



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

**(11) PI 0910511-5 B1**



**(22) Data do Depósito: 23/06/2009**

**(45) Data de Concessão: 01/06/2021**

**(54) Título:** APARELHO E MÉTODO PARA DECODIFICAR E CODIFICAR UM SINAL DE ÁUDIO

**(51) Int.Cl.:** G10L 19/20; G10L 21/02; G10L 19/008.

**(30) Prioridade Unionista:** 08/10/2008 US 61/103,820; 11/07/2008 US 61/079,841.

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

**(72) Inventor(es):** MAX NEUENDORF; BERNHARD GRILL; ULRICH KRAEMER; MARKUS MULTRUS; HARALD POPP; NIKOLAUS RETTELBACH; FREDERIK NAGEL; MARKUS LOHWASSER; MARC GAYER; MANUEL JANDER; VIRGILIO BACIGALUPO.

**(86) Pedido PCT:** PCT EP2009004522 de 23/06/2009

**(87) Publicação PCT:** WO 2010/003545 de 14/01/2010

**(85) Data do Início da Fase Nacional:** 07/01/2011

**(57) Resumo:** APARELHO E MÉTODO PARA DECODIFICAÇÃO DE UM SINAL DE ÁUDIO CODIFICADO. Um aparelho para decodificar (100) um sinal de 5 áudio codificado (102) . Um primeiro decodificador (110a) decodifica uma primeira porção (104a), de acordo com um primeiro algoritmo de decodificação, para uma primeira porção de tempo do sinal codificado (102) para obter um primeiro sinal decodificado (114a) . Um segundo decodificador (110b) decodifica uma segunda porção (104b) , de acordo com um segundo algoritmo de decodificação, para uma segunda porção de tempo do sinal codificado (102) para obter um segundo sinal decodificado (114b). Um módulo BWE (130) tem uma frequência cruzada (fx) controlável e está configurado para efetuar um algoritmo de extensão de largura de banda usando o primeiro sinal decodificado (114a) e parâmetros de BWE (106) para a primeira porção (104a), bem como para efetuar um algoritmo de extensão de largura de banda usando o segundo sinal decodificado (114b) e os parâmetros de extensão de largura de banda (106) para a segunda porção (104b). Um controlador (140) controla a frequência cruzada (fx) para o módulo BWE (130), de acordo com uma informação de modo de codificação (108).

**"APARELHO E MÉTODO PARA DECODIFICAR E CODIFICAR  
UM SINAL DE ÁUDIO"**

**ESPECIFICAÇÃO**

[0001] A presente invenção se refere a um aparelho e a um método para decodificar um sinal de áudio codificado, um aparelho para codificar, um método para codificação e um sinal de áudio.

[0002] Na técnica, os esquemas de codificação de domínio de frequência, tais como MP3 ou AAC, são conhecidos. Os referidos codificadores de domínio de frequência são baseados em uma conversão de domínio de tempo/domínio de frequência, uma etapa subsequente de quantização, na qual o erro de quantização é controlado utilizando-se a informação proveniente do módulo psicoacústico, bem como uma etapa de codificação, na qual os coeficientes espectrais quantizados e a informação colateral respectiva são codificados por entropia utilizando-se tabelas de código.

[0003] Por outro lado, há codificadores que são bastante adequados para o processamento de fala, tais como AMR-WB+, conforme descrito em 3GPP TS 26.290. Os referidos esquemas de codificação de fala realizam uma filtragem Preditiva Linear de um sinal de domínio de tempo. A referida filtragem PL é derivada de uma análise de Predição Linear do sinal de entrada de domínio de tempo. Em seguida, os coeficientes do filtro de LP resultantes são codificados e transmitidos como informação colateral. O processo é conhecido como Codificação de Predição Linear (LPC). Na saída do filtro, o sinal residual de predição ou sinal de erro de predição,

que é conhecido também como sinal de excitação, é codificado usando-se as etapas de *analysis-by-synthesis* [análise pela síntese] do codificador ACELP ou, de modo alternativo, é codificado usando-se um codificador de transformadas que utiliza uma transformada de Fourier com uma sobreposição. A escolha entre a codificação ACELP e a codificação de excitação de Transformada Codificada, que também é denominada codificação TCX, é realizada usando-se um algoritmo de circuito fechado ou de circuito aberto.

[0004] Os esquemas de codificação de áudio de domínio de frequência, tais como o esquema de codificação AAC de alta eficiência, que combina um esquema de codificação AAC e uma técnica de replicação de largura de banda espectral, também podem ser combinados a um *joint stereo* ou a uma ferramenta de codificação de múltiplos canais, que é conhecida pelo termo "MPEG surround". Por outro lado, os codificadores de fala, tais como o AMR-WB+, também têm uma etapa de aprimoramento de alta frequência e uma funcionalidade estéreo.

[0005] A referida replicação da banda espectral (SBR) compreende uma técnica que ganhou popularidade como um *add-on* para a percepção popular de áudio codificado, tais como MP3, e a codificação avançada de áudio (AAC). A SBR compreende um método de extensão de largura de banda (BWE) no qual a banda baixa (banda de base ou banda central) do espectro é codificada usando-se uma codificação existente, considerando que a banda superior (ou banda alta) é grosseiramente parametrizada usando-se alguns parâmetros. A SBR faz uso de uma correlação entre a banda baixa e a banda alta, para prever o sinal de banda alta proveniente da extração

das características da banda mais baixa.

[0006] A SBR é, por exemplo, usada em HE-AAC ou AAC+SBR. Na SBR, é possível alterar dinamicamente a frequência cruzada (frequência inicial de BWE), bem como a resolução temporal que significa o número de conjuntos de parâmetros (envelopes) por frame. O AMR-WB+ implementa uma extensão de largura de banda de domínio de tempo em combinação com um codificador central de domínio de tempo/frequência comutado, fornecendo boa qualidade de áudio, especialmente para sinais de fala. Um fator limitante para a qualidade de áudio do AMR-WB+ é a largura de banda de áudio comum para os dois codificadores centrais e a frequência inicial de BWE que é um quarto da frequência de amostragem interna do sistema. Enquanto o modelo de fala ACELP é capaz de modelar os sinais de fala tão bem sobre a largura de banda total, o codificador de áudio de domínio de frequência falha ao fornecer uma qualidade adequada para alguns sinais gerais de áudio. Assim, os esquemas de codificação de fala mostram uma alta qualidade para sinais de fala, inclusive em baixas taxas de bit, porém mostram uma baixa qualidade para sinais de música em taxas de bit baixas.

[0007] Os esquemas de codificação de domínio de frequência, tais como HE-AAC, são vantajosos por mostrar uma alta qualidade em baixas taxas de bits para sinais de música. Entretanto, a qualidade de sinais de fala é problemática em taxas de bits baixas.

[0008] Portanto, diferentes categorias de sinal de áudio demandam diferentes características de ferramenta de extensão de largura de banda.

[0009] Este é o objeto da presente invenção para fornecer um conceito de codificação/decodificação aprimorado.

[00010] O referido objeto é atingido por meio de um decodificador de áudio, de acordo com a reivindicação 1, um método de decodificação de áudio, em conformidade com a reivindicação 13, um codificador, de acordo com a reivindicação 8, um método para a codificação, em conformidade com a reivindicação 14, um sinal codificado, em conformidade com a reivindicação 15 ou um programa de computador, em conformidade com a reivindicação 16.

[00011] A presente invenção está baseada no achado de que a frequência cruzada ou a frequência inicial de BWE é um parâmetro que influencia a qualidade de áudio. Enquanto os codificadores de domínio de tempo (fala) normalmente codificam a banda de frequência inteira para uma taxa de amostragem dada, a largura de banda de áudio é um parâmetro de afinação para codificadores baseados em transformadas (por exemplo, codificadores para música), uma vez que diminuem o número total de linhas espectrais para codificar ao mesmo tempo em que aumentarão o número de bits por linha espectral disponível para codificação, que significa que é feito um *trade-off* entre a qualidade em relação à largura de banda de áudio. Portanto, na nova abordagem, diferentes codificadores centrais, com larguras de banda de áudio variáveis, são combinados a um sistema comutado com um módulo comum de BWE, no qual o módulo BWE deve contar para diferentes larguras de banda de áudio.

[00012] Um modo direto poderia estar para localizar a mais baixa de todas as larguras de banda do codificador central e

utilizá-la como frequência inicial de BWE, porém isso poderia deteriorar a qualidade de áudio percebida. Ademais, a eficiência de codificação poderia ser reduzida, uma vez que as seções de tempo onde um codificador central está ativo, o qual tem uma largura de banda superior à frequência inicial de BWE, algumas regiões de frequência poderiam ser duas vezes representadas, por meio do codificador central, bem como a BWE que introduz redundância. Portanto, uma melhor solução é adaptar a frequência inicial de BWE para a largura de banda de áudio do codificador central utilizado.

[00013] Portanto, de acordo com as configurações da presente invenção, um sistema de codificação de áudio combina uma ferramenta de extensão de largura de banda com um codificador central dependente do sinal (por exemplo, codificador comutado de fala/áudio), no qual a frequência cruzada compreende um parâmetro variável. Uma saída classificadora de sinal, que controla a comutação entre diferentes modos de codificação central, também pode ser utilizada para comutar as características do sistema de BWE, tais como resolução temporal e *smearing*, a resolução espectral e a frequência cruzada.

