



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

(11) PI 0815972-6 B1



(22) Data do Depósito: 26/08/2008

(45) Data de Concessão: 04/02/2020

(54) Título: MÉTODO PARA RECUPERAÇÃO DE ESPECTRO EM DECODIFICAÇÃO ESPECTRAL DE UM SINAL DE ÁUDIO, MÉTODO PARA USO EM CODIFICAÇÃO ESPECTRAL DE UM SINAL DE ÁUDIO, DECODIFICADOR, E, CODIFICADOR

(51) Int.Cl.: G10L 21/038; G10L 19/028; G10L 19/035; G10L 19/032; G10L 19/02.

(52) CPC: G10L 21/038; G10L 19/028; G10L 19/035; G10L 19/032; G10L 19/0204.

(30) Prioridade Unionista: 27/08/2007 US 60/968134.

(73) Titular(es): TELEFONAKTIEBOLAGET LM ERICSSON (PUBL).

(72) Inventor(es): GUSTAF ULLBERG; MANUEL BRIAND; ANISSE TALEB.

(86) Pedido PCT: PCT SE2008050969 de 26/08/2008

(87) Publicação PCT: WO 2009/029037 de 05/03/2009

(85) Data do Início da Fase Nacional: 25/02/2010

(57) Resumo: MÉTODO PARA RECUPERAÇÃO DE ESPECTRO EM DECODIFICAÇÃO ESPECTRAL DE UM SINAL DE ÁUDIO, MÉTODO PARA USO EM CODIFICAÇÃO ESPECTRAL DE UM SINAL DE ÁUDIO, DECODIFICADOR, E, CODIFICADOR Método para recuperação de espectro em decodificação espectral de um sinal de áudio compreende a obtenção (210) de um conjunto inicial de coeficientes espectrais, representando o sinal de áudio e determinação (212) de uma frequência de transição. A frequência de transição é adaptada a um conteúdo espectral do sinal de áudio. Os furos espectrais no conjunto inicial de coeficientes espectrais abaixo da frequência de transição são preenchidos com ruídos (214) e o conjunto inicial de coeficientes espectrais é estendido em largura de banda (216) acima da frequência de transição. Decodificadores e codificadores sendo dispostos para realização de parte ou de todo o método também são ilustrados.

“MÉTODO PARA RECUPERAÇÃO DE ESPECTRO EM DECODIFICAÇÃO ESPECTRAL DE UM SINAL DE ÁUDIO, MÉTODO PARA USO EM CODIFICAÇÃO ESPECTRAL DE UM SINAL DE ÁUDIO, DECODIFICADOR, E, CODIFICADOR”

CAMPO TÉCNICO

[001] A presente invenção se refere, em geral, a métodos e dispositivos para codificação e decodificação de sinais de áudio e, em particular, a métodos e dispositivos para preenchimento de espectro.

FUNDAMENTOS

[002] Quando sinais de áudio devem ser armazenados e/ou transmitidos, uma abordagem padrão hoje é codificar os sinais de áudio em uma representação digital de acordo com esquemas diferentes. A fim de economizar na capacidade de armazenamento e/ou transmissão, é desejo geral reduzir o tamanho da representação digital necessária para permitir a reconstrução dos sinais de áudio com qualidade suficiente. A troca entre o tamanho do sinal codificado e qualidade do sinal depende da aplicação real.

[003] Codificadores de áudio baseados em transformação comprimem sinais de áudio pela quantificação dos coeficientes de transformação. Para permitir baixas taxas de bits, quantificadores poderiam concentrar os bits disponíveis nos coeficientes mais energéticos e perceptualmente relevantes e transmitir apenas aqueles, deixando "furos espectrais" de coeficientes não quantificados no espectro de frequências.

[004] A chamada tecnologia de SBR (Spectral Band Replication - Réplica de Banda Espectral), veja, por exemplo, 3GPP TS 26.404 V6.0.0 (2004-09), "Enhanced aacPlus general audio codec - encoder SBR part (Release 6)", 2004 [1], fecha a folga entre o sinal de banda limitada de um codificador perceptual convencional e a largura de banda audível de, aproximadamente, 15 kHz. A idéia geral atrás da SBR é recriar os conteúdos de alta frequência que estão faltando de um sinal decodificado em uma

maneira perceptualmente precisa. As frequências acima de 15 kHz são menos importantes de um ponto de vista psicoacústico, mas também podem ser reconstruídas. Contudo, a SBR não pode ser usada como um codec autônomo. Ela sempre opera em conjunto com um codec de forma de onda convencional, um chamado codec de núcleo. O codec de núcleo é responsável pela transmissão da parte inferior do espectro original, enquanto o decodificador - SBR, que é, principalmente, um pós-processo para o decodificador de forma de onda convencional, reconstrói a faixa de frequência não transmitida. Os valores espectrais da banda alta não são transmitidos diretamente, como em codecs convencionais. O sistema combinado oferece um ganho de codificação superior ao ganho do codec de núcleo sozinho.

[005] A metodologia de SBR conta com a definição de uma frequência de transição fixa entre baixas frequências, perceptualmente relevantes, codificadas, de banda baixa e frequências altas menos relevantes, não codificadas de banda alta. Contudo, na prática, essa frequência de transição conta com o conteúdo de áudio do sinal original. Em outras palavras, de um sinal para outro, a frequência de transição apropriada pode variar bastante. Esse é, por exemplo, o caso quando comparando o discurso limpo e sinais musicais de banda completa.

[006] Os "furos espectrais" do espectro decodificado podem ser divididos em duas espécies. A primeira é a de pequenos furos em frequências mais baixas devido ao efeito de mascaramento instantâneo, veja, por exemplo, J.D. Johnston, "Estimation of Perceptual Entropy Using Noise Masking Criteria", Proc. ICASSP, pp. 2524-2527, Maio de 1988 [2]. A segunda é uma de furos maiores em altas frequências, resultantes da saturação pelo limite absoluto de audição e a adição de mascaramento [2]. A SBR se preocupa, principalmente, com a segunda espécie.

[007] Além disso, um codec de áudio típico, baseado nesse método, que objetiva o preenchimento do "furo espectral", isto é, coeficientes não

codificados, para as frequências mais altas, isto é, a segunda espécie de "furos espectrais", de preferência, será capaz de preencher os furos espectrais através de todo o espectro. Na verdade, mesmo se um codec de SBR for capaz de distribuir um sinal de áudio de largura de banda completa, as altas frequências reconstruídas não mascararão os artefatos importunos introduzidos pela codificação, isto é, quantificação, da banda baixa, isto é, as baixas frequências perceptualmente relevantes.

SUMÁRIO

[008] Um objetivo geral da presente invenção é proporcionar métodos e dispositivos para permitir a supressão eficiente de artefatos perceptuais causados pelos furos espectrais através de um sinal de áudio de banda completa.

[009] Os objetivos acima são alcançados por métodos e dispositivos de acordo com as reivindicações de patente anexas. Em palavras gerais de acordo com um primeiro aspecto, um método para recuperação de espectro na decodificação espectral de um sinal de áudio compreende a obtenção de um conjunto inicial de coeficientes espectrais, representando o sinal de áudio e determinando uma frequência de transição. A frequência de transição é adaptada a um conteúdo espectral do sinal de áudio. Os furos espectrais no conjunto inicial de coeficientes espectrais abaixo da frequência de transição são ruídos cheios e o conjunto inicial de coeficientes espectrais é largura de banda estendida acima da frequência de transição.

