferem), o primeiro conjunto de comutação 402 está no modo de passagem, o segundo conjunto de comutação está no modo de soma e diferença, fazendo desse modo com que o sinal tenha a forma de M/S entre o segundo conjunto de comutação 405 e um estágio de soma e diferença 409 do estágio de upmix. Devido ao fato que o estágio de upmix será então uma etapa eficazmente passante, não é importante se (usando os respectivos comutadores 404, 411) é ativado o primeiro ou o segundo modificadores do domínio da frequência.

[00093] A figura 3 ilustra um sistema decodificador de acordo com uma modalidade da invenção que, com relação àquelas das figuras 2 e 4, representa uma abordagem diferente para a provisão de MDST de dados requeridos para upmix. Tal como com os sistemas decodificadores já descritos, o sistema da figura 3 compreende os módulos de desquantização 301, um primeiro conjunto de comutação 302 operável

em um modo passante ou de soma e diferença e os filtros de TNS 303 (síntese), que são todos arranjados em série a partir da extremidade de entrada do sistema decodificador. Os módulos a jusante desse ponto são utilizados seletivamente por meio de dois segundos comutadores 305, 310 que são operados de preferência em conjunto de modo que ambos estejam em suas posições superiores ou posições inferiores, tal como indicado na figura. Na extremidade de saída do sistema decodificador, há um estágio de soma e diferença 302 e, imediatamente a montante do mesmo, dois módulos de MDCT inversa 306, 311 para transformar uma representação do domínio de MDCT de cada canal em uma representação do domínio do tempo.

[00094] Na decodificação de predição complexa, em que o sistema decodificador é provido com um fluxo de bits que codifica um sinal estereofônico de downmix/residual e os valores de um coeficiente de predição complexo, o primeiro conjunto de comutação 302 é ajustado em seu modo passante e os segundos comutadores 305, 310 são ajustados na posição superior. A jusante dos filtros de TNS, os dois canais do sinal estereofônico (desquantizado, filtrado em TNS, MDCT) são processados de maneiras diferentes. O canal de downmix é provido, por um lado, a um multiplicador e somador 308, que adiciona a representação de MDCT do canal de downmix ponderada pela parte real α_R do coeficiente de predição da representação de MDCT do canal residual e, por outro lado, a um módulo 306 dos módulos de transformação MDCT inversa. A representação do domínio de tempo do canal de downmix M, que é emitida pelo módulo de transformação MDCT inversa 306, é provida ao estágio de soma e diferença final 312 e a um módulo de transformação MDST 307. Esse uso duplo da representação do domínio de tempo do canal de downmix é vantajoso do ponto de vista da complexidade computacional. A representação MDST do canal de downmix obtida desse modo é provida a um multiplicador e somador adicional 309

que, depois de ser ponderado pela parte imaginária α_I do coeficiente de predição adiciona esse sinal à combinação linear emitida pelo somador 308; desse modo, a saída do somador 309 é um sinal de canal laterais, $S = Re\{\alpha M\} + D$. Similarmente ao sistema decodificador mostrado na figura 2, os multiplicadores e somadores 308, 309 podem ser combinados de imediato para formar um somador de múltiplos sinais ponderado com entradas das representações de MDCT e MDST do sinal de downmix, a representação de MDCT do sinal residual e o valor do coeficiente de predição complexo. A jusante desse ponto na presente modalidade, somente uma passagem através do módulo de transformação MDCT inversa 311 permanece antes que o sinal do canal laterais seja provido ao estágio de soma e diferença final 312.

[00095] A sincronicidade necessária no sistema decodificador pode ser obtida ao aplicar os mesmos comprimentos de transformação e formas da janela em ambos os módulos de transformação MDCT inversa 306, 311, tal como já é a prática na codificação de M/S e L/R seletiva de frequência. Um atraso de um quadro é introduzido pela combinação de determinadas modalidades do módulo de MDCT inversa 306 e modalidades do módulo de MDST 307. Portanto, cinco blocos de atraso opcionais 313 (ou instruções de software para este efeito em uma implementação de computador) são providos, de modo que a parte do sistema localizada à direita da linha tracejada pode ser atrasada em um quadro com relação à parte da esquerda quando necessário. Aparentemente, todas as interseções entre a linha tracejada e as linhas de conexão são providas com blocos de atraso, com exceção da linha de conexão entre o módulo de MDCT inversa 306 e o módulo de transformação MDST 307, que é onde aparece o atraso que requer compensação.

[00096] A computação dos dados de MDST para um quadro de tempo requer os dados de um quadro da representação do domínio de

tempo. No entanto, a transformação MDCT inversa é baseada em um quadro (atual), dois (de preferência: precedente e atual) ou três (de preferência: precedente, atual e subsequente) quadros consecutivos. Em virtude do cancelamento de serrilhamento (*alias*) no domínio do tempo (TDAC) bem conhecido associado com a MDCT, a opção de três quadros acarreta a sobreposição completa dos quadros de entrada e desse modo provê a melhor (e possivelmente perfeita) exatidão, pelo menos nos quadros que contêm o serrilhamento no domínio do tempo. Claramente, a MDCT inversa de três quadros opera a um atraso de um quadro. Ao aceitar o uso de uma representação do domínio do tempo aproximada como entrada para a transformação MDST, é possível evitar esse atraso, e desse modo a necessidade de compensar atrasos entre partes diferentes do sistema decodificador. Na opção de dois quadros, a TDAC que habilita a sobreposição/adição ocorre na primeira metade do quadro, e o serrilhamento pode estar presente somente na segunda metade. Na opção de um quadro, a ausência de TDAC implica que o serrilhamento pode ocorrer por todo o quadro; no entanto, a representação de MDST obtida dessa maneira, e usada como um sinal intermediário na codificação de predição complexa, também pode prover uma qualidade satisfatória.

[00097] O sistema de decodificação ilustrado na figura 3 também pode ser operado em dois modos de decodificação sem predição. Para decodificar diretamente um sinal estereofônico codificado L/R, os segundos comutadores 305, 310 são ajustados na posição inferior e o primeiro conjunto de comutação 302 é ajustado no modo passante. Desse modo, o sinal tem a forma L/R a montante do estágio de soma e diferença 304, que converte o mesmo na forma M/S, com o que ocorrem a transformação MDCT inversa e uma operação de soma e diferença final. Para decodificar um sinal estereofônico provido na forma codificada M/S conjuntamente, o primeiro conjunto de comutação 302 é

ajustado de preferência em seu modo de soma e diferença, de modo que o sinal tenha a forma L/R entre o primeiro conjunto de comutação 302 e o estágio de soma e diferença 304, que é frequentemente mais apropriado do ponto da vista de filtragem de TNS do que seria uma forma M/S. O processamento a jusante do estágio de soma e diferença 304 é idêntico àquele no caso da decodificação de L/R direta.

[00098] A figura 14 consiste em três diagramas de blocos generalizados dos decodificadores de acordo com modalidade da invenção. Em contraste com vários outros diagramas de blocos que acompanham este pedido de patente, uma linha de conexão na figura 14 pode simbolizar um sinal de múltiplos canais. Em particular, tal linha de conexão pode ser arranjada de modo a transmitir um sinal estereofônico que compreenda canais esquerdo/direito, do meio/laterais, de downmix/residual, pseudo-esquerdo/pseudo-direito, e outras combinações.

[00099] A figura 14A mostra um sistema decodificador para decodificar uma representação do domínio da frequência (indicada, para a finalidade desta figura como uma representação de MDCT) de um sinal de entrada. O sistema decodificador é adaptado para prover como sua saída uma representação do domínio do tempo de um sinal estereofônico, que é gerado com base no sinal de entrada. Para poder decodificar um sinal de entrada codificado pela codificação estereofônica de predição complexa, o sistema decodificador é provido com um estágio de upmix 1410. No entanto, ele também é capaz de manipular um sinal de entrada codificado em outros formatos e possivelmente que alterna entre vários formatos de codificação com o passar do tempo, por exemplo, uma sequência de quadros de tempo codificados pela codificação de predição complexa pode ser seguida por uma parte de tempo codificada pela codificação esquerda/direita direta. A capacidade do sistema decodificador de manipular formatos diferentes de codificação é obtida pela provisão de uma linha de conexão (de pas-

sagem) arranjada em paralela com o dito estágio de upmix 1410. Por meio de um comutador 1411, é possível selecionar se a saída do estágio de upmix 1410 (posição do comutador inferior na figura) ou o sinal não processado disponível sobre a linha de conexão (posição do comutador superior na figura) deve ser provida aos módulos do decodificador arranjados mais a jusante. Nessa modalidade, um módulo de MDCT inversa 1412 é arranjado a jusante do comutador, o qual transforma uma representação de MDCT de um sinal em uma representação do domínio do tempo. Como um exemplo, o sinal provido ao estágio de upmix 1410 pode ser um sinal estereofônico na forma de downmix/residual. O estágio de upmix 1410 é adaptado então para derivar um sinal lateral e para executar a operação de a soma e diferença de modo que o sinal estereofônico esquerdo/direito (no domínio de MDCT) seja emitido.

[000100] A figura 14B mostra um sistema decodificador similar àquele da figura 14A. O presente sistema é adaptado para receber um fluxo de bits em seu sinal de entrada. O fluxo de bits é processada inicialmente por um demultiplexador e por um módulo de desquantização combinado 1420, que provê, como um primeiro sinal de saída, uma representação de MDCT de um sinal estereofônico de múltiplos canais para tratamento adicional, tal como determinado pela posição de um comutador 1422 que tem a funcionalidade análoga àquela do comutador 1411 da figura 14A. Mais precisamente, o comutador 1422 determina se a primeira saída do demultiplexador e a desquantização deve ser processada somente por um estágio de upmix 1421 e um módulo de MDCT inversa 1423 (posição inferior) ou pelo módulo de MDCT inversa 1423 (posição superior). O módulo demultiplexador e de desquantização 1420 também emite informações de controle. No presente caso, a informação de controle associada com o sinal estereofônico pode incluir os dados que indicam se a posição superior ou inferior do comutador 1422 é apropriada

para decodificar o sinal ou, de maneira mais abstrata, de acordo com qual formato de codificação o sinal estereofônico deve ser decodificado. A informação de controle também pode incluir parâmetros para ajustar as propriedades do estágio de upmix 1421, por exemplo, um valor do coeficiente de predição complexo usado na codificação de predição complexa tal como já foi descrito acima.

[000101] A figura 14C mostra um sistema decodificador que, além das entidades análogas àquelas na figura 14B, compreende primeiro e segundo dispositivos modificadores do domínio da frequência 1431, 1435 arranjados respectivamente a montante e a jusante de um estágio de upmix 1433. Para as finalidades dessa figura, cada dispositivo modificador do domínio da frequência é ilustrado por um filtro de TNS. No entanto, pela expressão dispositivo modificador do domínio da frequência também poderiam ser compreendidos outros processos além da filtragem de TNS que são suscetíveis de serem aplicados tanto antes quanto depois do estágio de upmix. Os exemplos de modificações do domínio da frequência incluem a predição, a adição do ruído, a extensão da largura de faixa, e o processamento não linear. As considerações de psicoacústicas e razões similares, que incluem possivelmente as propriedades do sinal a ser processado e/ou a configuração ou os ajustes de tal dispositivo modificador do domínio da frequência,, indicam algumas vezes que é vantajoso aplicar a dita modificação do domínio da frequência a montante do estágio de upmix 1433 e não a jusante. Em outros casos, pode ser estabelecido por considerações similares que a posição a jusante da modificação do domínio da frequência é preferível àquela a montante. Por meio dos comutadores 1432, 1436, os dispositivos modificadores do domínio da frequência 1431, 1435 podem ser seletivamente ativados de modo que, em resposta à informação de controle, o sistema decodificador possa selecionar a configuração desejada. Como um exemplo, a figura 14C mostra

uma configuração em que o sinal estereofônico do demultiplexador e do módulo de desquantização combinado 1430 é inicialmente processado pelo primeiro dispositivo modificador do domínio da frequência 1431, e então provido ao estágio de upmix 1433 e finalmente diretamente encaminhado a um módulo de MDCT inversa 1437, sem passar através do segundo dispositivo modificador do domínio da frequência 1435. Tal como explicado na seção Descrição Resumida, essa configuração é preferida em relação à opção de executar TNS depois da upmix na codificação de predição complexa.

II. Sistemas codificadores

[000102] Um sistema codificador de acordo com a invenção será descrito agora com referência à figura 5, que é um diagrama de blocos generalizado de um sistema codificador para codificar um sinal estereofônico da esquerda/direita (L/R) como um fluxo de bits de saída pela codificação de predição complexa. O sistema codificador recebe uma representação do domínio do tempo ou de domínio da frequência do sinal e envia este para um estágio de downmix e a um estimador do coeficiente da predição. As partes reais e imaginárias dos coeficientes da predição são providas ao estágio de downmix a fim de regular a conversão dos canais esquerdo e direito na downmix e dos canais residuais, que são então providas a um multiplexador final MUX. Se o sinal não for provido como uma representação do domínio da frequência ao codificador, ele é transformado em tal representação no estágio de downmix ou no multiplexador.

[000103] Um dos princípios na codificação da predição consiste na conversão do sinal à esquerda/direita na forma de meio/lateral, isto é,

$$\begin{cases} M = (L + R)/2 \\ S = (L - R)/2 \end{cases}$$

[000104] e então fazer uso da correlação restante entre esses canais, ou seja, ao configurar

$$S = \mathrm{Re}\{\alpha M\} + D,$$

[000105] onde α é o coeficiente de predição complexo a ser determinado e D é o sinal residual. É possível escolher α a fim de que a energia do sinal residual D = S Re{αM} seja minimizada. A minimização da energia pode ser efetuada com respeito à potência momentânea, uma energia de duração mais curta ou mais longa (média da potência), que no caso de um sinal discreto corresponde à otimização no sentido dos quadrados médios.

[000106] As partes real e imaginária α_R, α_I do coeficiente da predição podem ser quantizadas e/ou codificadas conjuntamente. De preferência, no entanto, as partes real e imaginária são quantizadas independente e uniformemente, tipicamente com um tamanho de etapa de 0,1 (número sem dimensão). A resolução da faixa de frequência usada para o coeficiente de predição complexo não é necessariamente a mesma que a resolução para faixas de fatores de escala (sfb; isto é, um grupo de linhas de MDCT que estão usando o mesmo tamanho de etapa de quantização de MDCT e a mesma faixa de quantização) de acordo com o padrão MPEG. Em particular, a definição da faixa de frequência para o coeficiente da predição pode ser uma que seja justificada psico-acusticamente, tal como a escala de Bark. Deve-se observar que a definição de faixa de frequência pode variar nos casos em que o comprimento da transformação varia.