[00014] Portanto, um aspecto da presente invenção é um decodificador de áudio para um sinal de áudio codificado, o sinal de áudio codificado que compreende uma primeira porção codificada em conformidade com o primeiro algoritmo de codificação, uma segunda porção codificada em conformidade com um segundo algoritmo de codificação, parâmetros de BWE para a primeira porção e para a segunda porção e uma informação de modo de codificação que indica

um primeiro algoritmo de decodificação ou um segundo algoritmo de decodificação, que compreende um primeiro decodificador, um segundo decodificador, um módulo BWE e um controlador. O primeiro decodificador decodifica a primeira porção em conformidade com o primeiro algoritmo de decodificação para uma primeira porção de tempo do sinal codificado para obter um primeiro sinal decodificado. O segundo decodificador decodifica a segunda porção em conformidade com o segundo algoritmo de decodificação para uma segunda porção de tempo do sinal codificado para obter um segundo sinal decodificado. O módulo BWE tem uma frequência cruzada controlável e está configurado para efetuar um algoritmo de extensão de largura de banda usando o primeiro sinal decodificado e os parâmetros de BWE para a primeira porção, e para efetuar um algoritmo de extensão de largura de banda usando o segundo sinal decodificado e o parâmetro de extensão de largura de banda para a segunda porção. O controlador controla a frequência cruzada para o módulo BWE, em conformidade com a informação de modo de codificação.

[00015] De acordo com outro aspecto da presente invenção, um aparelho para codificação de um sinal de áudio compreende um primeiro e um segundo codificador, uma etapa de decisão e um módulo BWE. O primeiro codificador está configurado para codificar, em conformidade com um primeiro algoritmo de codificação, o primeiro algoritmo de codificação que tem uma primeira largura de banda de frequência. O segundo codificador está configurado para codificar, em conformidade com um segundo algoritmo de codificação, o segundo algoritmo de codificação que

tem uma segunda largura de banda de frequência que é menor que a primeira largura de banda de frequência. A etapa de decisão indica o primeiro algoritmo de codificação para uma primeira porção do sinal de áudio e o segundo algoritmo de codificação para uma segunda porção do sinal de áudio, a segunda porção sendo diferente da primeira porção. O módulo de extensão de largura de banda calcula os parâmetros de BWE para o sinal de áudio, no qual o módulo BWE é configurado para ser controlado por meio da etapa de decisão para calcular os parâmetros de BWE para uma banda que não inclui a primeira largura de banda de frequência na primeira porção do sinal de áudio, e para uma banda que não inclui a segunda largura de banda de frequência na segunda porção do sinal de áudio.

[00016] Em oposição às configurações, a SBR, na técnica anterior, somente é aplicada para um codificador de áudio sem comutação, que resulta nas desvantagens a seguir. Tanto a resolução temporal quando a frequência cruzada poderia ser dinamicamente aplicada, porém as implementações do estado da técnica, tais como a fonte 3GPP, aplicam, normalmente, somente uma alteração de resolução temporária para transientes como, por exemplo, castanholas. Ademais, uma resolução temporal geral melhor poderia ser escolhida em taxas superiores, como um parâmetro de afinação dependente da taxa de bits. Não é realizada nenhuma classificação explícita para determinar a resolução temporal ou um limite de decisão que controla a resolução temporal, combinando melhor o tipo de sinal, como, por exemplo, música estacionária, tonal em relação à fala. As configurações da presente invenção

superam as referidas desvantagens. As configurações permitem, especialmente, uma frequência cruzada adaptada combinada com uma escolha flexível para o codificador central utilizado, assim o sinal codificado fornece uma qualidade perceptível significativamente superior comparada ao codificador/decodificador da técnica anterior.

#### **BREVE DESCRIÇÃO DOS DESENHOS**

[00017] As configurações preferidas da presente invenção são descritas a seguir em relação aos desenhos anexados, nos quais:

[00018] A Fig. 1 mostra um diagrama de bloco de um aparelho para decodificação, em conformidade com um primeiro aspecto da presente invenção;

[00019] A Fig. 2 mostra um diagrama de bloco de um aparelho para codificação, em conformidade com o primeiro aspecto da presente invenção;

[00020] A Fig. 3 mostra um diagrama de bloco de um esquema de codificação em mais detalhes;

[00021] A Fig. 4 mostra um diagrama de bloco de um esquema de decodificação em mais detalhes;

[00022] A Fig. 5 mostra um diagrama de bloco de um esquema de codificação, em conformidade com um segundo aspecto;

[00023] A Fig. 6 é um diagrama esquemático de um esquema de decodificação, em conformidade com um segundo aspecto;

[00024] A Fig. 7 ilustra uma etapa de LPC do lado do codificador que fornece informação preditiva em curto prazo e o sinal de erro de predição;

[00025] A Fig. 8 ilustra uma configuração adicional de um dispositivo de LPC para gerar um sinal ponderado;

[00026] As Figs. 9a e 9b mostram um codificador que compreende um switch de áudio/fala que resulta em resolução temporal diferente para um sinal de áudio; e

[00027] A Fig. 10 ilustra uma representação para um sinal de áudio codificado.

#### **DESCRIÇÃO DETALHADA DA INVENÇÃO**

[00028] A Fig. 1 mostra um aparelho decodificador 100 para decodificar um sinal de áudio codificado 102. O sinal de áudio codificado 102 compreende uma primeira porção 104a codificada em conformidade com o primeiro algoritmo de codificação, uma segunda porção 104b codificada em conformidade com um segundo algoritmo de codificação, parâmetro de BWE 106 para a primeira porção de tempo 104a e para a segunda porção de tempo 104b e uma informação de modo de codificação 108 que indica um primeiro algoritmo de decodificação ou um segundo algoritmo de decodificação para as respectivas porções de tempo. O aparelho para decodificação 100 compreende um primeiro decodificador 110a, um segundo decodificador 110b, um módulo BWE 130 e um controlador 140. O primeiro decodificador 110a é adaptado para decodificar a primeira porção 104a, de acordo com o primeiro algoritmo de decodificação para uma primeira porção de tempo do sinal codificado 102 para obter um primeiro sinal decodificado 114a. O segundo decodificador 110b está configurado para decodificar a segunda porção 104b, de acordo com o segundo algoritmo de decodificação, para uma segunda porção de tempo do sinal

codificado, para obter um segundo sinal decodificado 114b. O módulo BWE 130 tem uma frequência cruzada controlável  $f_x$  que ajusta o comportamento do módulo BWE 130. O módulo BWE 130 está configurado para efetuar um algoritmo de extensão de largura de banda para gerar componentes do sinal de áudio na banda de frequência superior baseada no primeiro sinal decodificado 114a e nos parâmetros de BWE 106 para a primeira porção, e para gerar componentes do sinal de áudio na banda de frequência superior baseada no segundo sinal decodificado 114b e no parâmetro de extensão de largura de banda 106 para a segunda porção. O controlador 140 está configurado para controlar a frequência cruzada  $f_x$  do módulo BWE 130, de acordo com a informação do modo de codificação 108.

[00029] O módulo BWE 130 pode compreender também um combinador que combina componentes de sinal de áudio da banda de frequência inferior e superior e fornece o sinal de áudio resultante 105.

[00030] A informação do modo de codificação 108 indica, por exemplo, qual porção de tempo do sinal de áudio codificado 102 é codificada por qual algoritmo de codificação. Esta informação pode, ao mesmo tempo, identificar o decodificador a ser utilizado para diferentes porções de tempo. Ademais, a informação do modo de codificação 108 pode controlar um *switch to switch* entre diferentes decodificadores para diferentes porções de tempo.

[00031] Portanto, a frequência cruzada  $f_x$  é um parâmetro ajustável, que é ajustado em conformidade com o decodificador utilizado, o qual pode, por exemplo, compreender um codificador de

fala como o primeiro decodificador 110a e um decodificador de áudio como o segundo decodificador 110b. Conforme acima mencionado, a frequência cruzada  $f_x$  para um decodificador de fala (como para o exemplo baseado em LPC) pode ser superior à frequência cruzada utilizada para um decodificador de áudio (por exemplo, para música). Assim, nas configurações adicionais, o controlador 220 está configurado para aumentar a frequência cruzada  $f_x$  ou para diminuir a frequência cruzada  $f_x$  dentro de uma das porções de tempo (por exemplo, a segunda porção de tempo), assim, a frequência cruzada pode ser alterada sem alteração do algoritmo de decodificação. Isso significa que uma alteração na frequência cruzada pode não estar relacionada a uma alteração no decodificador utilizado: a frequência cruzada pode ser alterada sem alteração do decodificador utilizado, ou vice-versa, o decodificador pode ser alterado sem alteração da frequência cruzada.

[00032] O módulo BWE 130 pode também compreender um switch que é controlado pelo controlador 140 e/ou pelo parâmetro de BWE 106, assim o primeiro sinal decodificado 114a é processado por meio do módulo BWE 130, durante a primeira porção de tempo, e o segundo sinal decodificado 114b é processado pelo módulo BWE 130 durante, a segunda porção de tempo. Este switch pode ser ativado por meio da alteração na frequência cruzada  $f_x$  ou por um bit explícito dentro do sinal de áudio codificado 102, indicando o algoritmo de codificação utilizado durante a respectiva porção de tempo.