[0010] De acordo com um segundo aspecto, um método para uso em codificação espectral de um sinal de áudio compreende a determinação de uma frequência de transição para um conjunto inicial de coeficientes espectrais, representando o sinal de áudio. A frequência de transição é adaptada a um conteúdo espectral do sinal de áudio. A frequência de transição define uma borda entre uma faixa de frequências, destinada a ser um objeto para preenchimento de ruído de furos espectrais e uma faixa de frequências,

destinada a ser um objeto para a extensão de largura de banda.

[0011] De acordo com um terceiro aspecto, um decodificador para decodificação espectral de um sinal de áudio compreende uma entrada para obtenção de um conjunto inicial de coeficientes espectrais, representando o sinal de áudio e conjunto de circuitos de determinação de transição disposto para determinação de uma frequência de transição. A frequência de transição é adaptada a um conteúdo espectral do sinal de áudio. A decodificação compreende um preenchedor de ruído para preenchimento de ruído de furos espectrais no conjunto inicial dos coeficientes espectrais abaixo da frequência de transição e um extensor de largura de banda disposto para largura de banda, estendendo o conjunto inicial de coeficientes espectrais acima da frequência de transição.

[0012] De acordo com um quarto aspecto, um codificador para codificação espectral de um sinal de áudio compreende o conjunto de circuitos de determinação de transição disposto para determinar uma frequência de transição para um conjunto inicial de coeficientes espectrais, representando o sinal de áudio. A frequência de transição é adaptada a um conteúdo espectral do sinal de áudio. A frequência de transição define uma borda entre uma faixa de frequências, destinada a ser um objeto para preenchimento de ruído de furos espectrais e uma faixa de frequências, destinada a ser um objeto para extensão de largura de banda.

[0013] A presente invenção tem um número de vantagens. Uma vantagem é que o uso da frequência de transição permite o uso de um preenchimento de espectro combinado, usando preenchimento de ruído e extensão de largura de banda. Além disso, a frequência de transição é definida adaptativamente, por exemplo, de acordo com o esquema de codificação usado, o que torna o preenchimento de espectro dependente, por exemplo, de resolução de frequência. Qualquer codec de fala e/ou de áudio usando esse método é capaz de distribuir uma alta qualidade, isto é, com artefatos

importunos reduzidos e sinal de áudio de largura de banda completa. O método é flexível no sentido de que pode ser combinado com qualquer espécie de representação de frequência (DCT, MDCTY, etc.) ou bancos de filtros, isto é, com qualquer codec (perceptual, paramétrico, etc.).

BREVE DESCRIÇÃO DOS DESENHOS

[0014] A invenção, junto com outros objetivos e suas vantagens, pode ser mais bem compreendida fazendo-se referência à descrição seguinte, tomada junto com os desenhos anexos, em que:

A figura 1 é um diagrama em blocos esquemático e um sistema de codec;

A figura 2 é um diagrama em blocos esquemático de uma modalidade de um codificador de sinal de áudio de acordo com a presente invenção;

A figura 3 é uma ilustração esquemática de coeficientes espectrais, seus grupos e bandas de frequência;

A figura 4 é um diagrama em blocos esquemático de uma modalidade de um decodificador de sinal de áudio de acordo com a presente invenção;

As figuras 5A - C são ilustrações de modalidades de princípios para encontrar uma frequência de transição;

A figura 6 é um diagrama de fluxo das etapas de uma modalidade de um método de acordo com a presente invenção; e

A figura 7 é um fluxograma de uma etapa de uma modalidade de um método de manipulação de sinal de acordo com a presente invenção.

DESCRIÇÃO DETALHADA

[0015] Por todos os desenhos, os mesmos números de referência são usados para elementos similares ou correspondentes.

[0016] Uma modalidade de um sistema geral de codec para sinais de áudio é ilustrada esquematicamente na figura 1. Uma fonte de áudio 10 dá

origem a um sinal de áudio 15. O sinal de áudio 15 é manipulado em um codificador 20, que produz um fluxo binário 25, compreendendo dados representando o sinal de áudio 15. O fluxo binário 25 pode ser transmitido como, por exemplo, no caso de comunicação multimídia por uma disposição de transmissão e/ou armazenamento 30. A disposição de transmissão e/ou armazenamento 30, opcionalmente, também pode compreender alguma capacidade de armazenamento. O fluxo binário 25 também só pode ser armazenado na disposição de transmissão e/ou armazenamento 30, exatamente introduzindo um retardo de tempo na utilização do furo passante. A disposição de transmissão e/ou armazenamento 30 é, assim, uma disposição introduzindo pelo menos um dentre um reposicionamento espacial ou retardo de tempo do fluxo binário 25. Quando sendo usado, o fluxo binário 25 é manipulado em um decodificador 40, que produz uma saída de áudio 35 dos dados compreendidos no fluxo binário. Tipicamente, a saída de áudio 35 parecerá o sinal de áudio original 15 bem como possível sob certas restrições.

[0017] Aplicações em tempo real, o retardo de tempo entre a produção do sinal de áudio original 15 e, tipicamente, não é permitido que a saída de áudio 35 produzida exceda um certo tempo. Se os recursos de transmissão, ao mesmo tempo, forem limitados, a taxa de bits disponível também é, tipicamente, baixa. A fim de utilizar a taxa de bits disponível da melhor maneira possível, a codificação de áudio perceptual foi desenvolvida. A codificação de áudio perceptual, portanto, tem se tornado uma parte importante para muitos serviços de multimídia hoje. O princípio básico é converter o sinal de áudio em coeficientes espectrais em um domínio da frequência e usando um modelo perceptual para determinar um mascaramento dependente de frequência e tempo dos coeficientes espectrais.

[0018] A figura 2 ilustra uma modalidade de um codificador de áudio 20 de acordo com a presente invenção. Nesta modalidade particular, o codificador de áudio perceptual 20 é um codificador espectral baseado em um

transformador perceptual ou banco de filtro perceptual. Uma fonte de áudio 15 é recebida, compreendendo os quadros de sinais de áudio $x[n]$.

[0019] Em um codificador espectral típico, um conversor 21 é disposto para converter o sinal de áudio de domínio de tempo 15 em um conjunto 24 de coeficientes espectrais $X_b[n]$ de um domínio da frequência. Em um codificador de transformação típico, a conversão pode ser realizada, por exemplo, por uma Transformada de Fourier Discreta (DFT), a Transformada de Cosseno Discreta (DCT) ou a Transformada de Cosseno Discreta Modificada (MDCT). O conversor 21 pode, assim, tipicamente, ser constituído por um transformador espectral. Os detalhes do transformador real não são de importância particular para as idéias básicas da presente invenção e, portanto, não são aqui discutidos.

[0020] O conjunto 24 de coeficientes espectrais, isto é, uma representação de frequência do sinal de áudio de entrada é proporcionado a uma sessão de quantificação e codificação 28, onde os coeficientes espectrais são quantificados e codificados. Tipicamente, a codificação está operando para concentrar os bits disponíveis nos coeficientes mais energéticos e perceptualmente relevantes. Isso pode ser realizado usando, por exemplo, diferentes espécies de limites de mascaramento ou reduções de largura de banda. O resultado, tipicamente, será os "furos espectrais" de coeficientes não codificados no espectro de frequência. Em outras palavras, alguns dos coeficientes são deixados de propósito, uma vez que eles são, perceptualmente, menos importantes para não ocupar recursos de transmissão mais necessários para outras finalidades. Esses furos espectrais podem, então, através de diferentes estratégias de reconstrução ser corrigidos ou reconstruídos no lado do decodificador. Tipicamente, furos espectrais de duas espécies aparecem. A primeira espécie compreende furos espectrais individuais ou uns poucos vizinhos, que ocorrem em diferentes locais, principalmente na região de baixa frequência. O segundo tipo é um grupo

mais ou menos contínuo de furos espectrais na extremidade de alta frequência do espectro.