[000107] Tal como já foi observado, o sistema codificador de acordo com a invenção pode ter uma liberdade se deve aplicar a codificação estereofônica da predição ou não, em que este último caso implica um recuo para a codificação de L/R ou M/S. Tal decisão pode ser tomada em uma base de quadro do tempo ou mais precisa, em uma base da faixa de frequência dentro de um quadro de tempo. Tal como observado acima, um resultado negativo da decisão pode ser comunicado à entidade de decodificação de várias maneiras, por exemplo, pelo valor de um bit de indicador dedicado em cada quadro, ou pela ausência (ou

valor zero) de um valor do coeficiente da predição. Uma decisão positiva pode ser comunicada de maneira análoga. Uma implementação particularmente vantajosa, que permite o recuo sem nenhuma extrapolação, faz uso de um quarto valor reservado do campo de dois bits ms_mask_present (vide MPEG-2 AAC, documento ISO/IEC 131818-7), que é transmitido por quadro de tempo e definido tal como segue:

Tabela 1: Definição de ms_mask_present em USAC	
Valor	Significado
00	codificação de L/R para todas as faixas de frequência
01	um bit de sinalização por faixa é usado para indicar L/R ou M/S
10	codificação de M/S para todas as faixas de frequência
11	Reservado

[000108] Com a redefinição do valor 11 para que signifique "codificação de predição complexa", o codificador pode ser operado em todos os modos de legado, em particular a codificação de M/S e L/R, sem nenhuma penalidade da taxa de bits, e ainda é capaz de sinalizar a codificação de predição complexa para os quadros onde é vantajoso.

[000109] A decisão substantiva pode ser baseada em um fundamento da taxa de dados e da qualidade de áudio. Como uma medida da qualidade, os dados obtidos ao usar um modelo psicoacústico incluído no codificador (tal como é frequentemente o caso de codificadores de áudio à base de MDCT disponíveis) podem ser usados. Em particular, algumas modalidades do codificador provêm uma seleção otimizada de distorção da taxa do coeficiente de predição. Por conseguinte, em tais modalidades, a parte imaginária - e possivelmente também a parte real - do coeficiente de predição é ajustada em zero se o aumento no ganho da predição não economizar bits suficientes para a codificação do sinal residual para justificar o gasto dos bits requeridos para a codificação do coeficiente de predição.

[000110] As modalidade do codificador podem codificar as informações que estão relacionadas a TNS no fluxo de bits. Tais informações

podem incluir valores dos parâmetros de TNS a serem aplicados pelos filtros de TNS (síntese) no lado do decodificador. Se conjuntos idênticos de parâmetros de TNS tiverem que ser usados para ambos os canais, é econômica a inclusão de um bit de sinalização que indique que essa identidade dos conjuntos de parâmetros ao invés de transmitir os dois conjuntos de informações dos parâmetros independentemente. Informações também podem ser incluídas quanto ao fato se TNS deve ser aplicada antes ou depois do estágio de upmix, tal como apropriado com base, por exemplo, em uma avaliação psicoacústica das duas opções disponíveis.

[000111] Tal como ainda uma outra característica opcional, que é potencialmente benéfica de um ponto de vista de complexidade e da taxa de bits, o codificador pode ser adaptado para usar uma largura de faixa individualmente limitada para a codificação do sinal residual. As faixas de frequência acima desse limite não serão transmitidas ao decodificador, mas serão ajustadas em zero. Em determinados casos, as faixas de frequência mais elevadas têm um teor de energia tão pequeno que são já são quantizadas em zero. A prática normal (de acordo com o parâmetro max_sfb no padrão MPEG) envolve o uso da mesma limitação de largura de faixa para a downmix e sinais residuais. Agora, os autores da presente invenção verificaram empiricamente que o sinal residual, até uma extensão maior do que o sinal de downmix, tem o seu teor de energia localizado em faixas de frequência mais baixa. Portanto, ao colocar um limite da largura de faixa superior dedicada para o sinal residual, uma redução na taxa de bits é possível sem nenhuma perda significativa da qualidade. Por exemplo, isto pode ser obtido pela transmissão de dois parâmetros max_sfb independentes, um para o sinal de downmix e um para o sinal residual.

[000112] É indicado que, embora as questões da determinação ideal do coeficiente da predição, da quantização e da sua codificação, do

recuo para o modo de M/S ou L/R, da filtragem TNS e da limitação de largura de faixa superior, etc., sejam discutidas com referência ao sistema decodificador mostrado na figura 5, os mesmos fatos são igualmente aplicáveis às modalidades que serão apresentadas no que segue com referência às figuras subsequentes.

[000113] A Figura 6 mostra um outro sistema codificador de acordo com a invenção adaptado para executar a codificação estereofônica de predição complexa. O sistema recebe como entrada uma representação do domínio do tempo de um sinal estereofônico segmentado em quadros de tempo sucessivos possivelmente sobrepostos e compreende canais esquerdo e direito. Um estágio de soma e diferença 601 converte o sinal nos canais do meio e laterais. O canal do meio é provido a um módulo de MDCT 602 e a um módulo de MDST 603, ao passo que o canal lateral é provido apenas a um módulo de MDCT 604. Um estimador de coeficiente da predição 605 estima para cada quadro de tempo - e possivelmente para faixas de frequência individuais dentro de um quadro - um valor do coeficiente de predição complexo α tal como explicado acima. O valor do coeficiente α é provido como peso aos somadores ponderados 606, 607, que formam um sinal residual D como uma combinação linear das representações de MDCT e MDST do sinal do meio e da representação de MDCT do sinal lateral. De preferência, o coeficiente de predição complexo é provido aos somadores ponderados 606, 607 representados pelo mesmo esquema de quantização que será usado quando ele é codificado no fluxo de bits; isto provê obviamente uma reconstrução mais fiel, uma vez que ambos o codificador e o decodificador aplicam o mesmo valor do coeficiente de predição. O sinal residual, o sinal do meio (que pode ser mais apropriadamente chamado de sinal de downmix quando aparece em combinação com um sinal residual) e o coeficiente de predição são providos a um estágio combinado de quantização e multiplexador

combinado 608, que codifique estes e as possíveis informações adicionais como um fluxo de bits de saída.

[000114] A figura 7 mostra uma variação do sistema codificador da figura 6. Tal como fica evidente a partir da similaridade dos símbolos na figura, ele tem uma estrutura similar, mas também a funcionalidade adicional de operar em um modo de recuo de codificação de L/R direto. O sistema codificador é acionado entre o modo de codificação de predição complexa e o modo de recuo por meio de um comutador 710 provido imediatamente a montante do estágio de quantização e multiplexador combinado 709. Em sua posição superior, tal como mostrado na figura, o comutador 710 irá fazer com que o codificador opere no modo de recuo. Dos pontos imediatamente a jusante dos módulos de MDCT 702, 704, o sinal do meio/laterais é provido ao estágio de soma e diferença 705, que depois de ter convertido o mesmo na forma da esquerda/direita, passa o mesmo para o comutador 710, que conecta o mesmo ao estágio de quantização e multiplexador combinado 709.

[000115] A figura 8 mostra um sistema codificador de acordo com a presente invenção. Em contraste com os sistemas codificadores das figuras 6 e 7, esta modalidade deriva os dados de MDST requeridos para a codificação de predição complexa diretamente a partir dos dados de MDCT, isto é, por uma transformação de real em imaginário no domínio da frequência. A transformação de real em imaginário aplica qualquer uma das abordagens discutidas em relação aos sistemas decodificadores das figuras 2 e 4. É importante combinar o método de computação do decodificador com aquele do codificador, de modo que uma decodificação fiel possa ser obtida; de preferência, métodos de transformação de real em imaginário idênticos são usados no lado do codificador e no lado do decodificador. No que se refere às modalidades do decodificador, a parte A envolvida uma linha tracejada e que compreende a transformação de real em imaginário 804 pode ser

substituída por variantes aproximadas ou ao usar menos quadros de tempo de entrada como entrada. Do mesmo modo, a codificação pode ser simplificada ao usar qualquer uma das outras abordagens de aproximação descritas acima.

[000116] Em um nível mais elevado, o sistema codificador da figura 8 tem uma estrutura que difere daquela que deve seguir provavelmente por uma ação direta de substituição do módulo de MDST na figura 7 por um módulo (conectado apropriadamente) real a imaginário. A presente arquitetura é limpa e propicia a funcionalidade de comutar entre a codificação da predição e a codificação de L/R direta de uma maneira robusta e computacionalmente econômica. O sinal estereofônico de entrada é alimentado nos módulos de transformação de MDCT 801, que emitem uma representação do domínio da frequência de cada canal. Isso é alimentado a um comutador final 808 para acionar o sistema codificador entre os seus modos de predição e codificação direto e a um estágio de soma e diferença 802. Na codificação de L/R direta ou na codificação de M/S conjunta - que é realizada no quadro de tempo para o qual o coeficiente da predição α é ajustado em zero - esta modalidade sujeita o sinal de entrada à transformação MDCT, à quantização e à multiplexação somente, em que as últimas duas etapas são executadas por um estágio de quantização e multiplexador combinado 807 arranjado na extremidade da saída do sistema, onde um fluxo de bits é provido. Na codificação de predição, cada um dos canais é submetido a um processamento adicional entre o estágio de soma e diferença 802 e o comutador 808. A partir da representação de MDCT do sinal do meio, a transformação de real em imaginário 804 deriva os dados da MDST e encaminha os mesmos a um estimador de coeficiente da predição 803 e a um somador ponderado 806. Tal como nos sistemas codificadores mostrados nas figuras 6 e 7, um somador ponderado adicional 805 é usado para combinar o sinal lateral com as re-

presentações de MDCT e MDST ponderadas do sinal do meio para formar um sinal de canal residual, o qual é codificado em conjunto com o sinal de canal do meio (isto é, de downmix) e o coeficiente de predição pelo módulo de quantização e multiplexador combinado 807.

[000117] Voltando agora para a figura 9, será ilustrado que cada uma das modalidades do sistema codificador pode ser combinada com um ou mais filtros de TNS (análise). De acordo com as discussões precedentes, é frequentemente vantajoso aplicar a filtragem TNS ao sinal em sua forma de downmix. Desse modo, tal como mostrado na figura 9, a adaptação do sistema codificador da figura 7 para incluir TNS é efetuada ao adicionar os filtros de TNS 911 imediatamente a montante do módulo de quantização e multiplexador combinado 909.

[000118] Em vez do filtro de TNS direito/residual 911, dois filtros de TNS separados (não mostrados) podem ser providos imediatamente a montante da parte do comutador 910 adaptada para manipular o canal direito ou residual. Desse modo, cada um dos dois filtros de TNS será provido com os respectivos dados de sinal de canal todas as vezes, permitindo a filtragem de TNS baseada em mais quadros de tempo do que somente o atual. Tal como já foi observado, os filtros de TNS são apenas um exemplo de dispositivos modificadores do domínio da frequência, principalmente dispositivos que baseiam o seu processamento em mais quadros do que o atual, que podem se beneficiar de uma colocação tanto quanto ou mais do que no filtro de TNS.

[000119] Como uma outra alternativa possível à modalidade mostrada na figura 9, os filtros de TNS para a ativação seletiva podem ser arranjados em mais de um ponto para cada canal. Isto é similar à estrutura do sistema decodificador mostrado na figura 4, onde conjuntos diferentes de filtros de TNS podem ser conectados por meio de comutadores. Isto permite a seleção do estágio disponível mais apropriado para a filtragem TNS para cada quadro. Em particular, pode ser vanta-

joso comutar entre posições diferentes de TNS em relação à comutação entre a codificação estereofônica da predição complexa e outros modos de codificação.

[000120] A figura 11 mostra uma variação baseada no sistema codificador da figura 8, em que uma segunda representação do domínio da frequência do sinal de downmix é derivada por meio de um de uma transformação de real em imaginário 1105. Similarmente ao sistema decodificador mostrado na figura 4, este sistema codificador também inclui módulos modificadores de domínio da frequência seletivamente ativáveis, um deles 1102 provido a montante do estágio de downmix e um outro 1109 provido a jusante do mesmo. Os módulos de domínio da frequência 1102, 1109, que nesta figura são exemplificados por filtros de TNS, podem ser conectados em cada uma das passagens de sinal ao usar os quatro comutadores 1103a, 1103b, 1109a e 1109b.

III. Modalidade sem aparelho

[000121] As modalidades do terceiro e de um quarto aspectos da invenção são mostradas nas figuras 15 e 16. A figura 15 mostra um método para a decodificação de um fluxo de bits em um sinal estereofônico, o qual compreende as seguintes etapas:

1. Um fluxo de bits é inserido.

2. O fluxo de bits é desquantizado, por meio do que é obtida uma primeira representação do domínio da frequência de canais de downmix e residuais de um sinal estereofônico.

3. Uma segunda representação do domínio da frequência do canal de downmix é computada.

4. Um sinal do canal laterais é computado com base nas três representações do domínio da frequência dos canais.

5. Um sinal estereofônico, de preferência de forma esquerda/direita, é computado com base nos canais laterais e de downmix.

6. O sinal estereofônico obtido desse modo é emitido.

[000122] As etapas 3 a 5 podem ser consideradas como um processo de upmix. Cada uma das etapas 1 a 6 é análoga à funcionalidade correspondente em qualquer um dos sistemas decodificadores apresentados nas partes precedentes deste texto, e mais detalhes que estão relacionados à sua implementação podem ser recuperados nas mesmas partes.

[000123] A figura 16 mostra um método para a codificação de um sinal estereofônico como um sinal do fluxo de bits, o qual compreende as seguintes etapas:

1. Um sinal estereofônico é inserido.

2. O sinal estereofônico é transformado em uma primeira representação do domínio da frequência.

3. Um coeficiente de predição complexo é determinado.

4. É feito o upmix da representação do domínio da frequência.

5. Os canais de downmix e residuais são codificados como um fluxo de bits em conjunto com o coeficiente de predição complexo.