[00033] Nas demais configurações, o switch está

configurado para comutar entre a primeira e a segunda porção de tempo, a partir do primeiro decodificador para o segundo decodificador, assim o algoritmo de extensão de largura de banda é aplicado tanto ao primeiro sinal decodificado quanto ao segundo sinal decodificado. De modo alternativo, o algoritmo de extensão de largura de banda é aplicado ao primeiro e/ou ao segundo sinal decodificado e o switch é colocado após este, assim um dos sinais de largura de banda estendida é transmitido.

[00034] A Fig. 2 mostra um diagrama de bloco para um aparelho 200 para codificação de um sinal de áudio 105. O aparelho para codificação 200 compreende um primeiro codificador 210a, um segundo codificador 210b, uma etapa de decisão 220 e um módulo de extensão de largura de banda (módulo BWE) 230. O primeiro codificador 210a está operante para codificar, de acordo com um primeiro algoritmo de codificação, que tem uma primeira largura de banda de frequência. O segundo codificador 210b está operante para codificar, de acordo com um segundo algoritmo de codificação, que tem uma segunda largura de banda de frequência menor que a primeira largura de banda de frequência. O primeiro codificador pode, por exemplo, ser um codificador de fala, tais como um codificador baseado em LPC, considerando que o segundo codificador 210b pode compreender um codificador de áudio (música). A etapa de decisão 220 está configurada para indicar o primeiro algoritmo de codificação para uma primeira porção 204a do sinal de áudio 105, e para indicar o segundo algoritmo de codificação para uma segunda porção 204b do sinal de áudio 105, onde a segunda porção de tempo é diferente da primeira porção de tempo. A primeira porção de

tempo 204a pode referir-se à primeira porção de tempo e a segunda porção 204b pode referir-se à segunda porção de tempo, que é diferente da primeira porção de tempo.

[00035] O módulo BWE 230 está configurado para calcular os parâmetros de BWE 106 para o sinal de áudio 105 e está configurado para ser controlado pela etapa de decisão 220 para calcular o parâmetro de BWE 106 para uma primeira banda que não inclui a primeira largura de banda de frequência na primeira porção de tempo 204a do sinal de áudio 105. O módulo de BWE 230 está adicionalmente configurado para calcular o parâmetro de BWE 106 para uma segunda banda, não incluindo a segunda largura de banda na segunda porção de tempo 204b do sinal de áudio 105. A primeira (segunda) banda compreende, portanto, componentes de frequência do sinal de áudio 105 que estão fora da primeira (segunda) largura de banda de frequência e estão limitados em direção à extremidade inferior do espectro por meio da frequência cruzada  $f_x$ . A primeira ou a segunda largura de banda pode, portanto, ser definida por meio de uma frequência cruzada variável que é controlada pela etapa de decisão 220.

[00036] Ademais, o módulo BWE 230 pode compreender um switch controlado pela etapa de decisão 220. A etapa de decisão 220 pode determinar um algoritmo de codificação preferido para uma porção de tempo dada e controla o switch, assim, durante uma dada porção de tempo, o codificador preferido é utilizado. A informação modificada do modo de codificação 108' compreende o sinal do switch respectivo. Ademais, o módulo BWE 230 pode também compreender um filtro para obter componentes do sinal de áudio 105

na banda de frequência inferior/superior, que estão separados pela frequência cruzada  $f_x$ , que pode compreender um valor de cerca de 4 kHz ou 5 kHz. Finalmente, o módulo BWE 130 também pode compreender uma ferramenta de análise para determinar o parâmetro de BWE 106. A informação modificada de modo de codificação 108' pode ser equivalente (ou igual) à informação do modo de codificação 108. A informação do modo de codificação 108 indica, por exemplo, o algoritmo de codificação utilizado para as respectivas porções de tempo no fluxo de bits do sinal de áudio codificado 105.

[00037] De acordo com configurações adicionais, a etapa de decisão 220 compreende uma ferramenta classificadora de sinal que analisa o sinal original de saída 105 e gera a informação de controle 108 que aciona a seleção dos diferentes modos de codificação. A análise do sinal de entrada 105 depende da implementação com o objetivo de escolher o modo de codificação central ideal para um dado frame de sinal de entrada. A saída do classificador de sinal pode (opcionalmente) também ser utilizada para influenciar o comportamento de outras ferramentas, por exemplo, MPEG surround, SBR aprimorada, banco de filtros *time-warped* e outros. A entrada para a ferramenta classificadora de sinal compreende, por exemplo, o sinal de entrada original não modificado 105, mas também, opcionalmente, parâmetros dependentes da implementação. A saída da ferramenta classificadora de sinal compreende o sinal de controle 108 para controlar a seleção do codificador central (por exemplo, domínio de frequência não filtrado PL, tempo filtrado por PL, codificação de domínio de frequência ou algoritmos adicionais de codificação).

[00038] De acordo com as configurações, a frequência cruzada  $f_x$  é dependente do sinal ajustado, que está combinado com a decisão de switch para o uso de um algoritmo diferente de codificação. Portanto, um sinal simples de switch pode simplesmente ser uma alteração (um salto) na frequência cruzada  $f_x$ . Ademais, a informação do modo de codificação 108 também pode compreender a alteração da frequência cruzada  $f_x$  indicando, ao mesmo tempo, um esquema de codificação preferido (por exemplo, fala/áudio/música).

[00039] De acordo com configurações adicionais, a etapa de decisão 220 é operante para analisar o sinal de áudio 105, uma primeira saída do primeiro codificador 210<sup>a</sup>, uma segunda saída do segundo codificador 210b, um sinal obtido pela decodificação de um sinal de saída do codificador 210a ou um segundo codificador 210b em relação à informação-alvo. A etapa de decisão 220 pode, opcionalmente, ser operante para efetuar uma discriminação de fala/música, de modo que uma escolha para fala seja favorecida em relação a uma escolha para música, assim uma decisão para fala é tomada, por exemplo, inclusive quando uma porção inferior a 50% de um frame para o primeiro switch é de fala e uma porção superior a 50% do frame para o primeiro switch é de música. Portanto, a etapa de decisão 220 pode compreender uma ferramenta de análise que analisa o sinal de áudio para decidir se o sinal de áudio é fundamentalmente um sinal de fala ou fundamentalmente um sinal de música, assim, com base no resultado da etapa de decisão, é possível decidir qual é o melhor codificador a ser utilizado para a porção de tempo analisada do sinal de áudio.

[00040] As Figs. 1 e 2 não mostram diversos dos referidos detalhes para o codificador/decodificador. Os exemplos detalhados possíveis para o codificador/decodificador são mostrados nas figuras seguintes. Ademais, para o primeiro e o segundo decodificador 110a, b da Fig. 1, decodificadores adicionais podem ser apresentados, os quais podem ou não utilizar, por exemplo, algoritmos adicionais de codificação. Da mesma forma, também o codificador 200 da Fig. 2 pode compreender codificadores adicionais, que podem utilizar algoritmos adicionais de codificação. A seguir, o exemplo com dois codificadores/decodificadores será explicado com mais detalhes.

[00041] A Fig. 3 ilustra, em mais detalhes, um codificador que tem dois switches em cascata. Um sinal mono, um sinal estéreo ou um sinal de múltiplos canais é introduzido em uma etapa de decisão 220 e em um switch 232, que é parte do módulo BWE 230 da Fig. 2. O switch 232 é controlado pela etapa de decisão 220. De modo alternativo, a etapa de decisão 220 também pode receber uma informação colateral que é incluída no sinal mono, no sinal estéreo ou no sinal de múltiplos canais ou, ao menos, associada a um determinado sinal, onde a informação existente, que foi, por exemplo, gerada ao produzir originalmente o sinal mono, o sinal estéreo ou o sinal de múltiplos canais

[00042] A etapa de decisão 220 ativa o switch 232 para alimentar um sinal tanto na porção de codificação de frequência 210b, ilustrada agora na ramificação superior da Fig. 3, quanto uma porção de codificação de domínio LPC 210a, ilustrada na ramificação inferior da Fig. 3. Um elemento principal da

ramificação da codificação do domínio de frequência é um bloco de conversão de espectro 410, que é operante para converter um sinal de saída da etapa de pré-processamento comum (conforme discutido adiante) em um domínio espectral. O bloco de conversão de espectro pode incluir um algoritmo MDCT, QMF, um algoritmo de FFT, uma análise de wavelet ou um banco de filtros, tais como um banco de filtros criticamente amostrado que tem um número determinado de canais de banco de filtros, onde os sinais de sub-bandas, neste banco de filtros, podem ser sinais avaliados reais ou sinais avaliados complexos. A saída do bloco de conversão de espectro 410 está codificada usando-se um codificador espectral de áudio 421 que pode incluir blocos de processamento, conhecidos do esquema de codificação AAC.