[0021] De acordo com a presente invenção, é favorável tratar essas duas espécies diferentes de furos espectrais de diferentes maneiras, a fim de obter um preenchimento de espectro tão eficiente quanto possível. Um parâmetro para determinar é, então, uma frequência de transição, em que diferentes abordagens de preenchimento se encontram, uma chamada frequência de transição. Uma vez que os furos espectrais diferem entre diferentes espécies de sinais de áudio, a escolha ótima de frequência de transição também difere. De acordo com a presente versão, a sequência de transição é adaptada a um conteúdo espectral do sinal de áudio. Tipicamente, a frequência de transição é adaptada a um conteúdo espectral de um quadro presente do sinal de áudio, porém, a frequência de transição também pode depender de conteúdos de espectro de quadros anteriores do sinal de áudio e, se não houver sérias exigências de retardo, a frequência de transição também pode depender de conteúdos espectrais de quadros futuros do sinal de áudio. Essa adaptação pode ser realizada no lado do codificador através de um conjunto de circuitos de determinação de transição 60, tipicamente integrado com a seção de quantificação e codificação 28. Contudo, em modalidades alternativas, o conjunto de circuitos de determinação de transição 60 pode ser proporcionado como uma seção que opera separadamente, pelo que apenas um parâmetro representando a frequência de transição é proporcionado para as diferentes funcionalidades do codificador 20. A frequência de transição pode ser usada no lado do codificador, por exemplo, para fornecimento de uma codificação de envelope apropriada para os intervalos de frequência nos diferentes lados da frequência de transição.

[0022] A seção de quantificação e codificação 28 é ainda disposta para empacotamento dos coeficientes espectrais codificados junto com a informação e lado adicional em um fluxo de bits de acordo com o padrão de

transmissão ou armazenamento que será usado. Um fluxo binário 25, tendo dados representando o conjunto de coeficientes espectrais é, assim, transmitido da seção de quantificação e codificação 28. Uma vez que a frequência de transição é derivável diretamente do conteúdo espectral, a mesma derivação pode ser realizada em ambos os lados da interface de transmissão, isto é, no codificador e no decodificador. Isso significa que o valor da própria frequência de transição não que ser, necessariamente, transmitido entre a informação de lado adicional. Contudo, naturalmente, também é possível assim fazer, se houver capacidade de taxa de bits disponível.

[0023] Em uma modalidade particular, uma transformação de MCDT é usada. Após a ponderação realizada por um modelo psicoacústico, os coeficientes de MDCT são quantificados usando quantificação de vetor. Na quantificação de vetor, VQ, os coeficientes espectrais são divididos em pequenos grupos. Cada grupo de coeficiente pode ser visto como um vetor único e cada vetor é quantificado individualmente.

[0024] Por exemplo, devido às altas restrições na taxa de bits, o quantificador pode focalizar os bits disponíveis nos grupos mais energéticos e perceptualmente relevantes, resultando em que alguns grupos são ajustados para zero. Esses grupos formam furos espectrais no espectro quantificado. Isso é ilustrado na figura 3. Na presente modalidade, os grupos 70 compreendem o mesmo número de coeficientes espectrais 71, nesse caso quatro. Contudo, nos grupos de modalidades alternativas, tendo número diferente de coeficientes espectrais, também pode ser possível. Em uma modalidade particular, todos os grupos compreendem apenas um coeficiente espectral, cada um, isto é, o grupo, é o mesmo que o próprio coeficiente espectral. Os grupos quantificados 72 são ilustrados na figura através de retângulos não preenchidos, enquanto os grupos ajustados para zero 73 são ilustrados com retângulos pretos. Tipicamente, são apenas os grupos

codificados 72 que são transmitidos para qualquer usuário final.

[0025] Os grupos 70 de coeficientes são, por sua vez, divididos em diferentes bandas de frequências 74. Essa divisão é realizada, de preferência de acordo com algum critério psicoacústico. Grupos tendo propriedades psicoacústicas essencialmente similares podem, assim, ser tratados coletivamente. O número de elementos de cada banda de frequência 74, isto é, o número de grupos 70 associados com as bandas de frequência 74 podem, portanto, diferir. Se grandes porções de frequência têm propriedades similares, uma banda de frequência cobrindo dessas frequência pode ter uma grande faixa de frequência. Se essas propriedades psicoacústicas mudarem rápido através das frequências, isso, na verdade chama bandas de frequências de uma faixa de frequências menor. As rotinas para o preenchimento de espectro, de preferência, podem depender da banda de frequências a ser preenchida, conforme discutido em mais detalhes abaixo.

[0026] No estágio de decodificação, a operação inversa é obtida, basicamente. Na figura 4, uma modalidade de um decodificador de áudio 40 de acordo com a presente invenção, é ilustrado. Um fluxo binário 25 é recebido, o qual tem propriedades causadas pelo codificador descrito aqui acima. A desquantificação e a decodificação do fluxo binário 25 recebido, por exemplo, um fluxo de bits, são realizadas em um decodificador de coeficiente espectral 41. O decodificador de coeficiente espectral 41 é disposto para decodificação de coeficientes espectrais recuperados do fluxo binário nos coeficientes espectrais decodificados $X^Q[n]$ de um conjunto inicial de coeficientes espectrais 42, agrupados, possivelmente, em grupos de frequências $X^Q_b[n]$. O conjunto inicial de coeficientes espectrais 42, de preferência, parece o conjunto de coeficientes espectrais proporcionados pelo conversor do lado do codificador, possivelmente após o pós-processamento, tal como, por exemplo, limites de mascaramento ou reduções de largura de banda.

[0027] Conforme discutido acima, a aplicação de limites de mascaramento ou reduções de bandas de frequências no codificador, tipicamente, resulta em que o conjunto de coeficientes espectrais 42 é incompleto naquele sentido que compreende, tipicamente, os chamados "furos espectrais". Os "furos espectrais" correspondem a coeficientes espectrais que não são recebidos no fluxo binário. Em outras palavras, os furos espectrais são elemento de sustentação não definidos ou não codificados $X^Q[n]$ ou coeficientes espectrais automaticamente ajustados para um valor predeterminado, tipicamente, zero, pelo decodificador de coeficiente espectral 41. Para evitar artefatos audíveis esses coeficientes têm que ser substituídos por estimativas (preenchidas) no decodificador.

[0028] Os furos espectrais, frequentemente, estão em dois tipos. Furos espectrais pequenos, tipicamente, estão nas baixas frequências e um ou poucos furos espectrais ocorrem nas altas frequências.

[0029] Para minimizar os artefatos no sinal de áudio decodificado, o decodificador "preenche" o espectro por meio da substituição dos furos espectrais no espectro com estimativas dos coeficientes. Essas estimativas podem estar baseadas em informação lateral transmitida pelo decodificador e/ou podem ser dependentes do próprio sinal. Exemplos dessa informação lateral útil poderiam ser o envelope de potência do espectro e da tonalidade, isto é, medição do espectro, dos coeficientes que faltam.