6. O fluxo de bits é emitido.

[000124] Cada uma das etapas 1 a 5 é análoga à funcionalidade correspondente em qualquer um dos sistemas codificadores apresentados nas partes precedentes deste texto, e mais detalhes que estão relacionados à sua implementação podem ser recuperados nas mesmas partes.

[000125] Ambos os métodos podem ser expressos como instruções que podem ser lidas por computador na forma de programas de software e podem ser executados por um computador. O âmbito de proteção da presente invenção estende-se a tais produtos de software e programa de computador para distribuir tal software.

IV. Avaliação empírica

[000126] Várias das modalidades aqui apresentadas foram avaliadas empiricamente. As partes mais importantes do material experimental obtido neste processo serão resumidas nesta subseção.

[000127] A modalidade usada para os experimentos tinha as seguintes características:

(i) Cada espectro de MDST (para um quadro de tempo) foi computado pela filtragem da resposta de impulso finita bidimensional dos espectros de MDCT presente, precedente e seguinte.

(ii) Um modelo psicoacústico do codificador estereofônico USAC foi usado.

(iii) As partes real e imaginária do coeficiente de predição complexo α foram transmitidas em lugar dos parâmetros de PS ICC, CLD e IPD. As partes real e imaginária foram manipuladas independentemente, foram limitadas à faixa [-3,0, 3,0] e quantizadas ao usar um tamanho de etapa de 0,1. Elas foram então codificadas diferencialmente no tempo e finalmente codificada segundo Huffman ao usar o livro de códigos de fator de escala da USAC. Os coeficientes de predição foram atualizados a cada segunda faixa do fator de escala, o que resultou em uma definição da frequência similar àquela de Surround MPEG (vide, por exemplo, ISO/IEC 23003-1). Esse esquema de quantização e de codificação resultou em uma taxa de bits média de cerca de 2 kb/s para esta informação lateral estereofônico em uma configuração típica com uma taxa de bits alvo de 96 kb/s.

(iv) O formato do fluxo de bits foi modificado sem quebrar os fluxos de bits USAC atuais, uma vez que o elemento de fluxo de bits ms_mask_present tem atualmente somente três valores possíveis. O uso do quarto valor para indicar a predição complexa permite um modo de recuo da codificação do meio/laterais básica sem nenhum bit desperdiçado (para mais detalhes sobre este assunto, vide a subseção precedente desta apresentação).

[000128] Os testes de audição foram realizados de acordo com a metodologia MUSHRA, envolvendo em particular a reprodução em fones de ouvido e o uso de 8 itens de teste com uma taxa da amostra-

gem de 48 kHz. Três, cinco ou seis indivíduos de teste participaram de cada teste.

[000129] O impacto de diferentes aproximações de MDST foi avaliado para ilustrar a troca prática de complexidade para qualidade que existe entre essas opções. Os resultados são encontrados nas figuras 12 e 13, em que a primeira mostra as contagens absolutas obtidas e a última mostra as contagens diferenciais em relação a 96s USAC cp1f, isto é, codificação estereofônica unificada de domínio de MDCT pela predição complexa ao usar um quadro de MDCT atual para computar uma aproximação da MDST. Pode-se ver que o ganho da qualidade obtido pela codificação estereofônica unificada à base de MDCT aumenta quando mais abordagens computacionalmente complexas para computar o espectro de MDST são aplicadas. Considerando a média de todos os testes, o sistema baseado em um só quadro 96s USAC cp1f provê um aumento significativo na eficiência da codificação em relação à codificação estereofônica convencional. Por sua vez, resultados até mesmo significativamente melhores são obtidos para 96s USAC cp3f, ou seja, a codificação estereofônica unificada de domínio de MDCT pela predição complexa ao usar quadros atuais, precedentes e seguintes para computar a MDST.

V. Modalidades

[000130] Além disso, a invenção pode ser incorporada como um sistema decodificador para a decodificação de um sinal do fluxo de bits em um sinal estereofônico pela codificação estereofônica de predição, em que o sistema decodificador compreende:

[000131] um estágio de desquantização (202; 401) para prover as primeiras representações do domínio da frequência de um sinal de downmix (M) e de um sinal residual (D) com base no sinal do fluxo de bits, em que cada uma das primeiras representações do domínio da frequência compreende o conteúdo espectral do sinal correspondente

que representa o conteúdo espectral do sinal correspondente expresso em um primeiro subespaço de um espaço multidimensional, em que os primeiros componentes espectrais são coeficientes de transformação arranjados em um ou mais quadros de tempo de coeficientes de transformação, em que cada bloco é gerado pela aplicação de uma transformação a um segmento do tempo de um sinal de domínio do tempo; e

[000132] um estágio de upmix (206, 207, 210, 211; 406, 407, 408, 409) arranjado a jusante do estágio de desquantização, adaptado para gerar o sinal estereofônico com base no sinal de downmix e no sinal residual, e compreende:

[000133] um módulo (206; 408) para computar uma segunda representação do domínio da frequência do sinal de downmix com base na sua primeira representação do domínio da frequência, em que a segunda representação do domínio da frequência compreende segundos componentes espectrais que representam o conteúdo espectral do sinal expresso em um segundo subespaço do espaço multidimensional que inclui uma parte do espaço multidimensional não incluído no primeiro subespaço, em que o dito módulo é adaptado para:

[000134] derivar um ou mais primeiros componentes intermediários de pelo menos alguns dos primeiros componentes espectrais;

[000135] formar uma combinação dos ditos um ou mais primeiros componentes espectrais de acordo com pelo menos uma parte de uma ou mais respostas de impulsos para obter um ou mais segundos componentes intermediários; e

[000136] derivar um ou mais ditos segundos componentes espectrais de um ou mais ditos segundos componentes intermediários;

[000137] um somador ponderado (210, 211; 406, 407) para computar um sinal lateral (8) com base na primeira e segunda representações do domínio da frequência do sinal de downmix, na primeira represen-

tação do domínio de frequência do sinal residual e um coeficiente de predição complexo (α) codificado no sinal do fluxo de bits; e

[000138] um estágio de soma e diferença (207; 409) para computar o sinal estereofônico com base na primeira representação do domínio da frequência do sinal de downmix e do sinal lateral.

[000139] Além disso, a invenção pode ser incorporada como um sistema decodificador para decodificar um sinal do fluxo de bits em um sinal estereofônico pela codificação estereofônica de predição complexa, em que o sistema decodificador compreende:

[000140] um estágio de desquantização (301) para prover as primeiras representações do domínio da frequência de um sinal de downmix (M) e de um sinal residual (D) com base no sinal do fluxo de bits, em que cada uma das primeiras representações do domínio da frequência compreende os primeiros componentes espectrais que representam o conteúdo espectral do sinal correspondente expresso em um primeiro subespaço de um espaço multidimensional; e

[000141] um estágio de upmix (306, 307, 308, 309, 312) arranjado a jusante do estágio de desquantização, adaptado para gerar o sinal estereofônico com base no sinal de downmix e no sinal residual, e compreende:

[000142] um módulo (306, 307) para computar uma segunda representação do domínio da frequência do sinal de downmix com base na sua primeira representação do domínio da frequência, em que a segunda representação do domínio da frequência compreende segundos componentes espectrais que representam o conteúdo espectral do sinal expresso em um segundo subespaço do espaço multidimensional que inclui uma parte do espaço multidimensional não incluída no primeiro subespaço, em que o módulo compreende:

[000143] um estágio de transformação inversa (306) para computar uma representação do domínio do tempo do sinal de downmix com

base na primeira representação do domínio da frequência do sinal de downmix no primeiro subespaço do espaço multidimensional; e

[000144] um estágio de transformação (307) para computar a segunda representação do domínio da frequência do sinal de downmix com base na representação do domínio do tempo do sinal;

[000145] um somador ponderado (308, 309) para computar um sinal lateral (S) com base na primeira e segunda representações do domínio da frequência, na primeira e segunda representações do domínio da frequência do sinal de downmix, na primeira representação do domínio da frequência do sinal residual e em um coeficiente de predição complexo (α) codificado no sinal do fluxo de bits; e

[000146] um estágio de soma e diferença (312) para computar o sinal estereofônico com base na primeira representação do domínio da frequência do sinal de downmix e do sinal lateral.

[000147] Além disso, a invenção pode ser incorporada como um sistema decodificador com as características apresentadas na concretização do sistema decodificador, em que o módulo para computar uma segunda representação do domínio da frequência do sinal de downmix compreende:

[000148] um estágio de transformação inversa (306) para computar uma representação do domínio do tempo do sinal de downmix e/ou do sinal lateral com base na respectiva primeira representação do domínio da frequência do sinal no primeiro subespaço do espaço multidimensional; e

[000149] um estágio de transformação (307) para computar a segunda representação do domínio da frequência do respectivo sinal com base na representação do domínio do tempo do sinal,

[000150] em que, de preferência, o estágio de transformação inversa (306) executa uma transformada discreta de cosseno modificada inversa e o estágio da transformação executa uma transformada discre-

ta de seno modificada.

[000151] No sistema decodificador acima, o sinal estereofônico pode ser representado no domínio de tempo e o sistema decodificador pode ainda compreender:

[000152] um conjunto de comutação (302) arranjado entre o dito estágio de desquantização e o dito estágio de upmix, operáveis para funcionar como qualquer um de:

(a) um estágio passante, para o uso na codificação estereofônica comum; ou

(b) um estágio de soma e diferença, para o uso na codificação estereofônica direta;

[000153] um estágio de transformação inversa adicional (311) arranjado no estágio de upmix, para computar uma representação do domínio do tempo do sinal lateral;

[000154] um arranjo de seletores (305, 310) arranjado a montante dos estágios de transformação inversa (306, 301), adaptados para conectar seletivamente estes a qualquer um entre:

(a) um estágio de soma e diferença adicional (304) que por sua vez é conectado a um ponto a jusante do conjunto de comutação (302) e a montante do estágio de upmix; ou

(b) um sinal de downmix obtido a partir do conjunto de comutação (302) e um sinal lateral obtido a partir do somador ponderado (308, 309).

VI. Observações de fechamento

[000155] Outras modalidades da presente invenção tornar-se-ão aparentes a um elemento versado na técnica após a leitura da descrição acima. Muito embora a presente descrição e os desenhos apresentem modalidades e exemplos, a invenção não fica restringida a esses exemplos específicos. Numerosas modificações e variações podem ser feitas sem que se desvie do âmbito da presente invenção, o

qual é definido pelas concretizações.

[000156] Deve-se observar que os métodos e o aparelho apresentados neste pedido de patente podem ser aplicados, após modificações apropriadas dentro das habilidades de um elemento versado na técnica, incluindo a experimentação rotineira, à codificação dos sinais que têm mais de dois canais. É particularmente enfatizado que todos os sinais, parâmetros e matrizes mencionados em conexão com as modalidades descritas podem ser variáveis na frequência ou invariáveis na frequência e/ou variáveis no tempo ou invariáveis no tempo. As etapas de computação descritas podem ser executadas à guisa de frequência ou para todas as faixas de frequência de cada vez, e todas as entidades podem ser incorporadas para que tenham uma ação seletiva quanto à frequência. Para as finalidades do pedido de patente, todos os esquemas de quantização podem ser adaptados de acordo com modelos psico-acústicos. Além disso, deve-se observar que as várias conversões de soma e diferença, isto é, a conversão da forma de downmix/residual na forma pseudo-L/R, bem como a conversão de L/R em M/S e a conversão de M/S em L/R, são todas da forma

$$g\begin{bmatrix}1 & 1\\ 1 & -1\end{bmatrix},$$

[000157] onde, meramente, o fator de ganho g pode variar. Desse modo, ao ajustar os fatores do ganho individualmente, é possível compensar um certo ganho de codificação por uma escolha apropriada do ganho de decodificação. Além disso, tal como é concluído por um elemento versado na técnica, um número par de transformações de soma e diferença arranjadas em série tem o efeito de um estágio passante, possivelmente com ganho da não unitário.

[000158] Os sistemas e os métodos apresentados acima podem ser implementados como software, firmware, hardware ou uma combinação destes. Determinados componentes ou todos os componentes podem ser implementados como software executado por um processador

ou um microprocessador de sinais digitais, ou ser executados como hardware ou como um circuito integrado específico de aplicação. Tal software pode ser distribuído em meios que podem ser lidos por computador, que podem compreender meios de armazenamento do computador e meios de comunicação. Como é bem sabido por um elemento versado na técnica, os meios de armazenagem do computador incluem 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 do armazenamento do computador incluem, mas sem ficar a eles limitados, RAM, ROM, EEPROM, memória rápida ou outra tecnologia de memória, CD-ROM, discos versáteis digitais (DVD) ou outro armazenamento em disco ótico, cassetes magnéticos, fitas magnéticas, armazenamento em disco magnético ou outros dispositivos de armazenamento magnético, ou qualquer outro meio que possa ser usado para armazenar as informações desejadas e que possa ser acessado por um computador. Além disso, é sabido pelo elemento versado na técnica que os 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 incluem quaisquer meios de entrega de informações.

REIVINDICAÇÕES

1. Sistema decodificador para fornecer um sinal de som estereofônico por predição complexo de codificação, o sistema decodificador **caracterizado pelo fato de que** compreende:

um estágio de upmix adaptado para gerar o sinal estereofônico com base nas primeiras representações no domínio da frequência de um sinal de downmix (M) e de um sinal residual (D), cada uma das primeiras representações no domínio da frequência compreendendo os primeiros componentes espectrais que representam o conteúdo espectral do sinal correspondente expresso em um primeiro subespaço de um espaço multidimensional, o estágio de upmix compreendendo:

um módulo para calcular uma segunda representação no domínio da frequência do sinal de downmix com base na primeira representação no domínio da frequência, a segunda representação no domínio da frequência compreendendo segundos componentes espectrais que representam conteúdo espectral do sinal expresso em um segundo subespaço do espaço multidimensional que inclui uma porção do espaço multidimensional não incluído no primeiro subespaço, em que o módulo é adaptado para determinar os segundos componentes espectrais do sinal de downmix aplicando um filtro de resposta de impulso finita (FIR) aos primeiros componentes espectrais do sinal de downmix;

um somador ponderado para calcular uma primeira representação no domínio da frequência de um sinal lateral (S), a primeira representação no domínio da frequência do sinal lateral (S) compreendendo primeiros componentes espectrais representando o conteúdo espectral do sinal lateral expresso no primeiro subespaço do espaço multidimensional, com base nas primeira e segunda representações no domínio da frequência do sinal de downmix, a primeira representa-

ção no domínio da frequência do sinal residual e um coeficiente de predição complexo (α) codificado em um sinal de fluxo de bits recebido pelo sistema decodificador; em que cada componente espectral representa uma faixa de frequências, e em que cada um dos primeiros componentes espectrais do sinal lateral é determinado a partir de componentes espectrais do sinal de downmix e do sinal residual representando a mesma faixa de frequências que o primeiro componente espectral do sinal lateral; e

um estágio de soma e diferença para calcular o sinal estereofônico com base na primeira representação no domínio da frequência do sinal de downmix e do sinal lateral,

em que o estágio de upmix é ainda operável em um modo de passagem, no qual os sinais de downmix e residual são fornecidos diretamente ao estágio de soma e diferença.