[00043] De modo geral, o processamento na ramificação 210b é um processamento com base em um modelo baseado na percepção ou modelo aprofundado de informação. Portanto, esta ramificação modela o sistema auditivo humano que recebe o som. Contrário a isso, o processamento na ramificação 210a é gerar um sinal na excitação, residual ou de domínio de LPC. Geralmente, o processamento na ramificação 210a é um processamento baseado em um modelo de fala ou em um modelo de geração de informação. Para sinais de fala, este modelo é um modelo do sistema de geração de fala/som humano gerando som. Entretanto, se um som proveniente de uma fonte diferente exigir que um modelo diferente de geração de som seja codificado, então o processamento na ramificação 210a pode ser diferente. Ademais, para as ramificações de codificação mostradas, as configurações adicionais compreendem ramificações

adicionais ou codificadores centrais. Por exemplo, diferentes codificadores podem, opcionalmente, ser apresentados para diferentes fontes, assim o som proveniente de cada fonte pode ser codificado empregando-se um codificador preferido.

[00044] Na ramificação de codificação inferior 210a, um elemento principal é um dispositivo de LPC 510 que fornece informação LPC, a qual é utilizada para controlar as características de um filtro LPC. A referida informação LPC é transmitida para um decodificador. O sinal de saída da etapa LPC 510 é um sinal de domínio de LPC que consiste de um sinal de excitação e/ou de um sinal ponderado.

[00045] O dispositivo LPC, geralmente, fornece um sinal de domínio de LPC que pode ser qualquer sinal no domínio de LPC ou qualquer outro sinal que foi gerado por meio da aplicação de coeficientes de filtro LPC para um sinal de áudio. Ademais, um dispositivo LPC também pode determinar os referidos coeficientes e pode também quantizar/codificar os referidos coeficientes.

[00046] A escolha na etapa de decisão 220 pode ser adaptável ao sinal, assim a etapa de decisão efetua uma discriminação de música/fala e controla o switch 232, de modo que os sinais de música são introduzidos na ramificação superior 210b, e os sinais de fala são introduzidos na ramificação inferior 210a. Em uma configuração, a etapa de decisão 220 alimenta sua informação de controle em um fluxo de bits de saída, assim um decodificador pode utilizar a referida informação de decisão para efetuar as operações corretas de decodificação. Esta informação de decisão pode, por exemplo, compreender a informação do modo de

codificação 108 que pode compreender também informação sobre a frequência cruzada fx ou uma alteração da frequência cruzada fx.

[00047] O referido decodificador é ilustrado na Fig. 4. A saída do sinal do codificador espectral de áudio 421 é, após a transmissão, introduzida em um decodificador espectral de áudio 431. A saída do decodificar espectral de áudio 431 é introduzida no conversor de domínio de tempo 440 (o conversor de domínio de tempo pode, de modo geral, ser um conversor de um primeiro para um segundo domínio). De modo análogo, a saída da ramificação de codificação do domínio de LPC 210a, da Fig. 3, foi recebida no lado do decodificador e processada pelos elementos 531, 533, 534 e 532 para obter um sinal de excitação de LPC. O sinal de excitação de LPC é introduzido em uma etapa de síntese de LPC 540, que recebe, como uma entrada adicional, a informação de LPC gerada pela etapa de análise de LPC respectiva 510. A saída do conversor de domínio do tempo 440 e/ou a saída da etapa de síntese de LPC 540 são introduzidas em um switch 132, que pode ser parte do módulo BWE 130, na Fig. 1. O switch 132 é controlado via sinal de controle do switch (tais como, a informação do modo de codificação 108 e/ou o parâmetro BWE 106) que foi, por exemplo, gerado pela etapa de decisão 220, ou que foi fornecido externamente como por um criador do sinal mono original, sinal estéreo ou sinal de múltiplos canais.

[00048] Na Fig. 3, o sinal de entrada no switch 232 e a etapa de decisão 220 podem ser um sinal mono, um sinal estéreo, um sinal de múltiplos canais ou, geralmente, qualquer sinal de áudio. De acordo com a decisão que pode ser derivada do sinal de entrada

do switch 232 ou de qualquer fonte externa, tais como produtor de sinal de áudio original subjacente à entrada de sinal na etapa 232, o switch comuta entre a ramificação de codificação de frequência 210b e a ramificação de codificação de LPC 210a. A ramificação de codificação de frequência 210b compreende uma etapa de conversão de espectro 410 e uma etapa de quantização/codificação subseqüentemente conectada 421. A etapa de quantização/codificação pode incluir qualquer funcionalidade conhecida dos codificadores modernos de domínio de frequência, tais como o codificador AAC. Ademais, a operação de quantização na etapa de quantização/codificação 421 pode ser controlada por meio de um módulo psicoacústico que gera informação psicoacústica, tais como limite de mascaramento psicoacústico sobre a frequência, onde esta informação é introduzida na etapa 421.

[00049] Na ramificação de codificação de LPC 210a, o sinal de saída do switch é processado por meio de uma etapa de análise de LPC 510 que gera uma informação colateral de LPC e um sinal de domínio de LPC. O codificador de excitação pode compreender um switch adicional para comutar o processamento adicional do sinal de domínio de LPC entre uma operação de quantização/codificação 522, no domínio de LPC, ou uma de quantização/codificação 524 que processa valores no domínio espectral de LPC. Para esta finalidade, um conversor de espectro 523 é fornecido na entrada da etapa de quantização/codificação 524. O switch 521 é controlado em uma configuração de circuito aberto ou em uma configuração de circuito fechado, conforme as configurações específicas, como, por exemplo, descritas na

especificação técnica do AMR-WB+.

[00050] Para o modo de controle de circuito fechado, o codificador inclui adicionalmente um codificador/quantizador inverso 531 para o sinal de domínio de LPC, um codificador/quantizador inverso 533 para o sinal de domínio espectral de LPC e um conversor de espectro inverso 534 para a saída do item 533. Os dois sinais codificados e, novamente decodificados nas ramificações de processamento, da segunda ramificação de codificação, são inseridos no dispositivo de controle do switch 525. No dispositivo de controle do switch 525, estes dois sinais de saída são comparados entre si e/ou para uma função-alvo ou uma função-alvo é calculada, que pode estar baseada em uma comparação da distorção nos dois sinais, assim o sinal que tem a distorção inferior é utilizado para escolher qual posição o switch 521 deve tomar. De modo alternativo, no caso as duas ramificações fornecem taxas de bits não constantes, a ramificação que fornece a taxa de bits mais baixa deve ser selecionada, inclusive quando a distorção ou a distorção perceptiva desta ramificação é inferior à distorção ou distorção perceptiva da outra ramificação (um exemplo para a distorção pode ser a relação sinal-ruído). De modo alternativo, a função-alvo pode ser utilizada, como uma entrada, a distorção de cada sinal e uma taxa de bits de cada sinal e/ou critérios adicionais para localizar a melhor decisão para uma meta específica. Se, por exemplo, a meta é que a taxa de bits deve ser a mais baixa possível, então a função-alvo deve demasiadamente recair na taxa de bits dos dois sinais fornecidos pelos elementos 531, 534. Entretanto, quando a meta

principal é ter a melhor qualidade para uma determinada taxa de bits, então o controle do switch 525 deve, por exemplo, descartar cada sinal que esteja acima da taxa de bits permitida e, quando os dois sinais estiverem abaixo da taxa de bits permitida, o controle do switch deve selecionar o sinal que tem a melhor qualidade subjetiva estimada, ou seja, que tem as menores distorções de quantização/codificação ou uma melhor relação sinal-ruído.

[00051] O esquema de decodificação, em conformidade com uma configuração é, conforme declarado anteriormente, ilustrado na Fig. 4. Para cada um dos três tipos de sinais de saída possíveis, existe uma etapa específica de decodificação/requantização 431, 531 ou 533. Enquanto a etapa 431 fornece um espectro de frequência que é convertido no domínio de tempo usando-se um conversor de frequência/tempo 440, a etapa 531 fornece um sinal de domínio de LPC, bem como o item 533 fornece um espectro de LPC. Para certificar-se de que os sinais de entrada no switch 532 estão nos dois domínios de LPC, é fornecido o LPC de espectro/conversor de LPC 534. Os dados de saída do switch 532 são novamente transformados no domínio de tempo usando-se uma etapa de síntese de LPC 540, a qual é controlada por meio da informação de LPC gerada e transmitida pelo lado do codificador. Em seguida, subsequente ao bloco 540, as duas ramificações têm informação de domínio de tempo que é comutada em conformidade com um sinal de controle do switch para, finalmente, obter um sinal de áudio, tais como um sinal mono, um sinal estéreo ou sinal de múltiplos canais que depende da entrada de sinal no esquema de codificação da Fig. 3.

[00052] As Figs. 5 e 6 mostram configurações adicionais para o codificador/decodificador, onde as etapas de BWE, como parte dos módulos BWE 130, 230 representam uma unidade comum de processamento.

[00053] A Fig. 5 ilustra um esquema de codificação, onde o esquema de pré-processamento comum conectado à entrada do switch 232 pode compreender um bloco *surround/joint stereo* 101 que gera, como uma saída, parâmetros de *joint stereo* e um sinal de saída mono, que é gerado por meio de *downmixing* do sinal de entrada, que é um sinal que tem dois ou mais canais. Geralmente, o sinal na saída do bloco 101 também pode ser um sinal que tem mais canais, porém devido à funcionalidade de *downmixing* do bloco 101, o número de canais na saída do bloco 101 será menor que o número dos canais introduzidos no bloco 101.