[0030] Dois métodos diferentes podem ser usados para preencher as diferentes espécies de furos espectrais. O preenchimento de ruído funciona bem para furos espectrais nas frequências mais baixas, enquanto a "extensão de largura de banda" é mais adequada em altas frequências. A presente invenção descreve um método para decidir onde o preenchimento de ruído e a extensão de largura de banda devem ser usados, respectivamente.

[0031] A presente invenção conta com a definição de uma frequência de transição entre partes baixas e altas relevantes do espectro. Com base nessa

informação, o algoritmo de codificação típico contando com um procedimento de "preenchimento de ruído" de alta qualidade será capaz de reduzir os artefatos que ocorrem para baixas taxas e também de regenerar um sinal de áudio de largura de banda completa, mesmo em baixas taxas e com um esquema de baixa complexidade, baseado em "extensão de largura de banda". Isso será discutido em mais detalhes abaixo.

[0032] O conjunto inicial de coeficientes espectrais 42 do decodificador de coeficientes espectrais 40, compreendendo, tipicamente, uma certa quantidade de furos espectrais, é proporcionado para um conjunto de circuitos de determinação de transição 60. O conjunto de circuitos de determinação de transição 60 é disposto para determinação de uma frequência de transição f_t .

[0033] O conjunto inicial de coeficientes espectrais 42 do decodificador 41 de coeficientes espectrais também é proporcionado para um preenchedor de espectro 43. O preenchedor de espectro 43 é disposto para preenchimento do espectro do conjunto inicial de coeficientes espectrais 42, dando origem a um conjunto completo 44 de coeficientes espectrais reconstruídos $X'_b[n]$. O conjunto 44 de coeficientes espectrais reconstruídos tem, tipicamente, todos os coeficientes espectrais dentro de uma certa faixa de frequências definida.

[0034] O preenchedor de espectro 43, por sua vez, compreende um preencher de ruídos 50. O preenchedor de ruídos 50 é disposto para proporcionar um processo para preenchimento de ruído de furos espectrais, de preferência, na região de baixa frequência, isto é, abaixo da frequência de transição (f_t). Um valor é assim atribuído aos coeficientes espectrais no conjunto inicial de coeficientes espectrais, abaixo da frequência de transição, que "estão faltando", como um resultado de não serem incluídos no fluxo de bits codificado recebido. Com essa finalidade, uma saída 65 do conjunto de circuitos de determinação de transição 60 é conectada ao preenchedor de

ruídos 50, proporcionando informação associada com a frequência de transição (f_t).

[0035] O preenchedor de espectro 43 também compreende um extensor de largura de banda 55, disposto para estender a largura de banda do conjunto inicial de coeficientes espectrais acima da frequência de transição, a fim de produzir o conjunto 44 de coeficientes espectrais reconstruídos. Portanto, a saída 65 no conjunto de circuitos de determinação de transição 60 também é conectada ao extensor de largura de banda 55.

[0036] Conforme mencionado acima, o resultado do preenchedor de espectro 43 é um conjunto completo 44 de coeficientes espectrais reconstruídos $X_b'[n]$, tendo todos os coeficientes espectrais dentro de uma certa faixa de frequências definida.

[0037] O conjunto 44 de coeficientes espectrais reconstruídos é proporcionado para um conversor 45, conectado ao preenchedor de espectros 43. O conversor 45 é disposto para conversão do conjunto 44 de coeficientes espectrais de um domínio de frequência em um sinal de áudio 46 de um domínio de tempo. O conversor 45, na presente modalidade, é baseado em um transformador perceptual, correspondente à técnica de transformação usada no codificador 20 (figura 2). Em uma modalidade particular, o sinal foi proporcionado de volta no domínio de tempo com uma transformação inversa, por exemplo, MDCT - IMDCT inversa ou DFT - IDFT inversa, etc. Em outras modalidades, um banco de filtros inversos pode ser utilizado. Como no lado do codificador, a técnica do conversor 45 como tal é conhecida na técnica anterior e não será discutida mais uma vez. Um sinal de áudio final perceptualmente reconstruído 34 é proporcionado em uma saída 35 para o sinal de áudio, possivelmente em etapas de tratamento adicionais.

[0038] O codec deve decidir em que bandas de frequência usar o preenchimento de ruído e em que bandas de frequências usar a extensão de largura de banda. O preenchimento de ruído dá o melhor resultado quando a

maior parte dos grupos da banda de frequências a serem preenchidos são quantificados e haver apenas furos espectrais menores na banda. A extensão de largura de banda é preferível, quando uma grande parte do sinal nas altas frequências é deixada não quantificada.

[0039] Um método básico seria estabelecer uma frequência de transição fixa entre o preenchimento de ruído e a extensão de largura de banda. Os furos espectrais nas bandas ou grupos de frequência através daquela frequência são preenchidos pelo extensão de largura de banda.

[0040] Um problema com essa abordagem, porém, é que a frequência de transição ótima não é a mesma para todos os sinais de áudio. Alguns sinais têm a maior parte da energia concentrada nas baixas frequências e uma grande parte do sinal poderia ser submetida à extensão de largura de banda. Outros sinais têm sua energia dispersa mais uniformemente através do espectro e esses sinais podem se beneficiar do uso apenas do preenchedor de ruídos.

[0041] De acordo com uma modalidade de um método de acordo com a presente invenção, a frequência de transição é dependente, adaptativamente, de uma distribuição de furos espectrais no referido conjunto inicial de coeficientes espectrais. Uma rotina para encontrar uma frequência de transição adequada poderia ser seguir através de todas as bandas de frequência, começando na mais alta (BN), descendo até 1. Se não houver coeficientes quantificados na banda corrente, ela será preenchida por extensão de largura de banda. Se houver coeficientes codificados na banda, os furos dessa banda, bem como das bandas seguintes, são preenchidos usando preenchedor de ruído. Desse modo, uma frequência de transição é estabelecida no limite superior da primeira banda de frequência às vistas do lado de alta frequência que tenham coeficiente com a que tem nela um coeficiente quantificado. Isso é ilustrado na figura 5A. Os furos espectrais 77 na banda N, isto é, acima da frequência de transição f_t são, assim, preenchidos com as abordagens de extensão de largura de banda. Os furos espectrais 76

abaixo da frequência de transição f_t são, na verdade, preenchidos através do preenchimento de ruído.

[0042] Uma modalidade alternativa é ilustrada na Fig. 5B. Aqui, a definição da frequência de transição é baseada diretamente nos grupos 70, negligenciando a divisão de banda de frequência. Aqui, a extensão de largura de banda é usada para todos os grupos das frequências maiores abaixo do grupo imediatamente acima do primeiro grupo quantizado 78. Os furos espectrais 76 abaixo da frequência de transição t_r são ao invés disso, cheios por preenchimento de ruído.

[0043] Esses métodos são mais adaptativos ao sinal de áudio e ao quantificador, isto é, o esquema de codificação, mas podem experimentar problemas menores quando o sinal é quantificado, por exemplo, de acordo com a figura 5C. Aqui, uma grande parte das altas frequências do sinal é estabelecida em zero e a extensão de largura de banda, de preferência, será usada da banda B9 até a banda B12. Contudo, uma vez que há um único grupo quantificado, codificado 79 na banda de frequências B11, a extensão de largura de banda será completamente desativada abaixo desse grupo quantificado 79 e o preenchimento de ruído será usado em todas as bandas, subindo até esse grupo 79.