2. Sistema decodificador, de acordo com a reivindicação 1, **caracterizado pelo fato de que** uma resposta de impulso do filtro FIR é determinada dependendo de uma função de janela aplicada para determinar a primeira representação no domínio da frequência do sinal de downmix.

3. Sistema decodificador, de acordo com a reivindicação 1, **caracterizado pelo fato de que** o sinal de downmix e o sinal residual são segmentados em quadros de tempo,

o estágio de upmix sendo ainda adaptado para receber, para cada quadro de tempo, um campo de dados de dois bits associado a esse quadro e para operar, em resposta ao valor do campo de dados, em seu modo ativo ou em seu modo de passagem.

4. Sistema decodificador, de acordo com a reivindicação 1, **caracterizado pelo fato de que** os sinais de downmix e residual são segmentados em quadros de tempo,

o estágio upmix sendo adaptado ainda mais para receber

em um fluxo de bits MPEG, para cada quadro de tempo, um campo *ms_mask_present* associado a esse quadro e para operar, em resposta ao valor do campo *ms_mask_present*, em seu modo ativo ou em seu modo de passagem.

5. Sistema decodificador, de acordo com a reivindicação 1, **caracterizado pelo fato de que**:

os primeiros componentes espectrais têm valores reais expressos no primeiro subespaço; e

os segundos componentes espectrais têm valores imaginários expressos no segundo subespaço;

opcionalmente, os primeiros componentes espectrais podem ser obtidos por um dos seguintes:

uma transformada discreta de cosseno (DCT), ou

uma transformada discreta de cosseno modificada (MDCT),

e, opcionalmente, os segundos componentes espectrais podem ser obtidos por um dos seguintes:

uma transformada discreta de seno (DST), ou

uma transformada discreta de seno modificada (MDST).

6. Sistema decodificador, de acordo com a reivindicação 5, **caracterizado pelo fato de que**:

o sinal de downmix é particionado em quadros de tempo sucessivos, cada um associado a um valor do coeficiente de predição complexo; e

o módulo para calcular uma segunda representação no domínio da frequência do sinal de downmix é adaptado para se desativar, responsivo ao valor absoluto da parte imaginária do coeficiente de predição complexo, sendo menor que uma tolerância predeterminada para um quadro de tempo, para que ele não gere saída para esse quadro de tempo.

7. Sistema decodificador, de acordo com a reivindicação 6,

caracterizado pelo fato de que os quadros de tempo do sinal de downmix são ainda particionados em faixas de frequência, cada uma acompanhada por um valor do coeficiente de predição complexo; e

o módulo para calcular uma segunda representação no domínio da frequência do sinal de downmix é adaptado para se desativar, responsivo ao valor absoluto da parte imaginária do coeficiente de predição complexo sendo menor que uma tolerância predeterminada para uma faixa de frequência de um quadro de tempo, de modo que não gere saída para essa faixa de frequência.

8. Método de decodificação para fornecer um sinal estereofônico por codificação estereofônica de predição complexa, o método **caracterizado pelo fato de que** compreende as etapas de:

receber primeiras representações no domínio da frequência de um sinal de downmix (M) e um sinal residual (D), cada uma das primeiras representações no domínio da frequência compreendendo primeiros componentes espectrais representando o conteúdo espectral do sinal correspondente expresso em um primeiro subespaço de um espaço multidimensional;

receber um sinal de controle; e

responsivo ao valor do sinal de controle:

(a) fazer upmix do sinal de downmix e dos sinais residuais, usando um estágio de upmix, para obter o sinal estereofônico, executando as sub etapas de:

computar uma segunda representação no domínio da frequência do sinal de downmix com base na primeira representação no domínio da frequência do mesmo, a segunda representação no domínio da frequência compreendendo segundos componentes espectrais que representam o conteúdo espectral do sinal expresso em um segundo subespaço do espaço multidimensional que inclui uma porção do espaço multidimensional não incluído no primeiro subespaço, em

que o cálculo de uma segunda representação no domínio da frequência do sinal de downmix inclui a determinar os segundos componentes espectrais do sinal de downmix aplicando um filtro de resposta de impulso finita (FIR) aos primeiros componentes espectrais do sinal de downmix;

computar uma primeira representação no domínio da frequência de um sinal lateral (S), a primeira representação no domínio da frequência do sinal lateral (S) compreendendo primeiros componentes espectrais que representam o conteúdo espectral do sinal lateral expresso no primeiro subespaço do espaço multidimensional, com base nas primeira e segunda representações no domínio da frequência do sinal de downmix, a primeira representação no domínio da frequência do sinal residual e um coeficiente de predição complexo (α) codificado em um sinal de fluxo de bits; em que cada componente espectral representa uma faixa de frequências, e em que cada um dos primeiros componentes espectrais do sinal lateral é determinado a partir de componentes espectrais do sinal de downmix e o sinal residual representando a mesma faixa de frequências que o primeiro componente espectral do sinal lateral; e

computar o sinal estereofônico aplicando uma transformada de soma e diferença à primeira representação no domínio da frequência do sinal de downmix e do sinal lateral,

ou

(b) suspender o a realização do upmix.

9. Método de decodificação, de acordo com a reivindicação 8, **caracterizado pelo fato de que**:

os primeiros componentes espectrais têm valores reais expressos no primeiro subespaço;

os segundos componentes espectrais têm valores imaginários expressos no segundo subespaço;

opcionalmente, os primeiros componentes espectrais podem ser obtidos por um dos seguintes:

uma transformada discreta de cosseno (DCT), ou

uma transformada discreta de cosseno modificada (MDCT),

e, opcionalmente, os segundos componentes espectrais podem ser obtidos por um dos seguintes:

uma transformada discreta de seno (DST), ou

uma transformada discreta de seno modificada (MDST).

10. Método de decodificação, de acordo com a reivindicação 9, **caracterizado pelo fato de que**:

o sinal de downmix é particionado em quadros de tempo sucessivos, cada um associado a um valor do coeficiente de predição complexo; e

o cálculo de uma segunda representação no domínio da frequência do sinal de downmix é suspenso em resposta ao valor absoluto da parte imaginária do coeficiente de predição complexo, sendo menor que uma tolerância predeterminada para um quadro de tempo, de modo que nenhuma saída seja gerada para esse período de tempo.

11. Método de decodificação, de acordo com a reivindicação 10, **caracterizado pelo fato de que**:

os intervalos de tempo do sinal de downmix são ainda particionados em faixas de frequência, cada uma acompanhada de um valor do coeficiente de predição complexo; e

o cálculo de uma segunda representação no domínio da frequência do sinal de downmix é suspenso em resposta ao valor absoluto da parte imaginária do coeficiente de predição complexo, sendo menor que uma tolerância predeterminada para uma faixa de frequência de um quadro de tempo, de modo que nenhuma saída seja gerada para essa banda de frequência.

12. Método de decodificação, de acordo com a reivindica-

ção 8, **caracterizado pelo fato de que** o sinal estereofônico é representado no domínio do tempo e o método compreende ainda:

omitir, em resposta ao sinal de fluxo de bits ser codificado por codificação estereofônica direta ou codificação estereofônica conjunta, a etapa de realização do upmix; e

transformar inversamente o sinal de fluxo de bits para obter o sinal estereofônico.

13. Método de decodificação, de acordo com a reivindicação 12, **caracterizado pelo fato de que** compreende ainda a etapa de:

omitir, em resposta ao sinal de fluxo de bits ser codificado por codificação estereofônica direta ou codificação estereofônica conjunta, as etapas de transformação da representação no domínio do tempo do sinal de downmix e de computação de um sinal lateral; e

transformar inversamente as representações no domínio da frequência de cada um dos canais codificados pelo sinal de fluxo de bits para obter o sinal estereofônico.

14. Aparelho para emitir um sinal de áudio estereofônico tendo um canal esquerdo e um canal direito, o aparelho **caracterizado pelo fato de que** compreende:

um demultiplexador configurado para receber um fluxo de bits de áudio e extrair dele um coeficiente de predição, em que o fluxo de bits de áudio é segmentado em quadros e um valor do coeficiente de predição pode mudar para cada um dos quadros;

um decodificador configurado para gerar um sinal de downmix e um sinal residual do fluxo de bits de áudio; e

um upmixer configurado para operar em um modo de predição ou em um modo sem predição com base em um parâmetro codificado no fluxo de bits de áudio e para emitir o canal esquerdo e o canal direito como sinal de áudio estereofônico,

em que, quando o upmixer opera no modo de predição, o sinal residual representa uma diferença entre um sinal lateral e uma versão prevista do sinal lateral, e o upmixer gera o canal esquerdo e o canal direito a partir de uma combinação do sinal downmix, o sinal residual e o coeficiente de predição, e

em que, quando o upmixer opera no modo de não-predição, o sinal residual representa o sinal lateral, o upmixer gera o canal esquerdo com base em uma soma do sinal downmix e o sinal residual, e o upmixer gera o canal direito com base em um diferença do sinal downmix e o sinal residual.

15. Aparelho, de acordo com a reivindicação 14, **caracterizado pelo fato de que** o parâmetro é o coeficiente de predição.

16. Aparelho, de acordo com a reivindicação 15, **caracterizado pelo fato de que** o upmixer opera no modo de não-predição por um quadro de tempo quando o valor do coeficiente de predição é igual a zero ou é menor que uma tolerância predeterminada e o upmixer opera no modo de predição durante o quadro de tempo para todos os outros valores do coeficiente de predição.

17. Aparelho, de acordo com a reivindicação 14, **caracterizado pelo fato de que** o coeficiente de predição reduz ou minimiza uma energia do sinal residual.

18. Aparelho, de acordo com a reivindicação 14, **caracterizado pelo fato de que** compreende ainda um modelador de ruído configurado para moldar um ruído associado ao sinal de downmix, em que o modelador de ruído está disposto a montante do upmixer.

19. Aparelho, de acordo com a reivindicação 18, **caracterizado pelo fato de que** o modelador de ruído é um modelador de ruído temporal configurado para modelar o ruído ao longo do tempo.

20. Aparelho, de acordo com a reivindicação 14, **caracterizado pelo fato de que**, quando o upmixer opera no modo de predição,

o upmixer gera o canal esquerdo e o canal direito usando um filtro com três derivações.

21. Aparelho, de acordo com a reivindicação 14, **caracterizado pelo fato de que** o sinal de downmix compreende um sinal intermediário formado por uma combinação linear de um canal esquerdo original e um canal direito original.

22. Aparelho, de acordo com a reivindicação 14, **caracterizado pelo fato de que** o coeficiente de predição é um coeficiente com valor real.

23. Aparelho, de acordo com a reivindicação 14, **caracterizado pelo fato de que** o coeficiente de predição é um coeficiente de valor complexo.

24. Aparelho, de acordo com a reivindicação 14, **caracterizado pelo fato de que** o upmixer combina o sinal lateral com o sinal de downmix adicionando uma versão do sinal de downmix com uma versão do sinal lateral para gerar o canal esquerdo e subtraindo a versão do sinal lateral da versão do sinal de downmix para gerar o canal direito.

25. Aparelho, de acordo com a reivindicação 14, **caracterizado pelo fato de que** o coeficiente de predição é codificado no fluxo de bits de áudio e o demultiplexador é ainda configurado para decodificar o coeficiente de predição.

26. Aparelho, de acordo com a reivindicação 14, **caracterizado pelo fato de que** o upmixer é ainda configurado para adicionar o sinal residual ao sinal lateral ao operar no modo de predição.

27. Método para emitir um sinal de áudio estereofônico tendo um canal esquerdo e um canal direito, o método **caracterizado pelo fato de que** compreende as etapas de:

receber um fluxo de bits de áudio e extrair dele um coeficiente de predição, em que o fluxo de bits de áudio é segmentado em

quadros e um valor do coeficiente de predição pode mudar para cada um dos quadros;

gerar, em um decodificador, um sinal de downmix e um sinal residual a partir do fluxo de bits de áudio;

realizar upmix, no modo de predição ou no modo de não-predição, com base em um parâmetro codificado no fluxo de bits de áudio; e

emitir o canal esquerdo e o canal direito como sinal de áudio estereofônico,

em que, quando o upmix opera no modo de predição, o sinal residual representa uma diferença entre um sinal lateral e uma versão prevista do sinal lateral, e o upmix gera o canal esquerdo e o canal direito a partir de uma combinação do sinal de dowmix, o sinal residual e o coeficiente de predição, e

em que, quando o upmix opera no modo de não-predição, o sinal residual representa o sinal lateral, e o upmix gera o canal esquerdo com base na soma do sinal descendente e no sinal residual passado pelo decodificador, e gera o canal direito com base na diferença entre o sinal de downmix e o sinal residual.

28. Meio legível por computador não transitório, **caracterizado pelo fato de que** quando executado por um processador, executa o método conforme definido na reivindicação 27.