[00054] O esquema de pré-processamento comum pode compreender, além do bloco 101, uma etapa de extensão de largura de banda 230. Na configuração da Fig. 5, a saída do bloco 101 é introduzida no bloco de extensão de largura de banda 230, que fornece um sinal limitado à banda, tais como o sinal de banda baixa ou o sinal baixo passante em sua saída. De modo preferido, este sinal também é submetido à *downsampling* (por exemplo, por um fator de dois). Ademais, para a banda alta da entrada de sinal no bloco 230, os parâmetros de extensão de largura de banda 106, tais como parâmetros de envelope espectral, parâmetros de filtragem inversa, parâmetros de patamar de ruído, etc., conhecidos do perfil HE-AAC do MPEG-4, são gerados e encaminhados para um multiplexador de fluxo de bits 800.

[00055] De modo preferido, a etapa de decisão 220 recebe a entrada do sinal no bloco 101 ou introduz no bloco 230 para decidir entre, por exemplo, um modo de música ou um modo de fala. No modo de música, a ramificação superior de codificação 210b (segundo codificador na Fig. 2) é selecionada, ao passo que no modo de fala, é selecionada a ramificação inferior de codificação 210a. De modo preferido, a etapa de decisão controla adicionalmente o bloco *joint stereo* 101 e/ou o bloco de extensão de largura de banda 230 para adaptar a funcionalidade destes blocos ao sinal específico. Portanto, quando a etapa de decisão 220 determina que uma determinada porção de tempo do sinal de entrada se refere ao primeiro modo, como o modo de música, então características específicas do bloco 101 e/ou do bloco 230 podem ser controladas pela etapa de decisão 220. De modo alternativo, quando a etapa de decisão 220 determina que o sinal se refere a um modo de fala ou, geralmente, em um segundo modo de domínio de LPC, então características específicas dos blocos 101 e 230 podem ser controladas em conformidade a saída da etapa de decisão. A etapa de decisão 220 produz também a informação de controle 108 e/ou a frequência cruzada *fx*, a qual também pode ser transmitida para o bloco de BWE 230 e, ademais, para um multiplexador de fluxo de bits 800, assim ela será transmitida ao lado do decodificador.

[00056] De modo preferido, a conversão de espectro da ramificação de codificação 210b é efetuada usando-se uma operação de MDCT que, inclusive de modo preferido, é uma operação de MDCT *time-warped*, onde a potência, ou geralmente, a potência de distorção pode ser controlada entre zero e uma potência de

distorção alta. Em uma potência de distorção zero, a operação de MDCT no bloco 411 é uma operação de MCT direcionada, conhecida na técnica. A potência de distorção de tempo conjuntamente com a informação colateral de distorção de tempo pode ser transmitida/introduzida no multiplexador de taxa de bits 800 como informação colateral.

[00057] Na ramificação de codificação de LPC, o codificador de domínio de LPC pode incluir um núcleo ACELP 526 que calcula uma amplitude de tom, um retardo de tom e/ou informação de codebook, tais como um índice e amplitude de codebook. O modo TCX, conhecido a partir do 3GPP TS 26.290, inclui um processamento de um sinal perceptivamente ponderado em um domínio por transformadas. Um sinal ponderado de transformada de Fourier é quantizado usando-se uma quantização de rede dividida em múltiplas faixas (quantização vetorial algébrica) com quantização de fator de ruído. Uma transformada é calculada em 1024, 512 ou 256 janelas de amostra. O sinal de excitação é recuperado por filtragem inversa do sinal ponderado quantizado por meio de um filtro de ponderação inversa. O modo TCX também pode ser utilizado na forma modificada, na qual a MDCT é utilizada com uma sobreposição aumentada, quantização escalar e codificador aritmético para linhas de codificação espectral.

[00058] Na ramificação de codificação de "música" 210b, um conversor de espectro, preferencialmente, compreende uma operação de MDCT especificamente adaptada que tem determinadas funções de janela seguidas por uma etapa de codificação por quantização/entropia que pode consistir de uma única etapa de

quantização vetorial, porém, preferencialmente, é um codificador escalar combinado de quantizador/entropia semelhante ao quantizador/codificador na ramificação de codificação de domínio de frequência, ou seja, no item 421 da Fig. 5.

[00059] Na ramificação de codificação de “fala” 210a, há o bloco de LPC 510 seguido por um switch 521, novamente seguido por um bloco ACELP 526 ou um bloco TCX 527. A descrição de ACELP está em 3GPP TS 26.190 e a descrição de TCX está no 3GPP TS 26.290. Geralmente, o bloco ACELP 526 recebe um sinal excitação de LPC, calculado por um procedimento conforme descrito na Fig. 7. O bloco TCX 527 recebe um sinal ponderado conforme gerado pela Fig. 8.

[00060] No lado do decodificador ilustrado na Fig. 6, após a transformada espectral inversa no bloco 537, a inversa do filtro de ponderação é aplicada, que é  $(1 - \mu z^{-1}) / (1 - A(z/\gamma))$ . Em seguida, o sinal é filtrado por meio de  $(1 - A(z))$  para avançar ao domínio de excitação de LPC. Assim, a conversão ao bloco de domínio de LPC 534 e ao bloco de  $TCX^{-1}$  537 inclui transformada inversa e, em seguida, filtragem pela fórmula a seguir:

$$\frac{(1 - \mu z^{-1})}{(1 - A(z / \gamma))}$$

[0001] Onde  $1 - A(z)$  refere-se a conversão a partir do domínio ponderado para o domínio de excitação.

[0002] Embora o item 510, nas Figs. 3 e 5, ilustre um bloco único, o bloco 510 pode fornecer sinais diferentes, desde que os referidos sinais estejam no domínio de LPC. O modo atual do bloco 510, tais como o modo do sinal de excitação ou modo de sinal ponderado, pode depender do estado atual do switch. De modo

alternativo, o bloco 510 pode conter dois dispositivos de processamento paralelo, onde um dispositivo é implementado de modo similar à Fig. 7 e o outro dispositivo é implementado conforme a Fig. 8. Portanto, o domínio de LPC na saída do 510 pode representar tanto um sinal de excitação de LPC quanto um sinal ponderado de LPC ou qualquer outro sinal de domínio de LPC.

[0003] Na segunda ramificação de codificação (ACELP/TCX) da Fig. 5, o sinal é preferencialmente pré-enfatizado por meio de um filtro  $1-\mu z^{-1}$ , antes da codificação. No decodificador ACELP/TCX na Fig. 6, é retirada a ênfase do sinal sintetizado com o filtro  $1/(1-\mu z^{-1})$ . Em uma configuração preferida, o parâmetro  $\mu$  tem o valor de 0,68. A pré-ênfase pode ser parte do bloco de LPC 510, onde o sinal é pré-enfatizado antes da análise de LPC e da quantização. De modo semelhante, a retirada da ênfase pode ser parte do bloco de síntese de LPC,  $LPC^{-1}$  540.

[0004] A Fig. 6 ilustra um esquema de decodificação que se refere ao esquema de codificação da Fig. 5. O fluxo de bits gerado pelo multiplexador de fluxo de bits 800 (ou interface de saída) da Fig. 5 é introduzido no demultiplexador de fluxo de bits 900 (ou interface de entrada). Conforme uma informação derivada, por exemplo, do fluxo de bits por meio de um bloco de detecção de modo 601 (por exemplo, parte do controlador 140, na Fig. 1), um switch do lado do decodificador 132 é controlado tanto por sinais encaminhados da ramificação superior quanto por sinais provenientes da ramificação inferior para o bloco de extensão de largura de banda 701. O bloco de extensão de largura de banda 701 recebe, de um demultiplexador de fluxo de bits 900, a informação

colateral e, baseado nesta informação colateral e na saída da detecção de modo 601, é reconstruída a banda alta baseada na saída de banda baixa pelo switch 132. O sinal de controle 108 controla a frequência cruzada *fx* utilizada.

[0005] O sinal de banda completa gerado pelo bloco 701 é introduzido na etapa de processamento de *joint stereo/surround* 702, que reconstrói dois canais estéreos ou vários múltiplos canais. Geralmente, o bloco 702 fornecerá mais canais do que foram introduzidos neste bloco. Conforme a aplicação, a entrada no bloco 702 pode, inclusive, incluir dois canais, tais como em um modo estéreo e pode, inclusive, incluir mais canais, desde que a saída deste bloco tenha mais canais que a entrada neste bloco.

[0006] O switch 232 na Fig. 5 foi mostrado para comutar entre ramificações, assim somente uma ramificação recebe um sinal para processar e a outra ramificação não recebe um sinal para processar. Entretanto, em uma configuração alternativa, o switch 232 também pode ser organizado subsequente ao, por exemplo, codificador de áudio 421 e ao codificador de excitação 522, 523, 524, que significa que as duas ramificações 210a, 210b processam o mesmo sinal em paralelo. Entretanto, para não duplicar a taxa de bits, somente a saída de sinal de uma das ramificações de codificação 210a ou 210b é selecionada para ser escrita no fluxo de bits de saída. Em seguida, etapa de decisão operará, assim o sinal escrito no fluxo de bits minimiza uma determinada função de custo, onde a função de custo pode ser a taxa de bits gerada, a distorção perceptual gerada ou a função de custo combinada de taxa/distorção. Portanto, tanto neste modo quanto no modo

ilustrado nas Figuras, a etapa de decisão também pode operar em um modo de circuito fechado, para garantir que, finalmente, apenas uma saída de ramificação de codificação seja escrita no fluxo de bits, o qual tem, para uma distorção perceptual dada, a taxa de bits mais baixa ou, para uma dada taxa de bits, a distorção percebida mais baixa. No modo de circuito fechado, a entrada de *feedback* pode ser derivada das saídas dos três blocos de quantizador/multiplicador de frequências 421, 522 e 424, na Fig. 3.