[0044] Para evitar também este problema, outra modalidade também é proposta, onde a frequência de transição f_t é selecionada dependente de uma proporção de furos espectrais nas bandas de frequência. Como nas modalidades anteriores, o codec segue através das bandas de frequência, começando na mais alta descendo até um. Para cada banda de frequência, o número de coeficientes espectrais ou grupos codificados é contado. Se o número de coeficientes ou grupos quantificados dividido pelo número total de coeficientes ou grupos espectrais, isto é, a proporção de coeficientes espectrais codificados da banda de frequências excede um certo limite, os furos espectrais daquela banda de frequência e das bandas de frequência

seguintes são preenchidos com o preenchedor de ruído. Caso contrário a extensão da largura de banda é usada. Analogamente, pode-se monitorar a proporção de furos espectrais nas bandas de frequências. Em outras palavras, uma banda de frequências de transição deve ser encontrada, a qual é uma banda de frequência mais alta em que uma proporção de furos espectrais é menor do que um primeiro limite.

[0045] Há, também, critérios alternativos para selecionar a banda de frequências de transição. Uma possibilidade é deixar o limite depender da frequência. Dessa maneira, uma certa proporção de furos espectrais pode ser aceita nas partes de alta frequência, ainda usando as técnicas de expansão de largura de banda, mas não nas partes de baixa frequência. Qualquer pessoa habilitada na técnica percebe que os detalhes na seleção dos critérios apropriados podem ser variados de muitas maneiras, por exemplo, sendo dependente das propriedades relacionadas com o sinal ou outra informação.

[0046] Em uma modalidade, a frequência de transição é dependente, e de preferência igual, de um limite de frequência superior da banda de frequência de transição. Contudo, há também várias alternativas. Uma alternativa é buscar o coeficiente espectral ou grupo codificado de frequências mais altas e estabelecer a frequência de transição no lado de alta frequência daquele grupo.

[0047] O algoritmo da modalidade descrita acima também pode ser descrito com o pseudocódigo a seguir:

```

For currentBand = N to 1
    ratio = numCodedCoeffInBand(currentBand) /
numCoeffInBand(currentBand)
    If ratio > threshold
        Transition is between currentBand and
currentBand + 1
        Return
    End if
Next
Transition is at the start of band 1

```

[0048] É preferido que a frequência de transição não varie demais entre quadros consecutivos. Mudanças muito grandes podem ser percebidas como perturbação. Portanto, em uma modalidade exemplificativa, a frequência de transição é ainda dependente de uma frequência de transição previamente usada. Por exemplo, será possível proibir a frequência de transição de mudar mais de uma quantidade absoluta ou relativa predeterminada dentro dos quadros consecutivos. Alternativamente, uma frequência de transição provisória poderia ser introduzida como um valor em um filtro junto com frequências de transição anteriores, dando uma frequência de transição modificada, tendo um comportamento de mudança mais amortecido. A frequência de transição, então, dependerá de mais de uma frequência de transição anterior.

[0049] Essas rotinas são realizadas, tipicamente, isto é, de preferência, na seção de quantificação e codificação do codificador, respectivamente.

[0050] A figura 6 é um fluxograma ilustrando etapas de uma modalidade de um método de acordo com a presente invenção. Um método para recuperação de espectros na decodificação espectral de um sinal de áudio começa na etapa 200. Na etapa 210, um conjunto inicial de coeficientes espectrais, representando o sinal de áudio, é obtido. Na etapa 212, uma frequência de transição é determinada. A frequência de transição é adaptada a um conteúdo espectral do sinal de áudio. O preenchimento de um ruído de furos espectrais no conjunto inicial de coeficientes espectrais abaixo da frequência de transição é realizado na etapa 214 e largura de banda estendendo-se do conjunto inicial de coeficientes espectrais acima da frequência de transição é realizada na etapa 216. O processo termina na etapa 249.

[0051] De modo análogo, a figura 7 é um fluxograma ilustrando uma etapa de uma modalidade de outro método de acordo com a presente invenção. Um método para uso na codificação espectral de um sinal de áudio começa na etapa 200. Na etapa 212, uma frequência de transição é

determinada. A frequência de transição para um conjunto inicial de coeficientes espectrais, representando o sinal de áudio, é adaptada a um conteúdo espectral do sinal de áudio. A frequência de transição definindo uma borda entre uma faixa de frequências, destinada a ser um objeto para o preenchimento de ruídos, e uma faixa de frequências destinada a ser um objeto para a extensão de largura de banda.

[0052] A presente invenção adquire o número de vantagens pela definição adaptativa da frequência de transição de acordo com um esquema de codificação usado. A frequência de transição adaptada permite o uso eficiente de um preenchimento de espectro combinado, usando preenchimento de ruído e extensão de largura de banda. Qualquer codec de fala e /ou áudio, usando esse método, é capaz de distribuir um sinal de áudio de largura de banda completa e de alta qualidade com artefatos importunos reduzidos. O método é flexível no sentido de que ele pode ser combinado com qualquer espécie de representação de frequência (DCT, MDCT, etc.) o banco de filtros, isto é, com qualquer codec (perceptual, paramétrico, etc).

[0053] As modalidades descritas acima devem ser compreendidas como uns poucos exemplos ilustrativos da presente invenção. Será compreendido por aqueles habilitados na técnica que várias modificações, combinações e mudanças podem ser feitas nas modalidades sem afastamento do escopo da presente invenção. Em particular, soluções de partes diferentes nas diferentes modalidades podem ser combinadas em outras configurações, onde tecnicamente possível. O escopo da presente invenção, porém, é definido pelas reivindicações anexas.

REFERÊNCIAS

[1] 3GPP TS 26.404 V6.0.0 (2004-09), "Enhanced aacPlus general audio codec - encoder SBR part (Release 6)", 2004

[2] J. D. Johnston, "Estimation of Perceptual Entropy Using Noise Masking Criteria", Proc. ICASSP, pp. 2524-2527, Maio de 1988

REIVINDICAÇÕES

1. Método para recuperação de espectro em decodificação espectral de um sinal de áudio, caracterizado pelo fato de compreender as etapas de:

obtenção (210) de um conjunto inicial (42) de coeficientes espectrais, representando o referido sinal de áudio;

determinação (212) de uma frequência de transição (f_t);

preenchimento de ruído (214) de furos espectrais no referido conjunto inicial (42) de coeficientes espectrais abaixo da referida frequência de transição (f_t); e

extensão de largura de banda (216) do referido conjunto inicial (42) de coeficientes espectrais acima da frequência de transição (f_t);

a frequência de transição (f_t) sendo adaptado a um conteúdo espectral do referido sinal de áudio;

a frequência de transição (f_t) sendo adaptativamente dependente de uma distribuição de furos espectrais no referido conjunto inicial (42) de coeficientes espectrais.

2. Método de acordo com a reivindicação 1, caracterizado pelo fato de a referida etapa de determinação da referida frequência de transição (f_t), por sua vez, compreender as etapas de:

divisão dos coeficientes espectrais do conjunto inicial (42) de coeficientes espectrais em uma pluralidade de bandas de frequência (74); e

seleção da frequência de transição (f_t) dependente de uma proporção de furos espectrais nas bandas de frequência (74).

3. Método de acordo com a reivindicação 2, caracterizado pelo fato de as referidas bandas de frequência (74) terem uma largura de frequência constante.