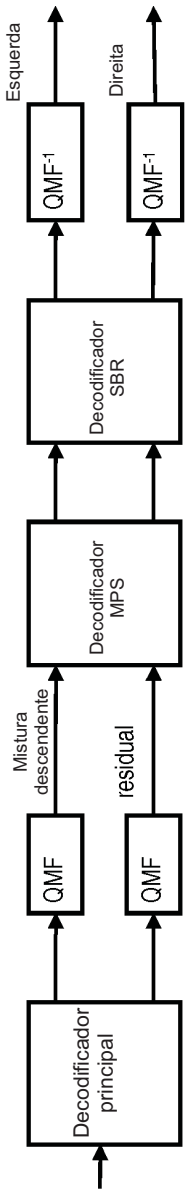

Esquerda
Direita
QMF-1
QMF-1
Decodificador SBR
Decodificador MPS
Mistura descendente
residual
QMF
QMF
Decodificador principal

Fig. 1A

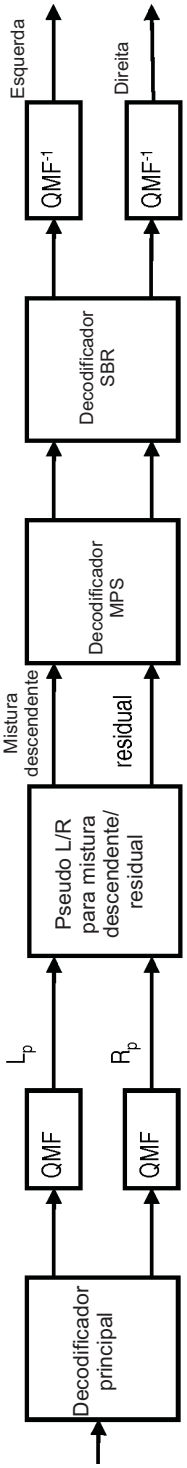

Esquerda
Direita
QMF-1
QMF-1
Decodificador SBR
Decodificador MPS
Mistura descendente
residual
Pseudo L/R para mistura descendente/ residual
L_p
R_p
QMF
QMF
Decodificador principal

Fig. 1B

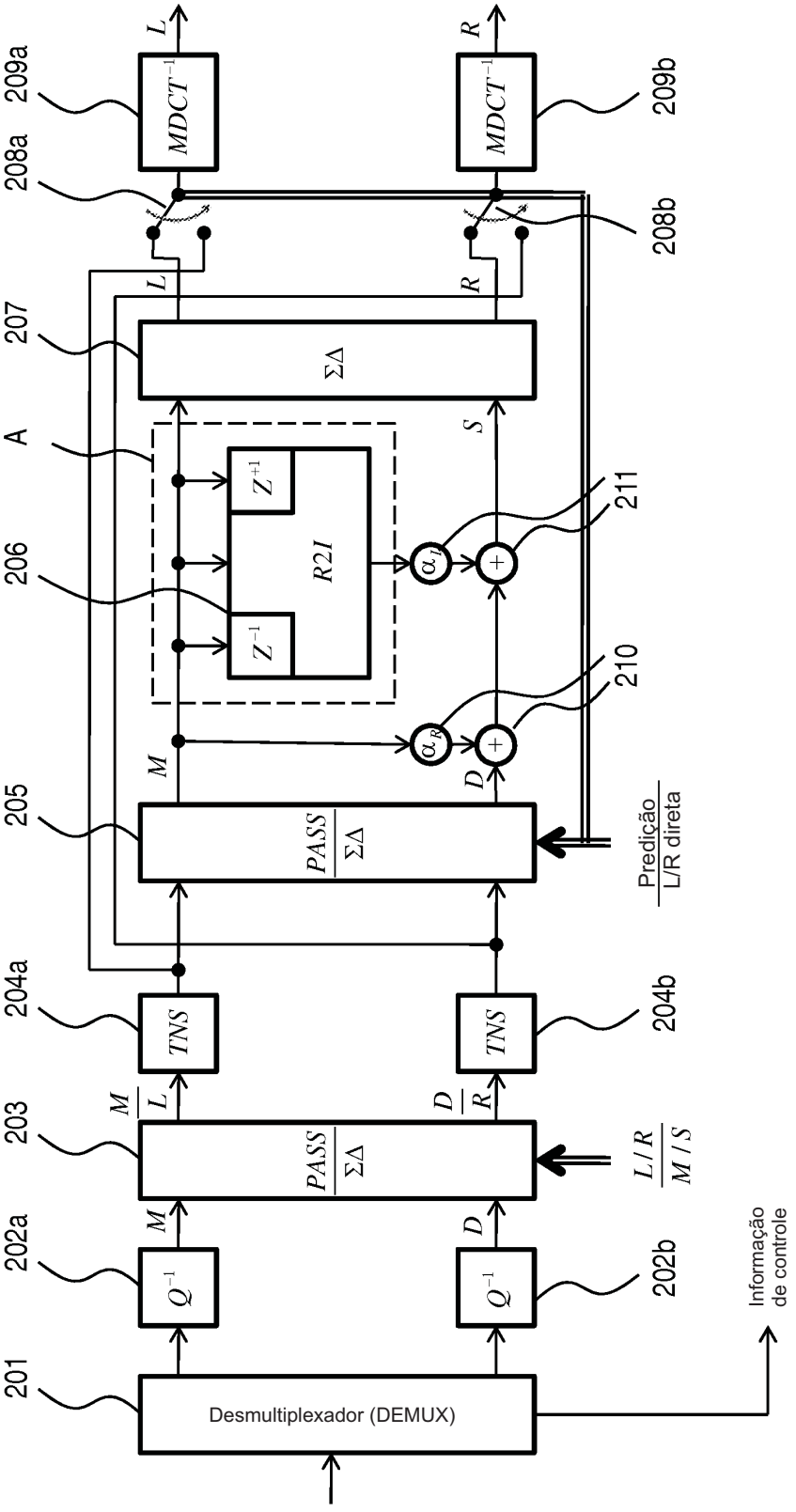

Desmultiplexador (DEMUX)
201
202a
202b
Q^{-1}
203
204a
204b
TNS
205
PASS / ΣΔ
206
A
Z^{-1}
R2I
Z^{+1}
207
ΣΔ
208a
208b
209a
209b
$MDCT^{-1}$
210
211
α_R
α_I
L
R
M
S
D
Predição L/R direta
L / R
M / S
Informação de controle

Fig. 2

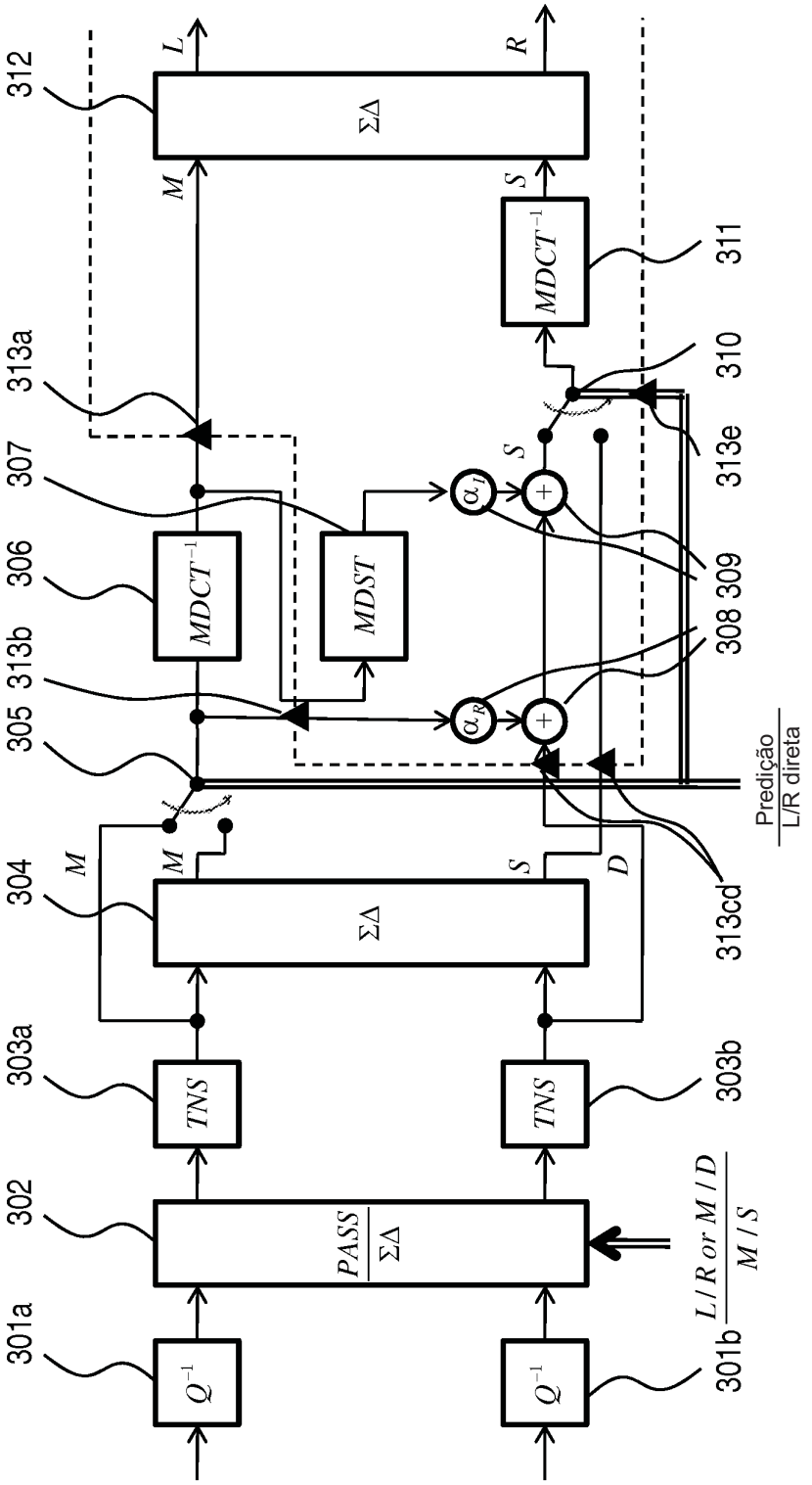

312
L
R
ΣΔ
M
S
MDCT⁻¹
311
313a
310
307
S
313e
306
MDCT⁻¹
MDST
309
313b
308
305
Predição
L/R direta
304
M
M
S
D
313cd
ΣΔ
303a
303b
TNS
TNS
302
PASS
ΣΔ
L/R or M/D
M/S
301a
301b
Q⁻¹
Q⁻¹

Fig. 3

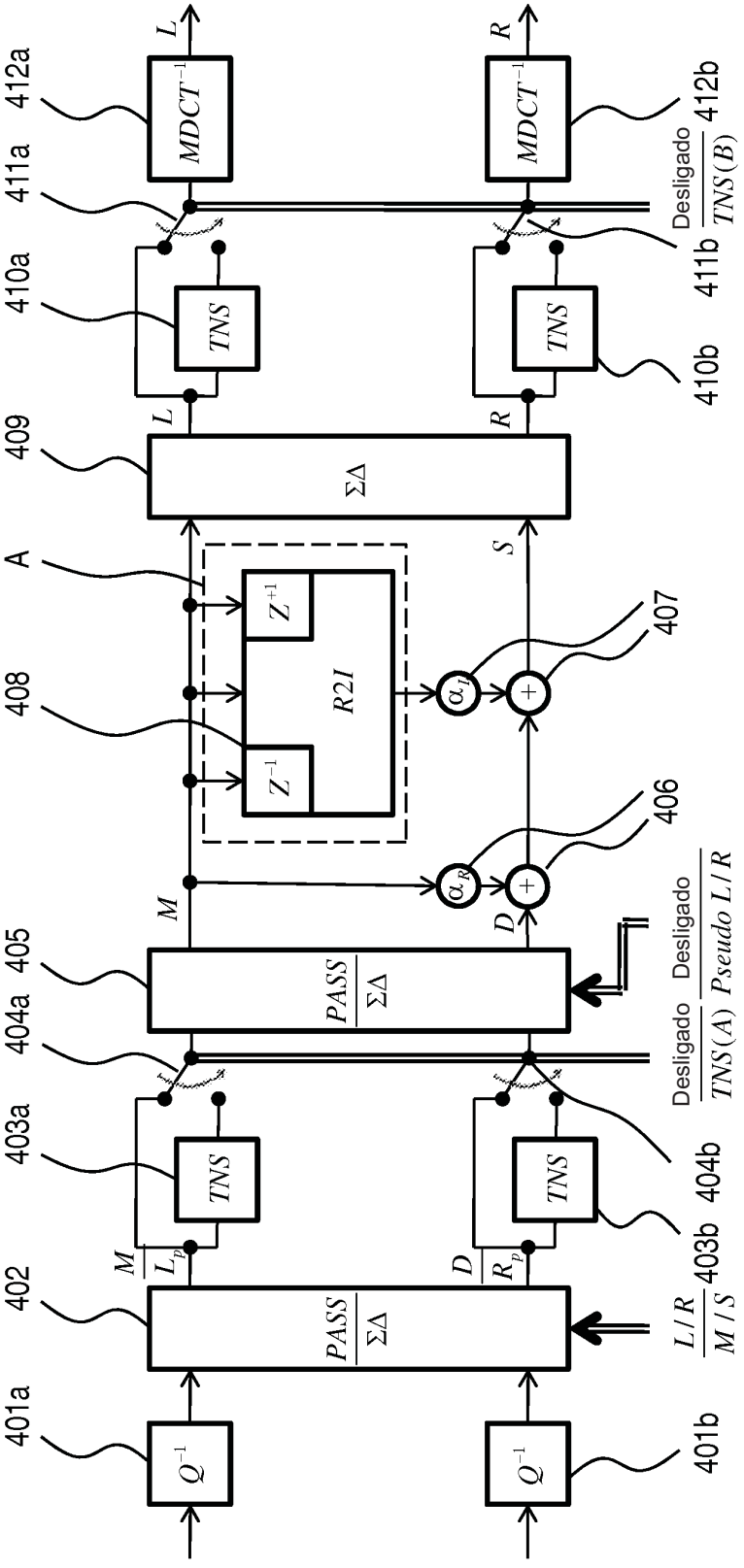
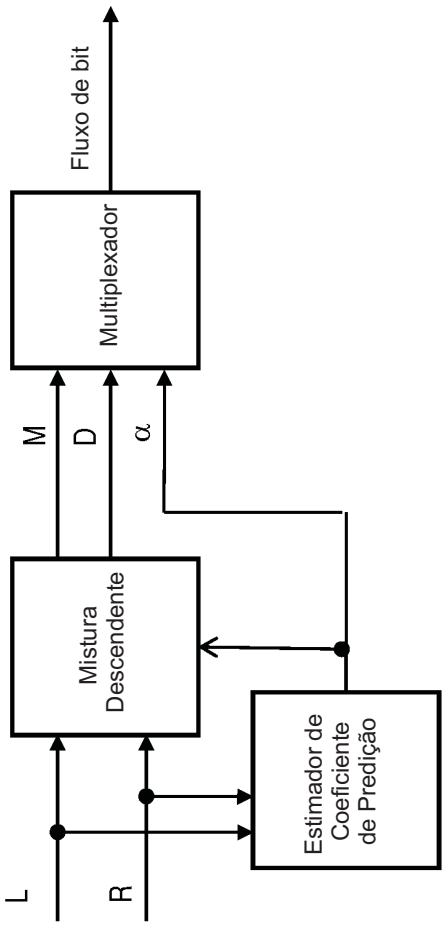