[0007] Também na configuração da Fig. 6, o switch 132 pode, em configurações alternativas, ser organizado após o módulo BWE 701, assim a extensão de largura de banda é efetuada em paralelo, para as duas ramificações, e o switch seleciona um dos dois sinais estendidos de largura de banda.

[0008] Na implementação que tem dois switches, ou seja, o primeiro switch 232 e o segundo switch 521, é preferido que a resolução de tempo, para o primeiro switch, seja menor que a resolução de tempo para o segundo switch. Diferentemente declarado, os blocos do sinal de entrada no primeiro switch, que podem ser comutados via operação de comutação, são maiores que os blocos comutados pelo segundo switch 521, que opera no domínio de LPC. De modo exemplar, os switches de domínio de frequência/domínio de LPC 232 podem comutar blocos de uma extensão de 1024 amostras, e o segundo switch 521 pode comutar blocos que tem 256 amostras cada.

[0009] A Fig. 7 ilustra uma implementação mais detalhada do bloco de análise de LPC 510. O sinal de áudio é introduzido em um bloco de determinação de filtro 83, que determina a informação

de filtro  $A(z)$ . Esta informação é fornecida como informação preditiva em curto prazo, exigida para um decodificador. A informação preditiva em curto prazo é exigida pelo filtro preditivo atual 85. No subtrator 86, uma amostra atual do sinal de áudio é introduzida e um valor predito para a amostra atual é subtraído, assim, para esta amostra, o sinal de erro de predição é gerado na linha 84.

[00010] Enquanto a Fig. 7 ilustra o modo preferido para calcular o sinal de excitação, a Fig. 8 ilustra o modo preferido para calcular o sinal ponderado. Em oposição à Fig. 7, o filtro 85 é diferente, quando  $y$  é diferente de 1. Um valor menor que 1 é preferido para  $y$ . Ademais, o bloco 87 está presente, e  $\mu$  é preferível com um número menor que 1. Geralmente, os elementos na Fig. 7 e 8 podem ser implementados conforme em 3GPP TS 26.190 e 3GPP TS 26.290.

[00011] De modo subsequente, um codificador CELP de *analysis-by-synthesis* é discutido para ilustrar as modificações aplicadas para este algoritmo. Este codificador CELP é discutido em detalhes em "Speech Coding: A Tutorial Review" ["Codificação de fala: Uma revisão tutorial"], Andreas Spanias, Proceedings of the IEEE [Procedimentos do IEEE], Vol. 82, No. 10, Outubro de 1994, páginas 1541-1582.

[00012] Para casos específicos, quando um frame é uma mistura de uma fala sem voz e com voz, ou quando ocorre uma fala sobre a música, uma codificação TCX pode ser mais adequada para codificar a excitação no domínio de LPC. A codificação TCX processa diretamente a excitação no domínio de frequência sem

fazer qualquer suposição da produção de excitação. A TCX é, em seguida, mais genérica que a codificação CELP e não está restrita a uma modelo de fonte com voz ou sem voz de excitação. A TCX ainda é um modelo de codificação de filtro de fonte que utiliza um filtro preditivo linear para modelar os formantes dos sinais semelhantes à fala.

[00013] Na codificação semelhante a AMR-WB+, uma seleção entre diferentes modos de TCX e ACELP ocorre conforme conhecido a partir da descrição de AMR-WB+. Os modos de TCX são diferentes na referida extensão de blocos do tipo Transformada Rápida de Fourier diferente para modos diferentes e o melhor modo pode ser selecionado por meio de análise para abordagem de síntese ou por meio de um modo direto de "alimentação".

[00014] Conforme discutido em relação à Fig. 5 e 6, a etapa de pré-processamento comum 100 inclui, preferencialmente, um *joint multi-channel* (dispositivo *surround/joint stereo*) 101 e, adicionalmente, uma etapa de extensão de largura de banda 230. De modo respectivo, um decodificador inclui uma etapa de extensão de largura de banda 701 e, subseqüentemente, uma etapa de *joint multi-channel* conectada 702. Preferencialmente, a etapa de *joint multi-channel* 101 é, em relação ao codificador, conectada antes que a etapa de extensão de largura de banda 230, e, no lado do decodificador, a etapa de extensão de largura de banda 701 é conectada antes da etapa de *joint multi-channel* 702 em relação à direção do processamento de sinal. Entretanto, de modo alternativo, a etapa de pré-processamento comum pode incluir uma etapa de *joint multi-channel* sem a subseqüente etapa de extensão

de largura de banda conectada ou uma etapa de extensão de largura de banda sem uma etapa de *joint multi-channel* conectada.

[00015] As Figs. 9a e 9b mostram uma vista simplificada de um codificador da Fig. 5, onde o codificador compreende uma unidade de escolha de switch 220 e uma unidade de codificação estéreo 101. Ademais, o codificador também compreende as ferramentas de extensão de largura de banda 230 como, por exemplo, um calculador de dados de envelope e módulos relativos à SBR. A unidade de escolha de switch 220 fornece um sinal de escolha de switch 108' que comuta entre o codificador de áudio 210b e o codificador de fala 210a. O codificador de fala 210a pode, ainda, ser dividido em um codificador com voz e sem voz. Cada um destes codificadores pode codificar o sinal de áudio na banda de frequência central usando números diferentes de valores de amostra (por exemplo, 1024 para uma resolução superior ou 256 para uma resolução inferior). O sinal de escolha de switch 108' também é fornecido para a ferramenta de extensão de largura de banda (BWE) 230. Em seguida, a ferramenta de BWE 230 utilizará a escolha de switch 108' para, por exemplo, ajustar o número de envelopes espectrais 104 e ligar/desligar um detector transiente opcional, bem como ajustar a frequência cruzada  $f_x$ . O sinal de áudio 105 é introduzido na unidade de escolha de switch 220 e é introduzido na codificação estéreo 101, assim, a codificação estéreo 101 pode produzir os valores de amostra que são introduzidos na unidade de extensão de largura de banda 230. Conforme a decisão 108' gerada pela unidade de escolha de unidade de switch 220, a ferramenta de extensão de largura de banda 230 gerará dados de replicação de

banda espectral que são, por sua vez, encaminhados tanto para um codificador de áudio 210b quanto para um codificador de fala 210a.

[00016] O sinal de escolha de switch 108' é dependente do sinal e pode ser obtido a partir da unidade de escolha de switch 220, por meio da análise de sinal de áudio, por exemplo, usando-se um detector transiente ou outros detectores que podem ou não compreender um limite variável. De modo alternativo, o sinal de escolha do switch 108' pode ser manualmente ajustado (por exemplo, por um usuário) ou obtido a partir de um fluxo de dados (incluído no sinal de áudio).

[00017] A saída do codificador de áudio 210b e o codificador de fala 210a podem, novamente, ser introduzidos no formatador de fluxo de dados 800 (vide Fig. 5).

[00018] A Fig. 9b mostra um exemplo para o sinal de escolha de switch 108', que detecta um sinal de áudio para um período de tempo antes de um período de tempo  $t_a$ , e após um segundo tempo  $t_b$ . Entre o primeiro tempo  $t_a$  e o segundo tempo  $t_b$ , a unidade de escolha de switch 220 detecta um sinal de fala resultante em diferentes valores discretos para o sinal de escolha de switch 108'.

[00019] A decisão de utilizar uma frequência cruzada  $f_x$  superior é controlada por meio da unidade de escolha de switch 220. Isso significa que o método descrito também é passível de utilização dentro de um sistema no qual o módulo de SBR é combinado somente com um codificador único central e uma frequência cruzada  $f_x$  variável.

[00020] Embora algumas das Figs. 1 até 9 sejam ilustradas

como diagramas de bloco de um aparelho, as referidas figuras são, simultaneamente, uma ilustração de um método, onde as funcionalidades de bloco se referem às etapas de método.

[00021] A Fig. 10 ilustra uma representação para um sinal de áudio codificado 102, que compreende a primeira porção 104a, a segunda porção 104b, a terceira porção 104c e uma quarta porção 104d. Nesta representação, o sinal de áudio codificado 102 é um fluxo de bits transmitido por um canal de transmissão que compreende ainda a informação de modo de codificação 108. Cada porção 104 do sinal de áudio codificado 102 pode representar uma porção de tempo diferente, embora diferentes porções 104 possam estar na frequência, bem como no domínio de tempo, assim o sinal de áudio codificado 102 pode não representar uma linha de tempo.