4. Método de acordo com a reivindicação 2, caracterizado pelo fato de pelo menos duas das referidas bandas de frequência (74) terem

larguras de frequência (f_t) diferentes.

5. Método de acordo com qualquer uma das reivindicações de 2 a 4, caracterizado pelo fato de a etapa de seleção da frequência de transição (f_t) compreender:

encontrar uma banda de frequências de transição, sendo uma banda de frequência mais alta em que a proporção é menor do que um primeiro limite.

6. Método de acordo com a reivindicação 5, caracterizado pelo fato de a etapa de seleção da referida frequência de transição (f_t) compreender adicionalmente:

ajuste da referida frequência de transição (f_t) dependente de um limite de frequência superior da banda de frequência de transição.

7. Método de acordo com a reivindicação 5 ou 6, caracterizado pelo fato de a referida etapa de ajuste da frequência de transição (f_t) ainda ser dependente de uma frequência de transição usada previamente.

8. Método de acordo com a reivindicação 7, caracterizado pelo fato de a etapa de ajuste da referida frequência de transição (f_t) ainda ser dependente de mais de uma frequência de transição previamente usada.

9. Método de acordo com a reivindicação 7 ou 8, caracterizado pelo fato de a referida frequência de transição (f_t) ser proibida de mudar mais de uma quantidade absoluta ou relativa predeterminada entre dois quadros consecutivos.

10. Método para uso em codificação espectral de um sinal de áudio, caracterizado pelo fato de compreender:

determinação (212) de uma frequência de transição (f_t) para um conjunto inicial (24; 42) de coeficientes espectrais representando o referido sinal de áudio;

a referida frequência de transição (f_t) definindo uma borda entre uma faixa de frequência, destinada a ser um objeto para preenchimento

de ruído de furos espectrais e uma faixa de frequências destinada a ser um objeto para extensão de largura de banda;

a frequência de transição (f_t) sendo adaptada a um conteúdo espectral do referido sinal de áudio;

a referida frequência de transição (f_t) sendo adaptativamente dependente de uma distribuição de furos espectrais no referido conjunto inicial (42) de coeficientes espectrais.

11. Decodificador (40) para decodificação espectral de um sinal de áudio, caracterizado pelo fato de compreender:

entrada para obtenção de um conjunto inicial (42) de coeficientes espectrais representando o referido sinal de áudio;

conjunto de circuitos de determinação de transição (60) disposto para determinação de uma frequência de transição (f_t);

um preenchedor de ruído (50) para preenchimento de ruído de furos espectrais no referido conjunto inicial (42) de coeficientes espectrais abaixo da referida frequência de transição (f_t); e

um extensor de largura de banda (55) disposto para extensão de largura de banda do referido conjunto inicial (42) de coeficientes espectrais acima da referida frequência de transição (f_t);

a frequência de transição (f_t) sendo adaptada a um conteúdo espectral do referido sinal de áudio;

o conjunto de circuitos de determinação de transição (60) sendo disposto para determinar, adaptativamente, a frequência de transição (f_t) dependente de uma distribuição de furos espectrais no conjunto inicial (42) de coeficientes espectrais.

12. Decodificador de acordo com a reivindicação 11, caracterizado pelo fato de o conjunto de circuitos de determinação de transição (60) ser disposto ainda para dividir os coeficientes espectrais do conjunto inicial de coeficientes espectrais em uma pluralidade de bandas de

frequências (74) e para selecionar a frequência de transição (f_t) dependente de uma proporção de furos espectrais nas bandas de frequências (74).

13. Decodificador de acordo com a reivindicação 12, caracterizado pelo fato de as referidas bandas de frequências (74) terem uma largura de frequência constante.

14. Decodificador de acordo com a reivindicação 12, caracterizado pelo fato de pelo menos duas das bandas de frequências (74) terem larguras de frequências diferentes.

15. Decodificador de acordo com qualquer uma das reivindicações de 12 a 14, caracterizado pelo fato de o conjunto de circuitos de determinação de transição (60) ser ainda disposto para

encontrar uma banda de frequências de transição, sendo uma banda de frequências mais altas em que a proporção é menor do que um primeiro limite.

16. Decodificador de acordo com a reivindicação 15, caracterizado pelo fato de o conjunto de circuitos de determinação de transição (6) ser ainda disposto para

estabelecer a frequência de transição (f_t) dependente de um limite de frequência superior da banda de frequências de transição.

17. Codificador (20) para codificação espectral de um sinal de áudio, caracterizado pelo fato de compreender:

conjunto de circuitos de determinação de transição (60) disposto para determinação de uma frequência de transição (f_t) para um conjunto inicial (24) de coeficientes espectrais, representando o sinal de áudio;

a frequência de transição (f_t) definindo uma borda entre uma faixa de frequências, destinada a ser um objeto para preenchimento de ruído de furos espectrais e uma faixa de frequências, destinada a ser um objeto para extensão de largura de banda;

a frequência de transição (f_t) sendo adaptada a um conteúdo

espectral do sinal de áudio;

o conjunto de circuitos de determinação de transição (60) sendo disposto para determinar, adaptativamente, a frequência de transição (f_t) dependente de uma distribuição de furos espectrais no conjunto inicial (42) de coeficientes espectrais.