401a
401b
Q^{-1}
402
$PASS$ / $\Sigma\Delta$
L/R / M/S
403a
403b
TNS
M / L_p
D / R_p
404a
404b
Desligado / $TNS(A)$
Desligado / $Pseudo\ L/R$
405
406
407
408
409
410a
410b
411a
411b
Desligado / $TNS(B)$
412a
412b
$MDCT^{-1}$
Z^{-1}
$R2I$
α_I
α_R
$\Sigma\Delta$
A
M
D
S
L
R

Fig. 4

Fig. 5


Fluxo de bit
Multiplexador
M
D
α
Mistura Descendente
Estimador de Coeficiente de Predição
L
R

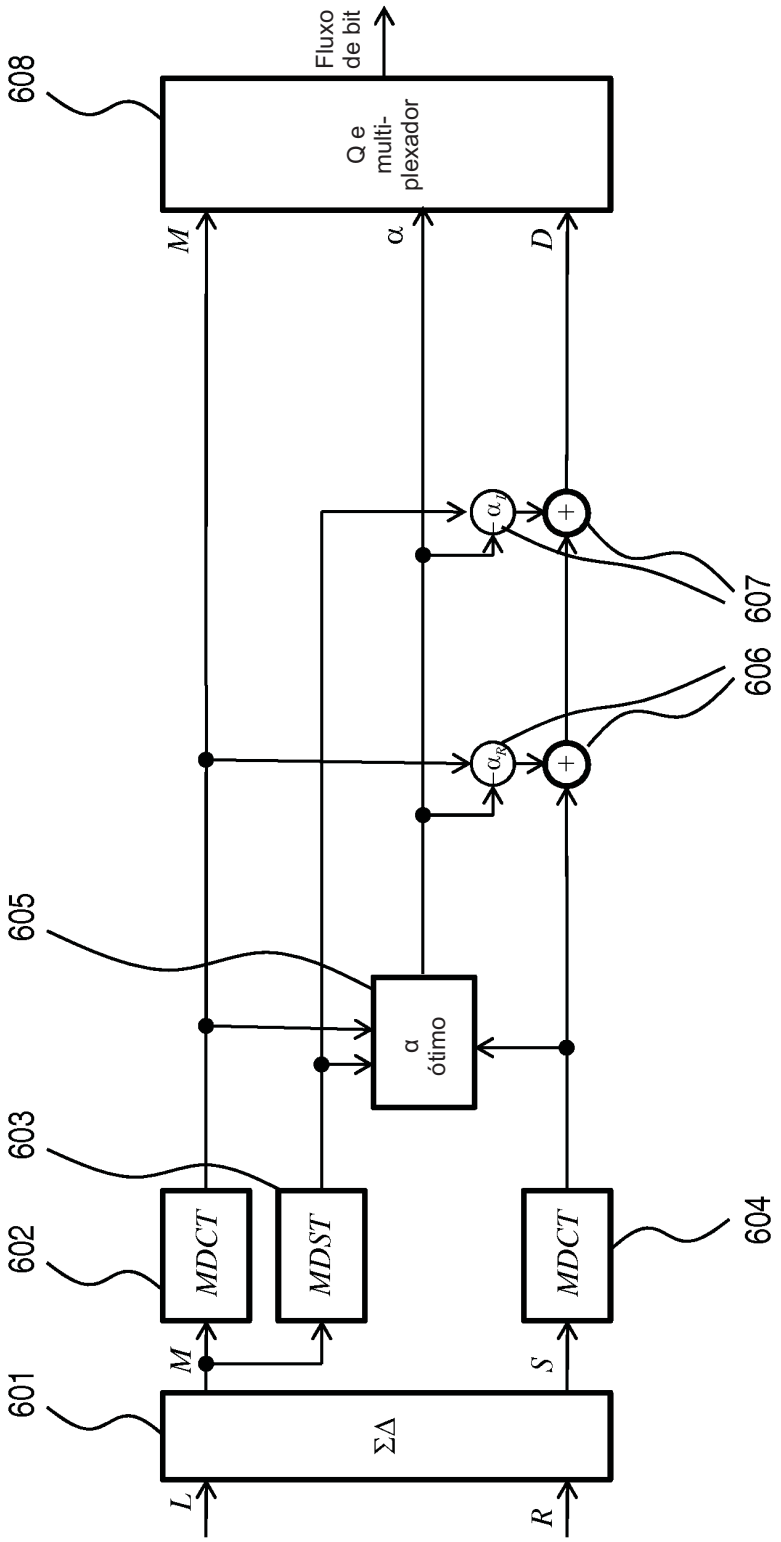

608
Fluxo de bit
Q e multi-plexador
M
α
D
607
606
605
α ótimo
603
602
MDCT
MDST
601
ΣΔ
L
M
S
R
MDCT
604

Fig. 6

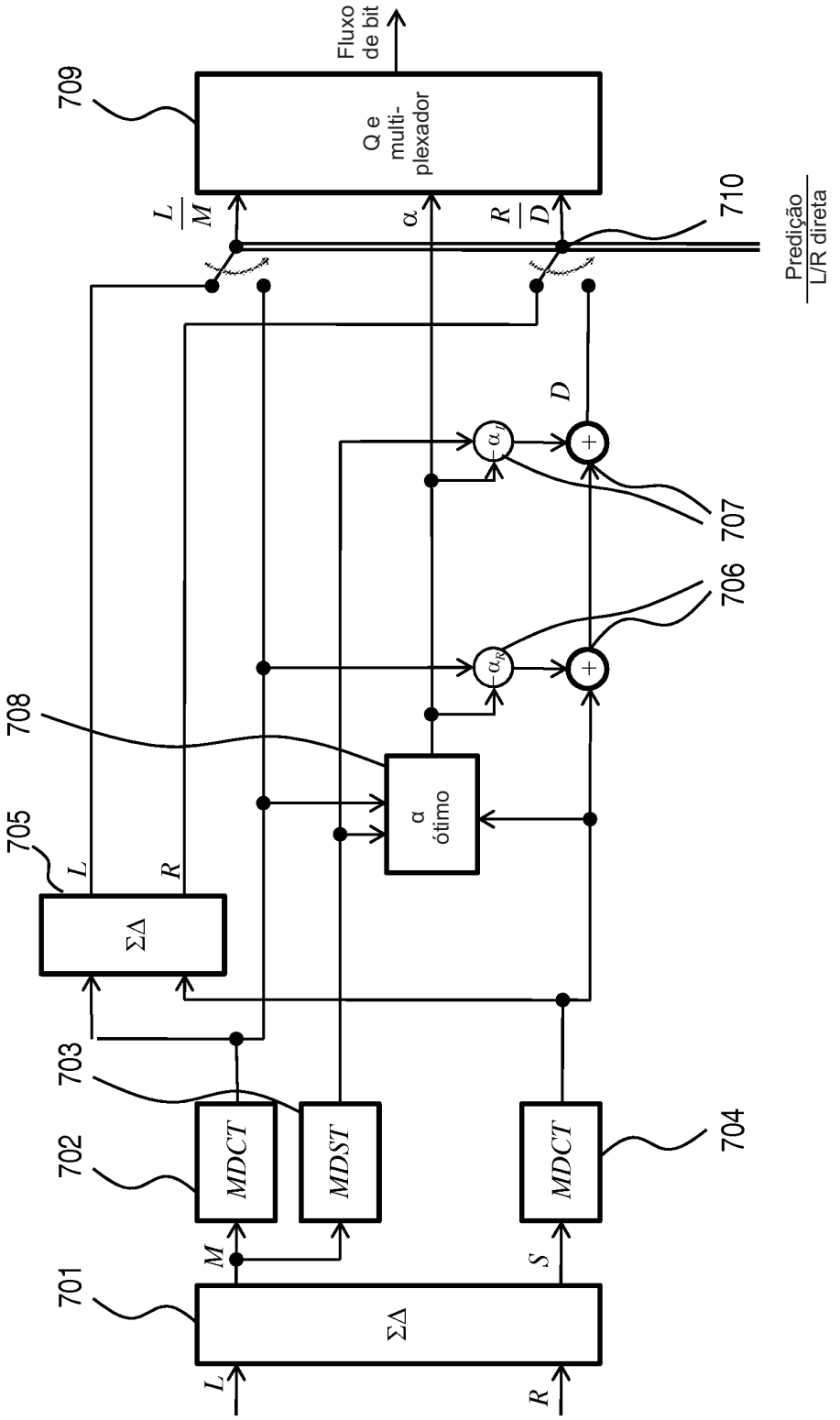

Fluxo de bit
709
Q e multi-plexador
L/M
α
R/D
710
Predição L/R direta
D
707
706
708
α ótimo
705
L
R
ΣΔ
703
702
MDCT
MDST
MDCT
704
M
S
701
ΣΔ
L
R

Fig. 7

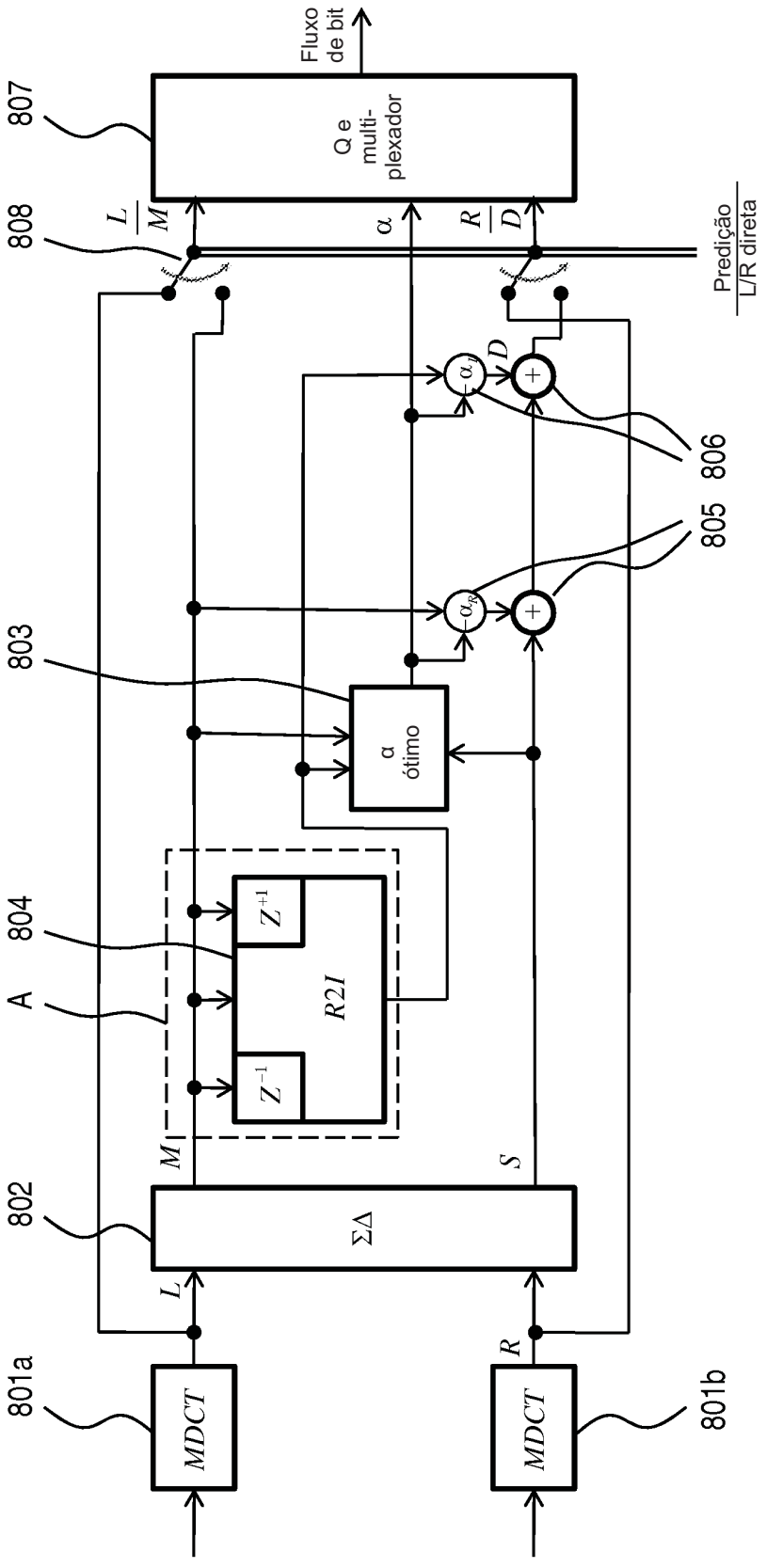

Fluxo de bit
807
Q e multi-plexador
808
L/M
α
R/D
Predição L/R direta
806
805
803
α ótimo
804
A
Z+1
R2I
Z-1
M
S
802
ΣΔ
L
R
801a
MDCT
801b
MDCT