[00022] Nesta configuração, o sinal de áudio codificado 102 compreende, ademais, uma primeira informação de modo de codificação 108a que identifica o algoritmo de codificação utilizado para a primeira porção 104a; uma segunda informação de modo de codificação 108b que identifica o algoritmo de codificação utilizado para a segunda porção 104b; uma terceira informação de modo de codificação 108c que identifica o algoritmo de codificação utilizado para a terceira porção 104c; e uma quarta informação de modo de codificação 108d que identifica o algoritmo de codificação utilizado para a quarta porção 104d. A primeira informação de modo de codificação 108a também pode identificar a primeira frequência cruzada fx1 utilizada dentro da primeira porção 104a, e a segunda informação de modo de codificação 108b também pode identificar a segunda frequência cruzada fx2 utilizada dentro da segunda porção 104b. Por exemplo, dentro da primeira porção 104a, o modo de codificação de "fala" pode ser utilizado e dentro da segunda

porção 104b o modo de codificação de “música” pode ser utilizado, assim, a primeira frequência cruzada  $fx_1$  pode ser superior a segunda frequência cruzada  $fx_2$ .

[00023] Nesta configuração exemplar, o sinal de áudio codificado 102 não compreende nenhuma informação de modo de codificação para a terceira porção 104c que indica que não há alteração no codificador utilizado e/ou na frequência cruzada  $fx$  entre a primeira e a terceira porção 104a, c. Portanto, a informação do modo de codificação 108 pode aparecer como cabeçalho somente para as porções 104 que utilizam um codificador central diferente e/ou frequência cruzada comparada à porção anterior. Em configurações adicionais, em vez de sinalizar os valores das frequências cruzadas para as porções diferentes 104, a informação de modo de codificação 108 pode compreender um único bit que indica o codificador central (primeiro ou segundo codificador 210a, b) utilizado para a respectiva porção 104.

[00024] Portanto, a sinalização do comportamento do switch entre as diferentes ferramentas de SBR pode ser feita, por exemplo, submetendo-se um bit específico dentro de um fluxo de bits, assim o referido bit pode ativar ou desativar um comportamento específico no decodificador. De modo alternativo, nos sistemas com dois codificadores centrais, de acordo com as configurações, a sinalização do switch também pode ser iniciada por meio da análise do codificador central. Neste caso, a submissão da adaptação das ferramentas de SBR é feita de modo implícito, o que significa que é determinada por meio de atividade respectiva do codificador central.

[00025] Mais detalhes sobre a descrição padrão dos elementos de fluxo de bits para a carga útil de SBR podem ser encontrados em ISO/IEC 14496-3, subcategorias 4.5.2.8. Uma modificação deste fluxo de bits padrão compreende uma extensão do índice para uma tabela de frequência mestre (para identificar a frequência cruzada utilizada). O índice utilizado é codificado, por exemplo, com quatro bits que permitem que a banda cruzada seja variável sobre uma faixa de 0 a 15 bandas.

[00026] As configurações da presente invenção podem, portanto, ser resumidas como segue. Sinais diferentes com diferentes características de tempo/frequência têm diferentes demandas sobre a característica da extensão de largura de banda. Os sinais transientes (por exemplo, dentro de um sinal de fala) precisam de uma boa resolução temporal da BWE, e a frequência cruzada  $f_x$  (a borda superior de frequência do codificador central) deve ser a mais alta possível (por exemplo, 4 kHz, 5 kHz ou 6 kHz). Especialmente na fala com voz, uma estrutura temporal distorcida pode diminuir a qualidade percebida. Os sinais tonais precisam de uma reprodução estável dos componentes espectrais e um padrão harmônico de combinação das porções reproduzidas de alta frequência. A reprodução estável das partes tonais limita a largura de banda do codificador central, porém não precisa de uma BWE com boa temporal, mas com uma resolução espectral melhor. No desenho do codificador central comutado de fala/áudio, também é possível utilizar a escolha do codificador central para adaptar as características temporais e espectrais da BWE, bem como para adaptar a frequência BWE inicial (frequência cruzada) para as

características de sinal. Portanto, as configurações fornecem uma extensão de largura de banda onde a escolha do codificador central atua como critério de adaptação para características de extensão de largura de banda.

[00027] A sinalização da frequência (cruzada) inicial alterada de BWE pode ser realizada explicitamente por meio do envio de informação adicional (como, por exemplo, a informação do modo de codificação 108) no fluxo de bits ou implicitamente por meio de derivação da frequência cruzada  $f_x$  diretamente a partir do codificador central utilizado (no caso do codificador central é, por exemplo, sinalizado dentro do fluxo de bits). Por exemplo, uma frequência  $f_x$  de BWE inferior para o codificador por transformadas (por exemplo, codificador de áudio/música) e uma frequência superior para um codificador de domínio de tempo (fala). Neste caso, a frequência cruzada pode estar na faixa entre 0 hz até a frequência de Nyquist.

[00028] Embora alguns aspectos sejam descritos no contexto do aparelho, fica claro que os referidos aspectos também representam uma descrição do método respectivo, onde um bloco ou dispositivo se refere a uma etapa de método ou a uma característica de uma etapa de método. De modo análogo, os aspectos descritos no contexto de uma etapa de método também representam uma descrição de um bloco, item ou características respectivas de um aparelho pertinente.

[00029] O sinal de áudio codificado inventado pode ser armazenado em um meio de armazenamento digital ou pode ser transmitido por um meio de transmissão, tais como um meio de

transmissão sem fio ou um meio de transmissão por cabo, tais como a Internet.

[00030] De acordo com determinados requisitos de implementação, as configurações da invenção podem ser implementadas em hardware ou software. A implementação pode ser efetuada utilizando-se um meio de armazenamento digital, por exemplo, um disco flexível, um DVD, CD, CD-ROM, PROM, EPROM, EEPROM ou um cartão de memória, que contem com sinais de controle que sejam eletronicamente passíveis de leitura do conteúdo armazenado, que sejam compatíveis (ou sejam capazes de compatibilidade) com um sistema programável de computador, para que assim o respectivo método seja realizado.

[00031] Algumas configurações, de acordo com a invenção, compreendem um dispositivo para o transporte de dados que conte com sinais de controle que sejam eletronicamente passíveis de leitura, os quais sejam compatíveis com um sistema programável de computador, para que assim um dos métodos ora descritos seja realizado.

[00032] De modo geral, as configurações da presente invenção podem ser implementadas como um produto de programa de computador com um código de programa, o código de programa sendo operante para efetuar um dos métodos, quando o produto do programa de computador seja executado em um computador. O código de programa pode, por exemplo, ser armazenado em um dispositivo de transporte passível de leitura em uma máquina.

[00033] Outras configurações compreendem o programa de computador para a realização de um dos métodos ora descritos,

armazenados em um dispositivo de transporte passível de leitura em uma máquina, ou seja, uma configuração do método de invenção é, portanto, um programa de computador que tem um código de programa para efetuar um dos métodos ora descritos, quando o programa de computador é executado em um computador.

[00034] Uma configuração adicional dos métodos de invenção é, portanto, um dispositivo de transporte de dados (ou um meio de armazenamento digital, ou um meio de leitura em computador) que compreende o conteúdo gravado em si, o programa de computador para efetuar um dos métodos ora descritos.

[00035] Uma configuração adicional do método de invenção é, portanto, um fluxo de dados ou uma seqüência de sinais que representam o programa de computador para efetuar um dos métodos ora descritos. O fluxo de dados ou a seqüência de sinais pode, por exemplo, ser configurada para ser transferida via conexão de comunicação de dados, por exemplo, via Internet.

[00036] Uma configuração adicional compreende um meio de processamento, por exemplo, um computador, um dispositivo lógico programável, configurado para ou adaptado para efetuar um dos métodos ora descritos.

[00037] Uma configuração adicional compreende um computador que tem instalado o programa de computador para efetuar um dos métodos ora descritos.

[00038] Em algumas configurações, um dispositivo lógico programável (por exemplo, um circuito FPGA - "Field Programmable Gate Array") pode ser utilizado para efetuar algumas ou todas as funcionalidades dos métodos ora descritos. Em algumas

configurações, um circuito FPGA pode ser compatível com um microprocessador para efetuar um dos métodos ora descritos. De modo geral, os métodos são preferencialmente realizados por qualquer aparelho de hardware.

[00039] As configurações acima descritas são meramente ilustrativas para os princípios da presente invenção. É entendido que as modificações e variações dos arranjos e dos detalhes ora descritos serão óbvios para os demais especialistas na técnica. Portanto, é pretendida a limitação somente pelo escopo das reivindicações de patente pendentes e não por detalhes específicos apresentado pela forma de descrição e explicação das configurações ora contidas.