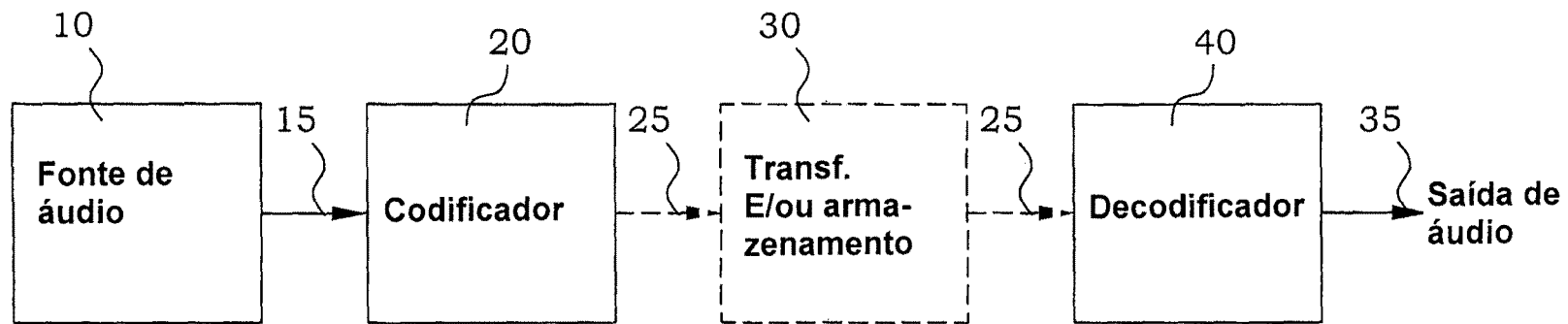


Fig. 1

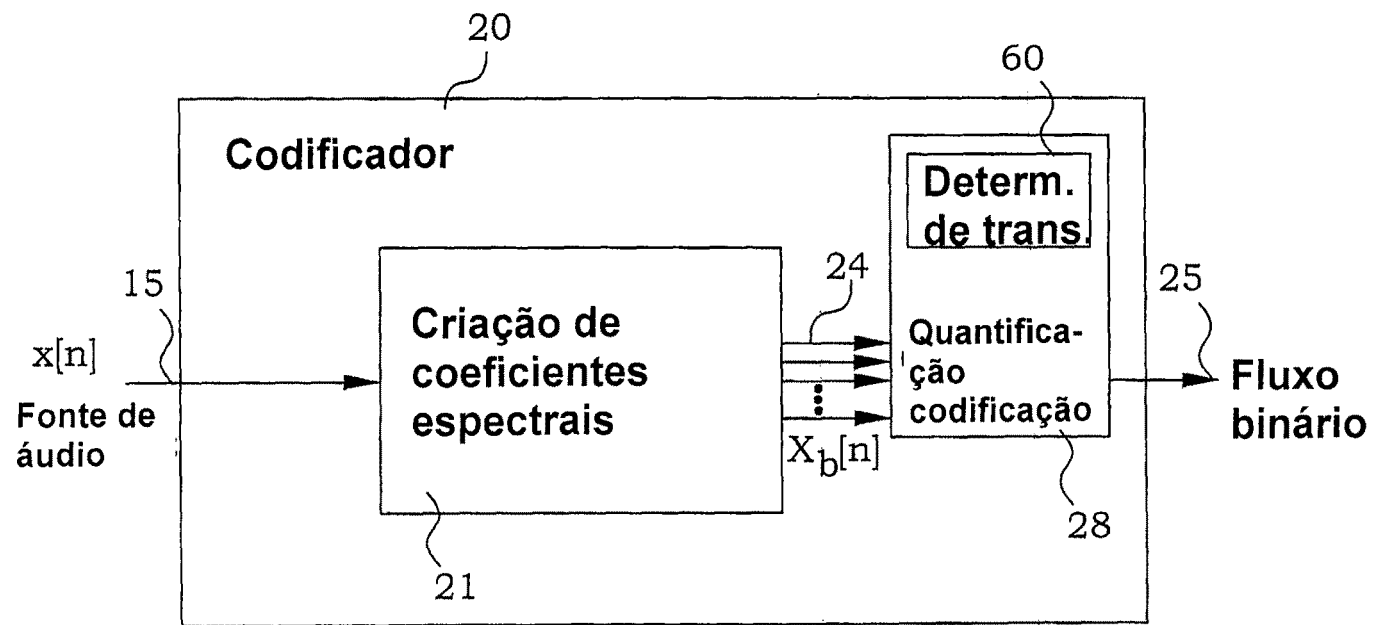


Fig. 2

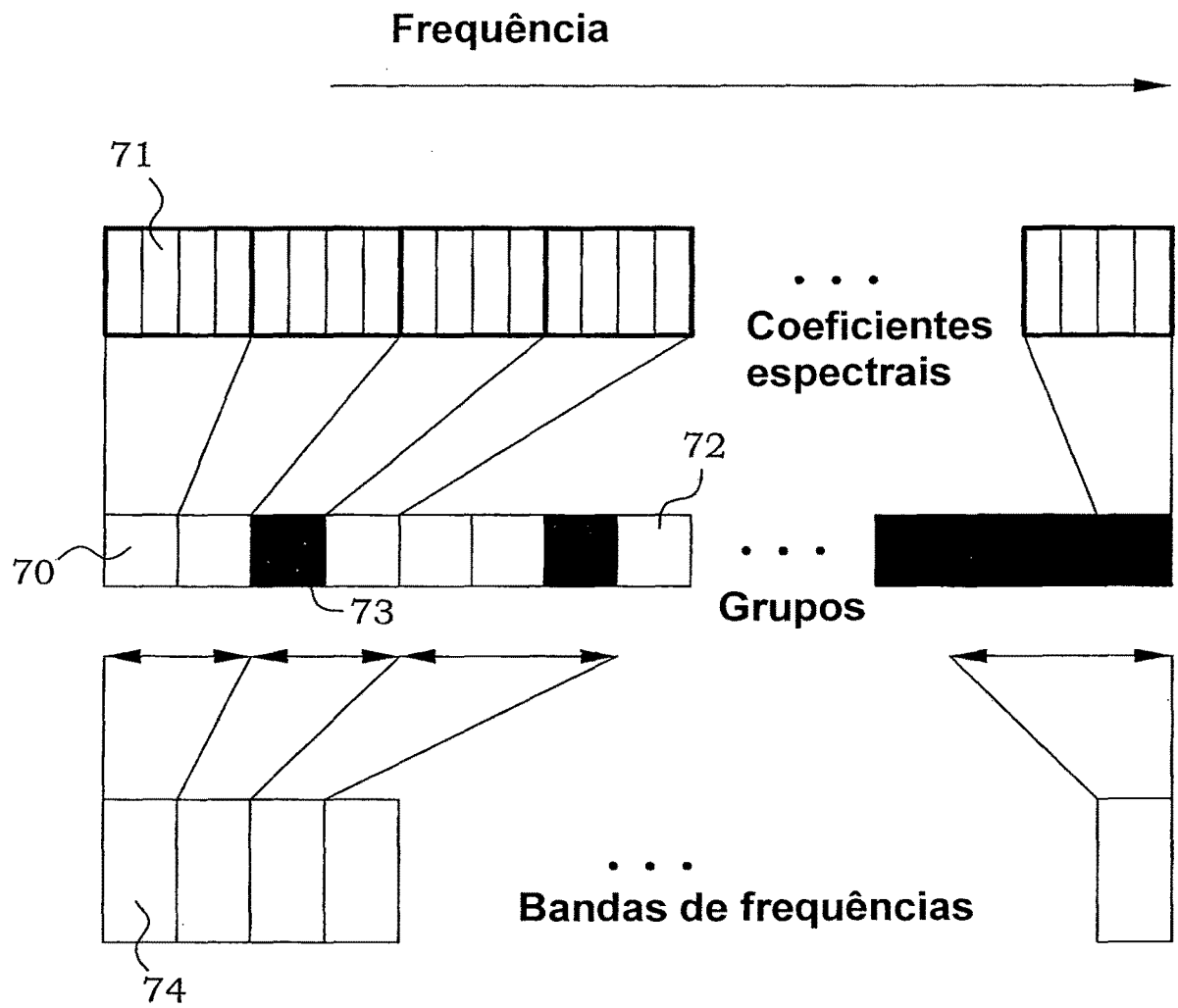


Fig. 3

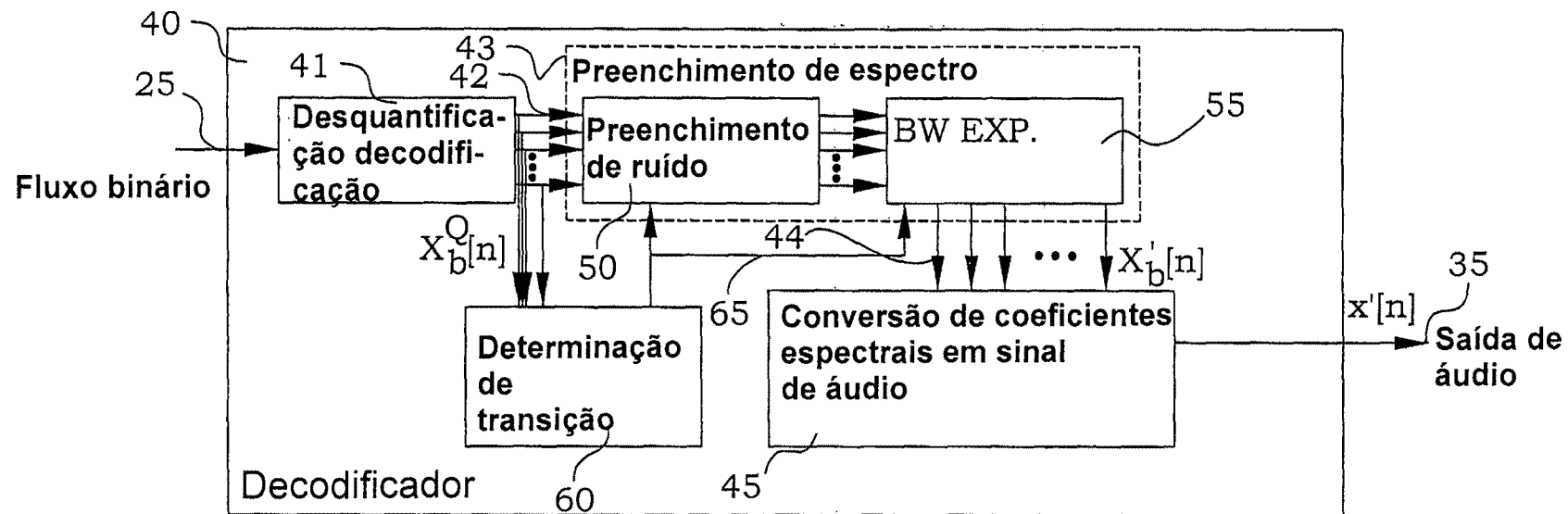