Fig. 8

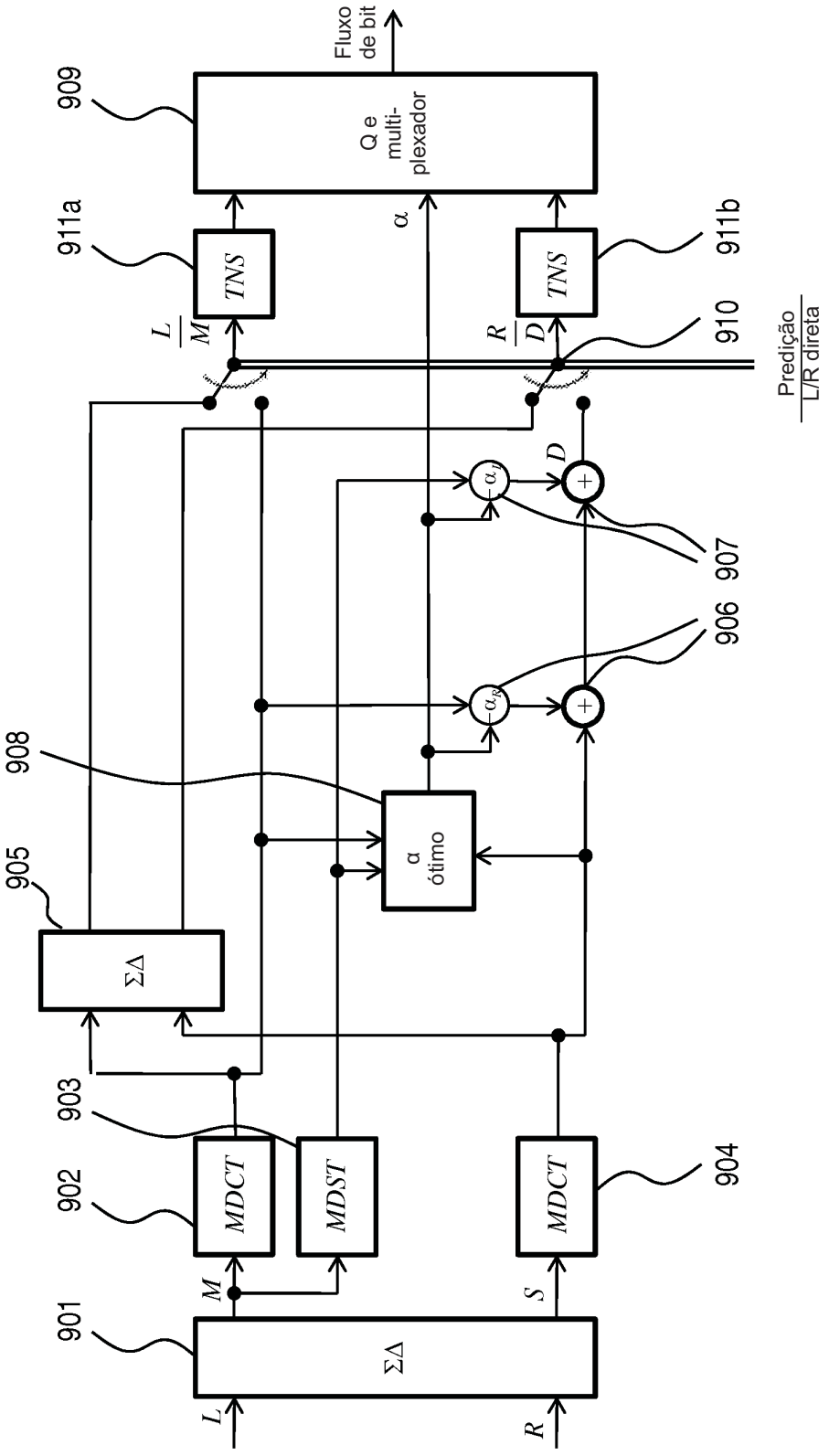

Fluxo de bit
909
Q e multi-plexador
911a
TNS
911b
TNS
α
L/M
R/D
910
Predição L/R direta
D
907
906
908
α ótimo
905
ΣΔ
903
902
MDCT
MDST
MDCT
904
M
S
901
ΣΔ
L
R

Fig. 9

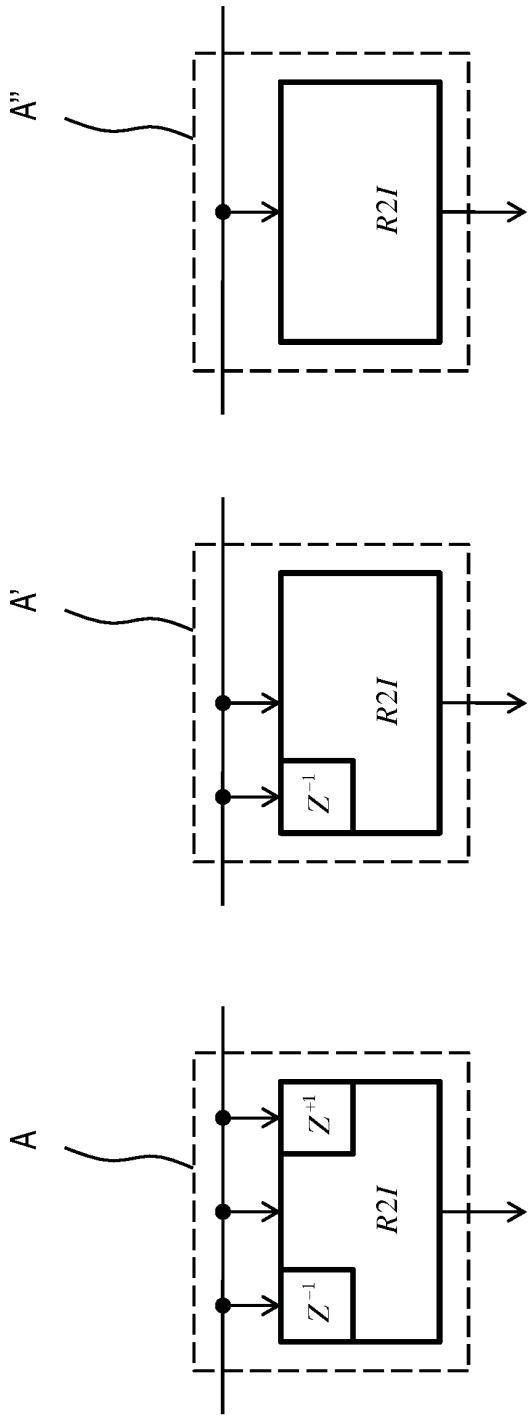


Fig. 10

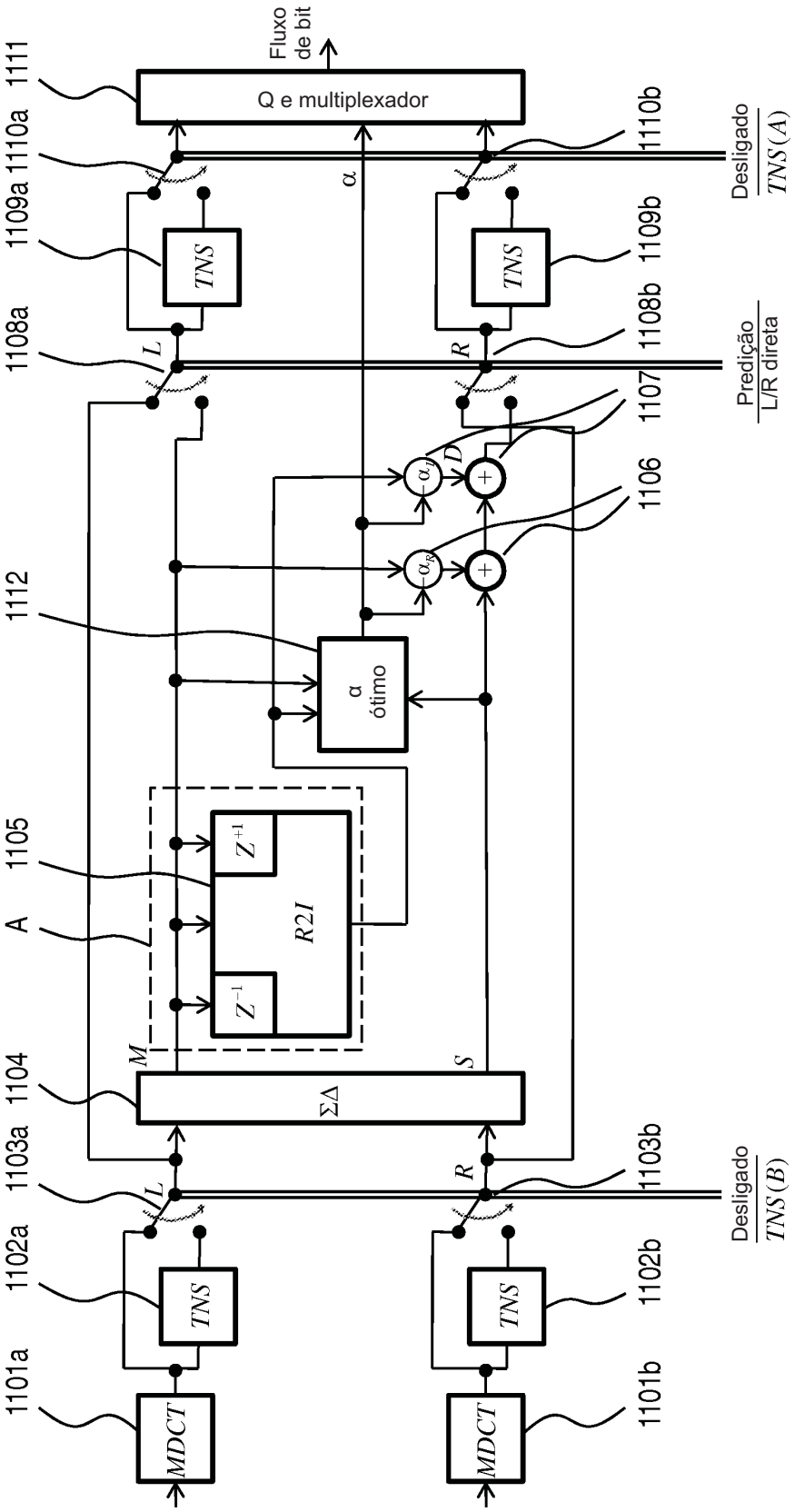

Fluxo de bit
Q e multiplexador
1111
1110a
1110b
Desligado
TNS(A)
1109a
1109b
TNS
1108a
1108b
L
R
Predição
L/R direta
1107
1106
α
1112
α ótimo
1105
A
Z+1
R2I
Z-1
M
S
1104
ΣΔ
1103a
1103b
Desligado
TNS(B)
1102a
1102b
1101a
1101b
MDCT

Fig. 11

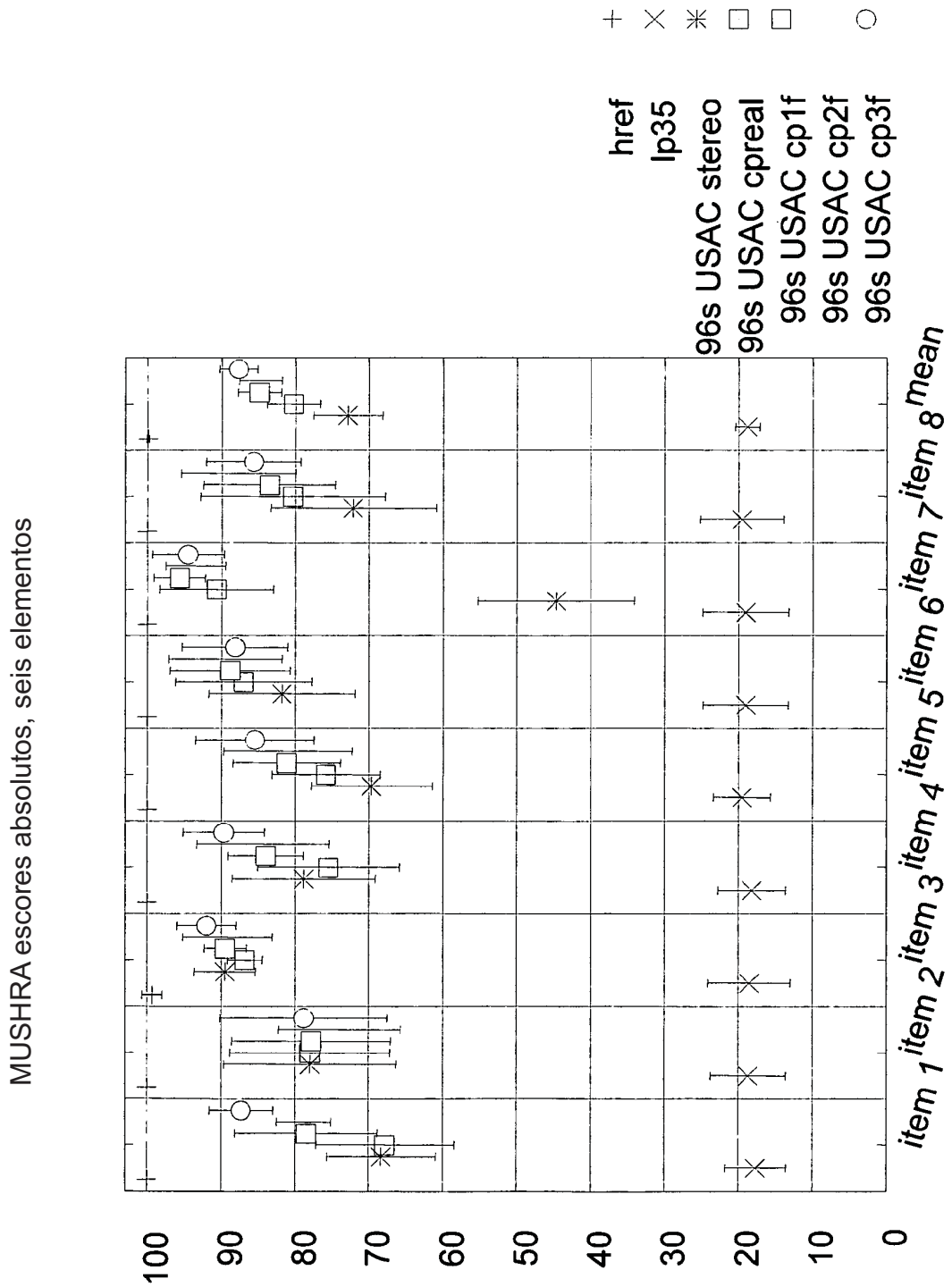

MUSHRA escores absolutos, seis elementos
100
90
80
70
60
50
40
30
20
10
0
item 1 item 2 item 3 item 4 item 5 item 6 item 7 item 8 mean
href
lp35
96s USAC stereo
96s USAC cpreal
96s USAC cp1f
96s USAC cp2f
96s USAC cp3f

Fig. 12

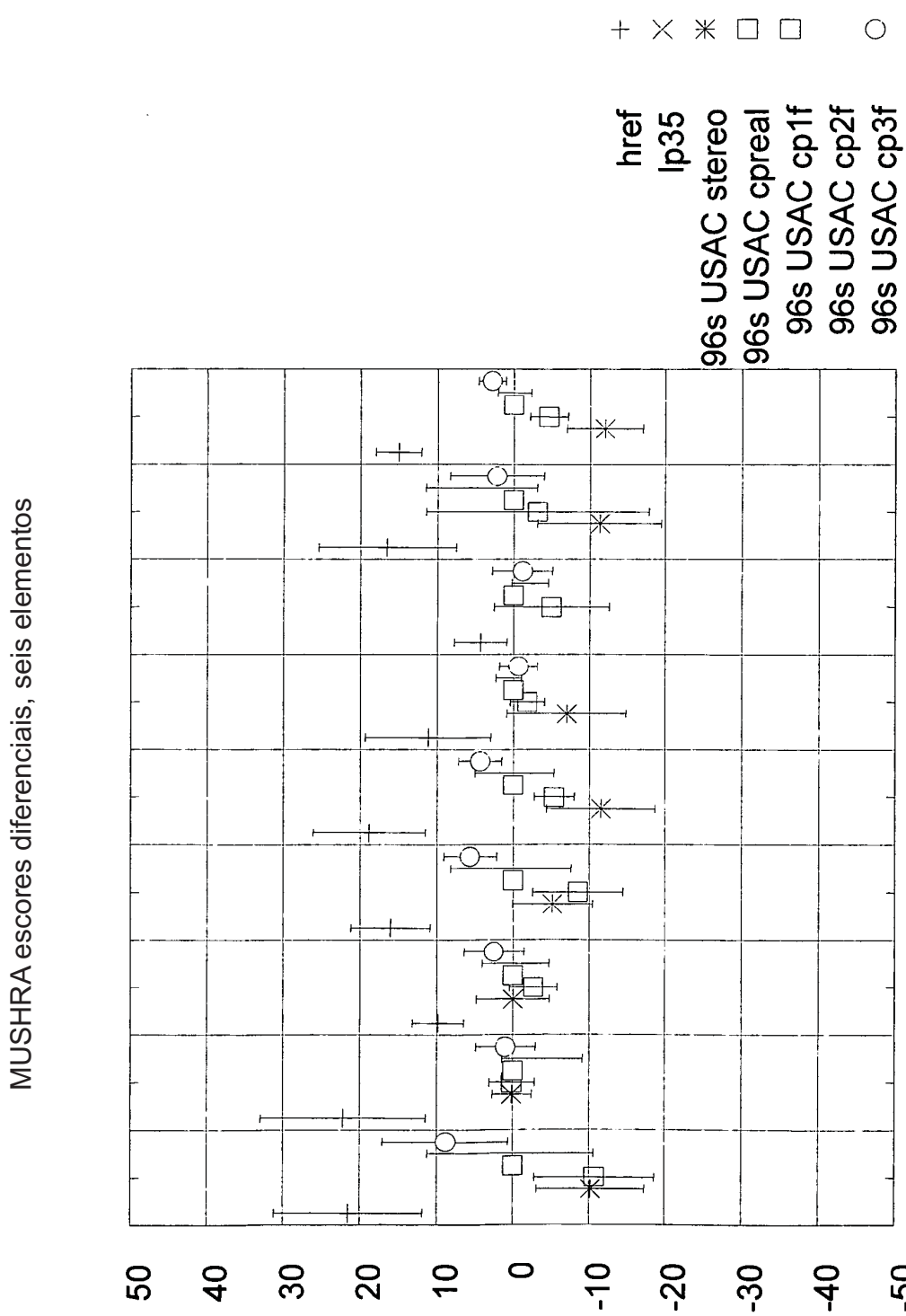

MUSHRA escores diferenciais, seis elementos
50
40
30
20
10
0
-10
-20
-30
-40
-50
item 1 item 2 item 3 item 4 item 5 item 6 item 7 item 8 mean
href
lp35
96s USAC stereo
96s USAC cpreal
96s USAC cp1f
96s USAC cp2f
96s USAC cp3f

Fig. 13

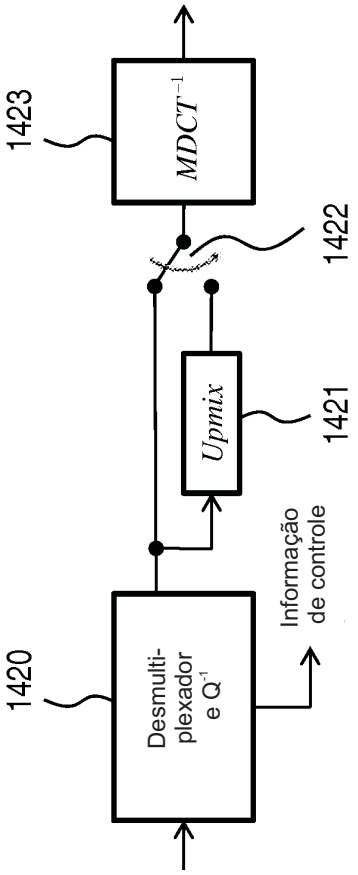


Fig. 14B

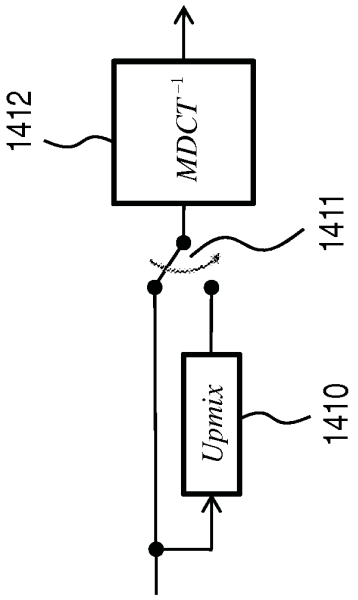


Fig. 14A

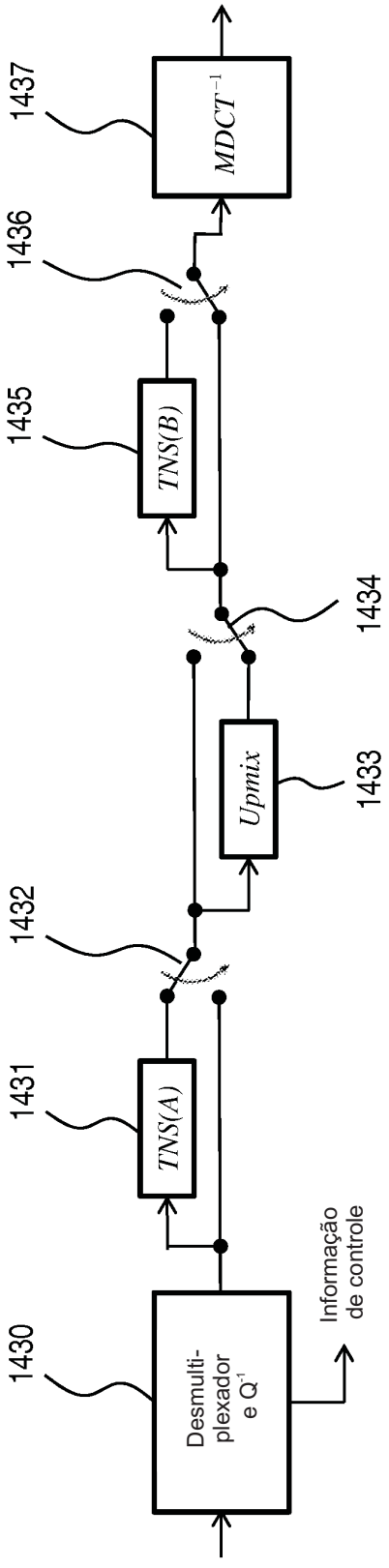


Fig. 14C

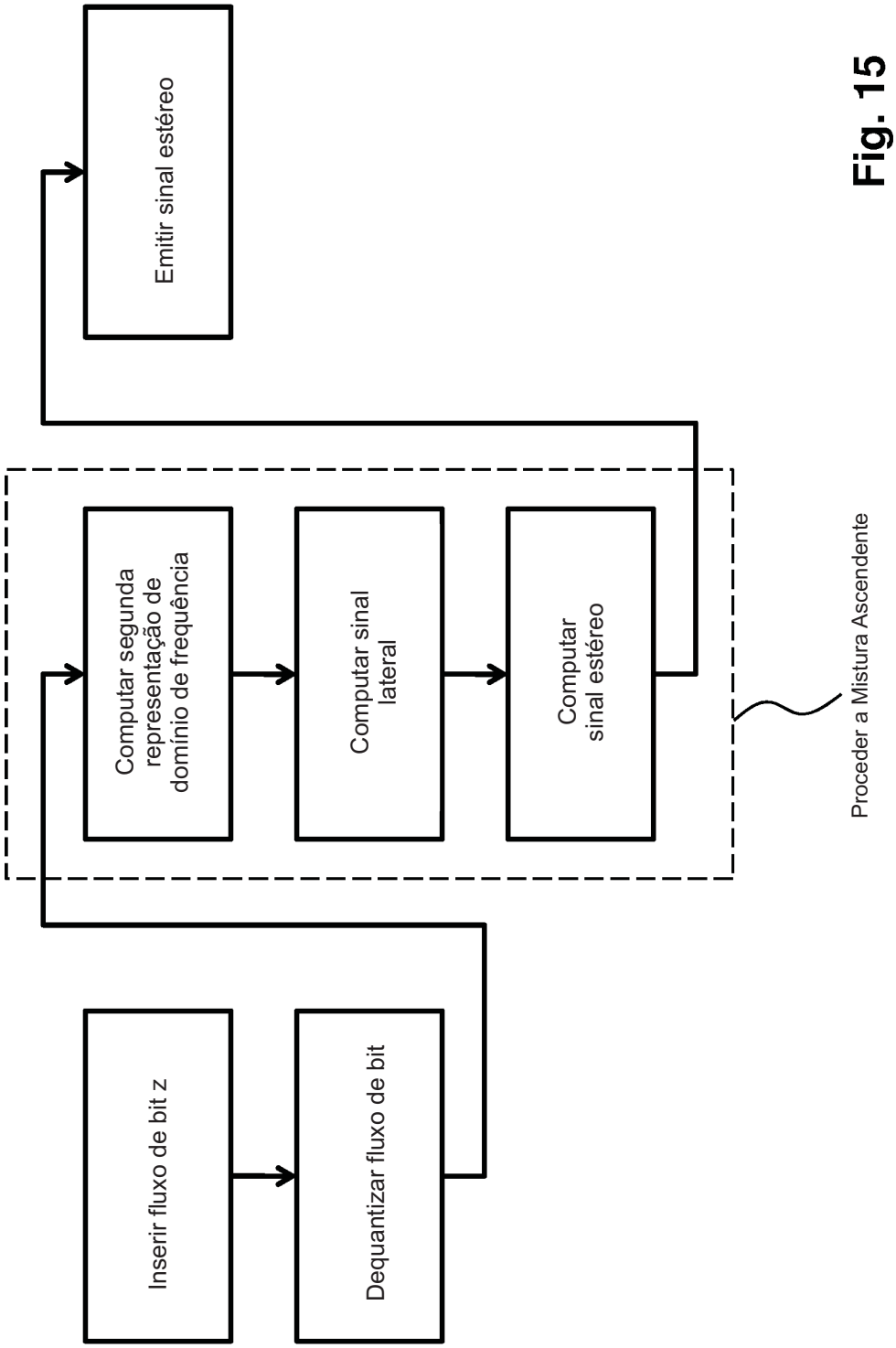

Emitir sinal estéreo
Computar segunda representação de domínio de frequência
Computar sinal lateral
Computar sinal estéreo
Proceder a Mistura Ascendente
Inserir fluxo de bit z
Dequantizar fluxo de bit

Fig. 15

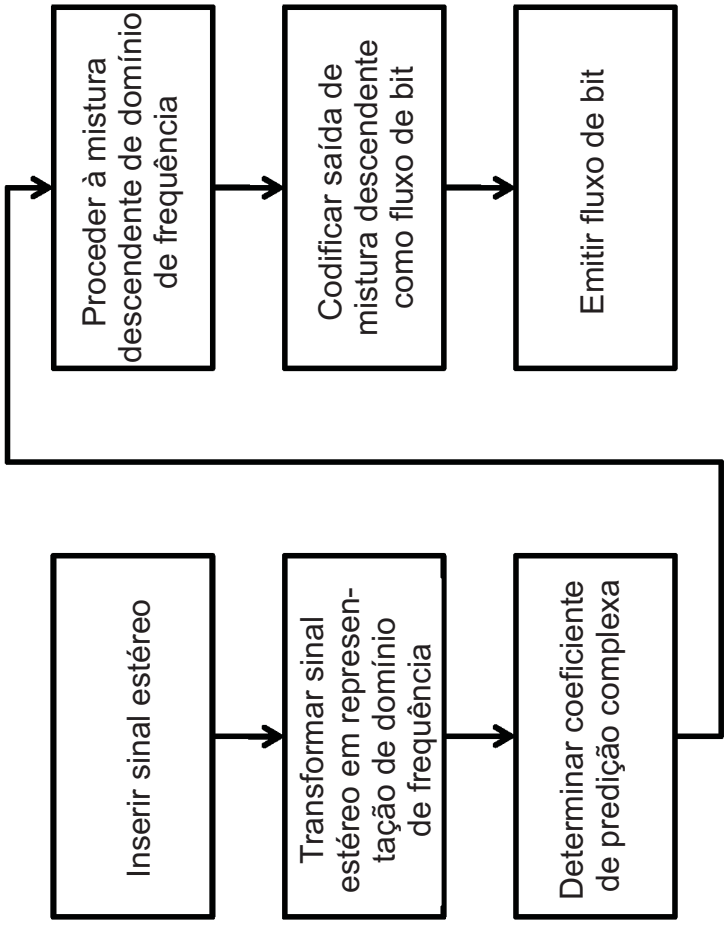

Inserir sinal estéreo
Transformar sinal estéreo em represen-tação de domínio de frequência
Determinar coeficiente de predição complexa
Proceder à mistura descendente de domínio de frequência
Codificar saída de mistura descendente como fluxo de bit
Emitir fluxo de bit

Fig. 16