Fig. 4

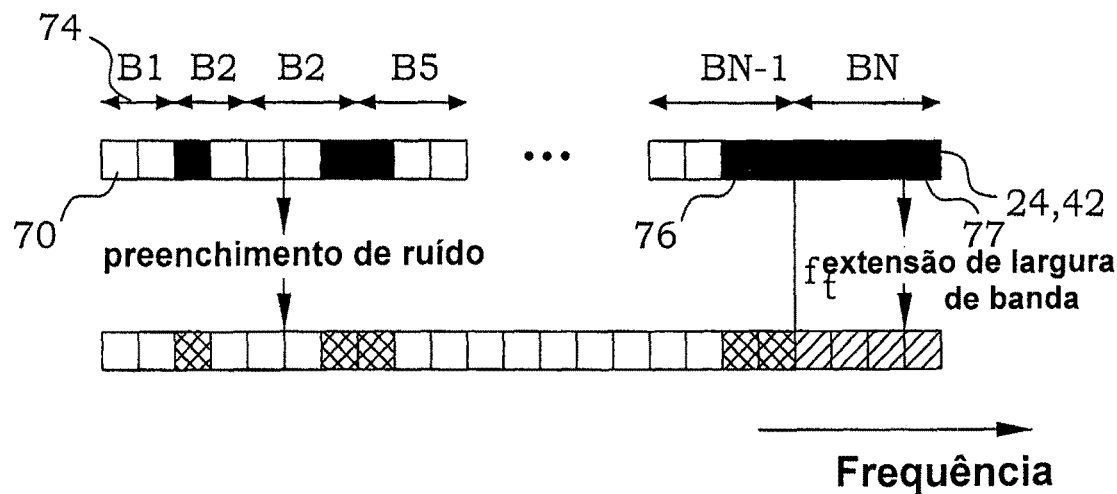


Fig. 5A

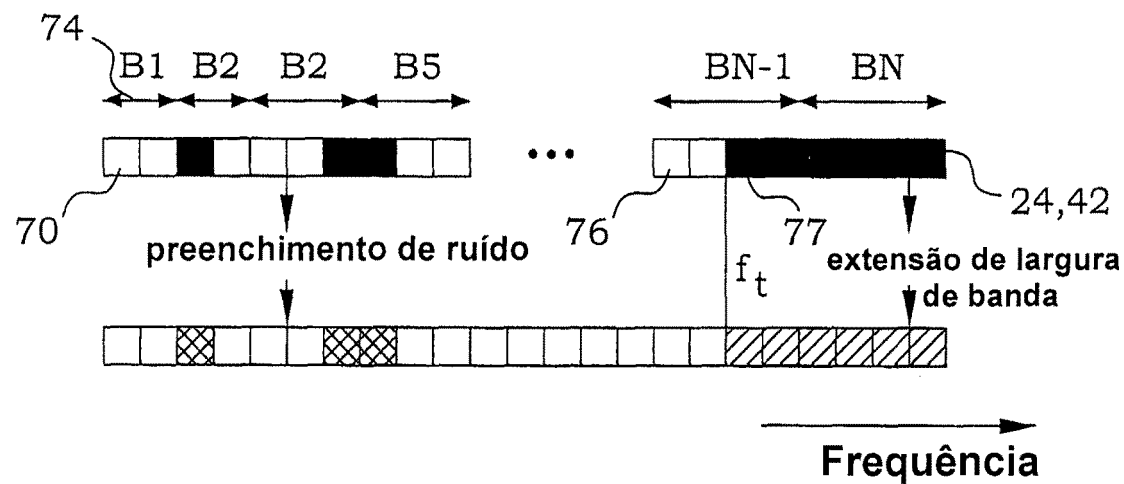


Fig. 5B

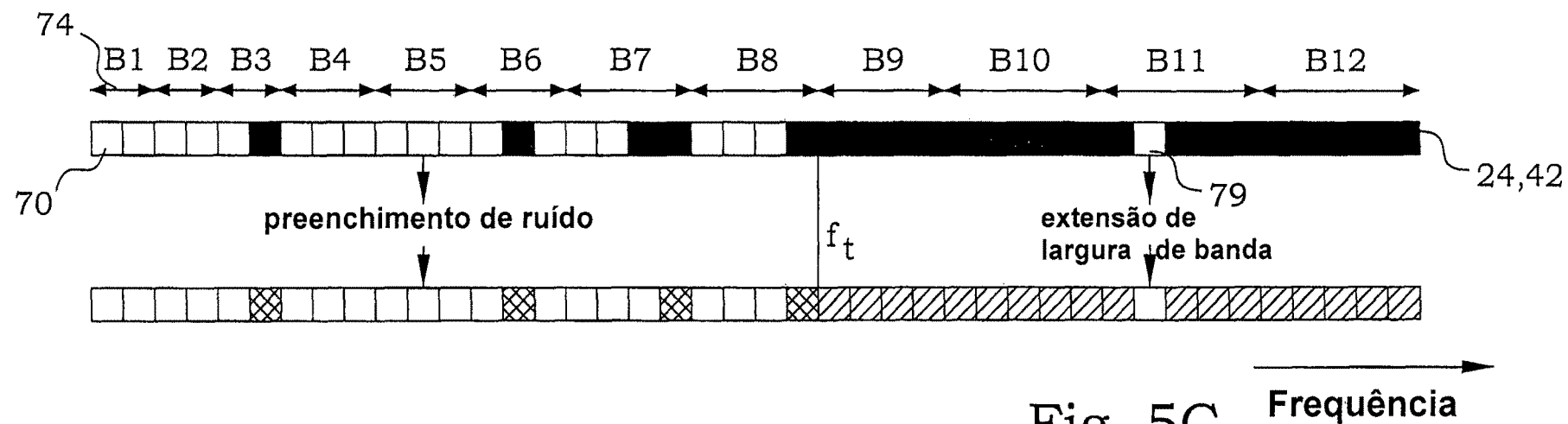


Fig. 5C