REIVINCAÇÕES

1. Um aparelho para decodificar (100) um sinal de áudio codificado (102), o sinal de áudio codificado compreendendo uma primeira parte (104a) codificada de acordo com um primeiro algoritmo de codificação, uma segunda parte (104b) codificada de acordo com um segundo algoritmo de codificação, parâmetros BWE (106) para a primeira parte e a segunda parte e uma informação de modo de codificação (108) indicando um primeiro algoritmo de decodificação ou um segundo algoritmo de decodificação, caracterizado pelo fato de que compreende:

um primeiro decodificador (110a) para decodificar a primeira parte de acordo com o primeiro algoritmo de decodificação para uma primeira parte de tempo do sinal codificado para adquirir um primeiro sinal decodificado (114a), em que o primeiro decodificador compreende um codificador baseado em LPC;

um segundo decodificador (110b) para decodificar a segunda parte de acordo com o segundo algoritmo de decodificação para uma segunda parte de tempo do sinal codificado para adquirir um segundo sinal decodificado , (114b) em que o segundo decodificador compreende um codificador baseado em transformação;

um módulo BWE (130) que compreende uma frequência de cruzamento controlável (fx), o módulo BWE sendo configurado para executar um algoritmo de extensão de largura de banda usando o primeiro sinal decodificado e os

parâmetros BWE para a primeira parte e para executar um algoritmo de extensão de largura de banda usando o segundo sinal decodificado e a extensão de largura de banda parâmetro para a segunda parte,

em que o módulo BWE é configurado para usar uma primeira frequência de cruzamento (fx1) para a extensão de largura de banda para o primeiro sinal decodificado e para usar uma segunda frequência de cruzamento (fx2) para a extensão de largura de banda para o segundo sinal decodificado,

em que a primeira frequência de crossover é maior do que a segunda frequência de crossover; e

um controlador (140) para controlar a frequência de cruzamento para o módulo BWE de acordo com as informações do modo de codificação.

2. Aparelho para decodificação, de acordo com a reivindicação 1, caracterizado pelo fato de que compreende ainda uma interface (900) de entrada para inserir o sinal de áudio codificado como um fluxo de bits.

3. Aparelho para decodificação, de acordo com a reivindicação 1, caracterizado pelo fato de que o módulo BWE compreende um comutador (132) que é configurado para alternar entre a primeira e a segunda porção de tempo do primeiro decodificador para o segundo decodificador de modo que o algoritmo de extensão de largura de banda seja aplicado ao primeiro decodificado sinal ou para o segundo sinal decodificado.

4. Aparelho para decodificação, de acordo

com a reivindicação 3, caracterizado pelo fato de que o controlador é configurado para controlar a chave dependente do algoritmo de decodificação indicado dentro das informações do modo de codificação.

5. Aparelho para decodificação, de acordo com a reivindicação 1, caracterizado pelo fato de que o controlador é configurado para aumentar a frequência de crossover dentro da primeira porção de tempo ou para diminuir a frequência de crossover dentro da segunda porção de tempo.

6. Aparelho para codificar (132) um sinal de áudio (105), caracterizado pelo fato de que compreende:

um primeiro codificador (210a) que é configurado para codificar de acordo com um primeiro algoritmo de codificação, o primeiro algoritmo de codificação compreendendo uma primeira largura de banda de frequência, em que o primeiro codificador compreende um codificador baseado em LPC;

um segundo codificador (210b) que é configurado para codificar de acordo com um segundo algoritmo de codificação, o segundo algoritmo de codificação compreendendo uma segunda largura de banda de frequência sendo menor do que a primeira largura de banda de frequência, em que o segundo codificador compreende um codificador baseado em transformada;

um estágio de decisão (220) para indicar o primeiro algoritmo de codificação para uma primeira parte (204a) do sinal de áudio e para indicar o segundo algoritmo

de codificação para uma segunda parte (204b) do sinal de áudio, a segunda parte sendo diferente da primeira parte; e um módulo de extensão de largura de banda (230) para calcular os parâmetros BWE (106) para o sinal de áudio, em que o módulo BWE é configurado para ser controlado pelo estágio de decisão para calcular os parâmetros BWE para uma banda que não compreende a primeira largura de banda de frequência na primeira parte do sinal de áudio e para uma banda que não compreende a largura de banda de segunda frequência na segunda parte do sinal de áudio,

em que a primeira ou a segunda largura de banda de frequência é definida por uma frequência de crossover variável ( $f_x$ ) e em que o estágio de decisão é configurado para emitir a frequência de crossover variável,

em que o módulo BWE é configurado para usar uma primeira frequência de cruzamento ( $f_{x1}$ ) para calcular os parâmetros BWE para um sinal codificado usando o primeiro codificador e para usar uma segunda frequência de cruzamento ( $f_{x2}$ ) para um sinal codificado usando o segundo codificador, em que a primeira frequência de cruzamento é mais alta do que a segunda frequência de cruzamento.

7. Aparelho para codificação, de acordo com a reivindicação 6, caracterizado pelo fato de que compreende ainda uma interface de saída (800) para emitir o sinal de áudio codificado, o sinal de áudio codificado compreendendo uma primeira parte (104a) codificada de acordo com um primeiro algoritmo de codificação, uma

segunda parte (104b) codificada de acordo com um segundo algoritmo de codificação , Parâmetros BWE (106) para a primeira parte e a segunda parte e informações do modo de codificação (106) indicando o primeiro algoritmo de decodificação ou o segundo algoritmo de decodificação.

8. Aparelho para codificação, de acordo com a reivindicação 6, caracterizado pelo fato de que a primeira ou a segunda largura de banda de frequência é definida por uma frequência de crossover variável e em que o estágio de decisão é configurado para emitir a frequência de crossover variável.

9. Aparelho para codificação, de acordo com a reivindicação 6, caracterizado pelo fato de que o módulo BWE compreende uma chave (232) controlada pelo estágio de decisão, em que a chave é configurada para alternar entre o primeiro e o segundo codificador de tempo de modo que o sinal de áudio seja para diferentes porções de tempo codificadas por o primeiro ou pelo segundo codificador.

10. Aparelho para codificação, de acordo com a reivindicação 6, caracterizado pelo fato de que o estágio de decisão (220) é operativo para analisar o sinal de áudio ou uma primeira saída do primeiro codificador ou uma segunda saída do segundo codificador ou um sinal adquirido pela decodificação de um sinal de saída do primeiro codificador ou o segundo codificador em relação a uma função alvo.

11. Método para decodificar um sinal de áudio codificado (102), o sinal de áudio codificado

compreendendo uma primeira parte (104a) codificada de acordo com um primeiro algoritmo de codificação, uma segunda parte (104b) codificada de acordo com um segundo algoritmo de codificação, parâmetros BWE (106) para a primeira parte e a segunda parte e uma informação de modo de codificação (108) indicando um primeiro algoritmo de decodificação ou um segundo algoritmo de decodificação, método caracterizado pelo fato de que compreende:

decodificar a primeira porção de acordo com o primeiro algoritmo de decodificação para uma primeira porção de tempo do sinal codificado para adquirir um primeiro sinal decodificado (114a), em que a decodificação da primeira porção compreende o uso de um codificador baseado em LPC;

decodificar a segunda porção de acordo com o segundo algoritmo de decodificação para uma segunda porção de tempo do sinal codificado para adquirir um segundo sinal decodificado (114b), em que a decodificação da segunda porção compreende o uso de um codificador baseado em transformação;

realizar um algoritmo de extensão de largura de banda por um módulo BWE (130) que compreende uma frequência de cruzamento controlável ( $f_x$ ), usando o primeiro sinal decodificado e os parâmetros BWE para a primeira parte, e executar, pelo módulo BWE que compreende a frequência de cruzamento controlável, um algoritmo de extensão de largura de banda usando o segundo decodificado sinal e o parâmetro de extensão de largura de banda para a

segunda parte,

em que uma primeira frequência de crossover (fx1) é usada para a extensão de largura de banda para o primeiro sinal decodificado e uma segunda frequência de crossover (fx2) é usada para a extensão de largura de banda para o segundo sinal decodificado, em que a primeira frequência de crossover é mais alta do que a segunda frequência de crossover; e

controlar a frequência de cruzamento para o módulo BWE de acordo com as informações do modo de codificação.

12. Método para codificar um sinal de áudio (105), caracterizado pelo fato de que compreende:

codificação de acordo com um primeiro algoritmo de codificação, o primeiro algoritmo de codificação compreendendo uma primeira largura de banda de frequência, em que a codificação de acordo com um primeiro algoritmo de codificação compreende o uso de um codificador baseado em LPC;

codificação de acordo com um segundo algoritmo de codificação, o segundo algoritmo de codificação compreendendo uma segunda largura de banda de frequência sendo menor do que a primeira largura de banda de frequência, em que a codificação de acordo com um segundo algoritmo de codificação compreende o uso de um codificador baseado em transformação;

indicando o primeiro algoritmo de codificação para uma primeira parte (204a) do sinal de

áudio e o segundo algoritmo de codificação para uma segunda parte (204b) do sinal de áudio, a segunda parte sendo diferente da primeira parte; e

cálculo de parâmetros BWE (106) para o sinal de áudio de modo que os parâmetros BWE sejam calculados para uma banda que não compreende a largura de banda de primeira frequência na primeira parte do sinal de áudio e para uma banda que não compreende a largura de banda de segunda frequência na segunda parte do sinal de áudio,

em que a primeira ou a segunda largura de banda de frequência é definida por uma frequência de crossover variável ( $f_x$ ),

em que o módulo BWE é configurado para usar uma primeira frequência ( $f_{x1}$ ) de cruzamento para calcular os parâmetros BWE para um sinal codificado usando o codificador baseado em LPC e para usar uma segunda frequência de cruzamento ( $f_{x2}$ ) para um sinal codificado usando o codificador baseado em transformação, em que a primeira frequência de cruzamento é maior que a segunda frequência de crossover.

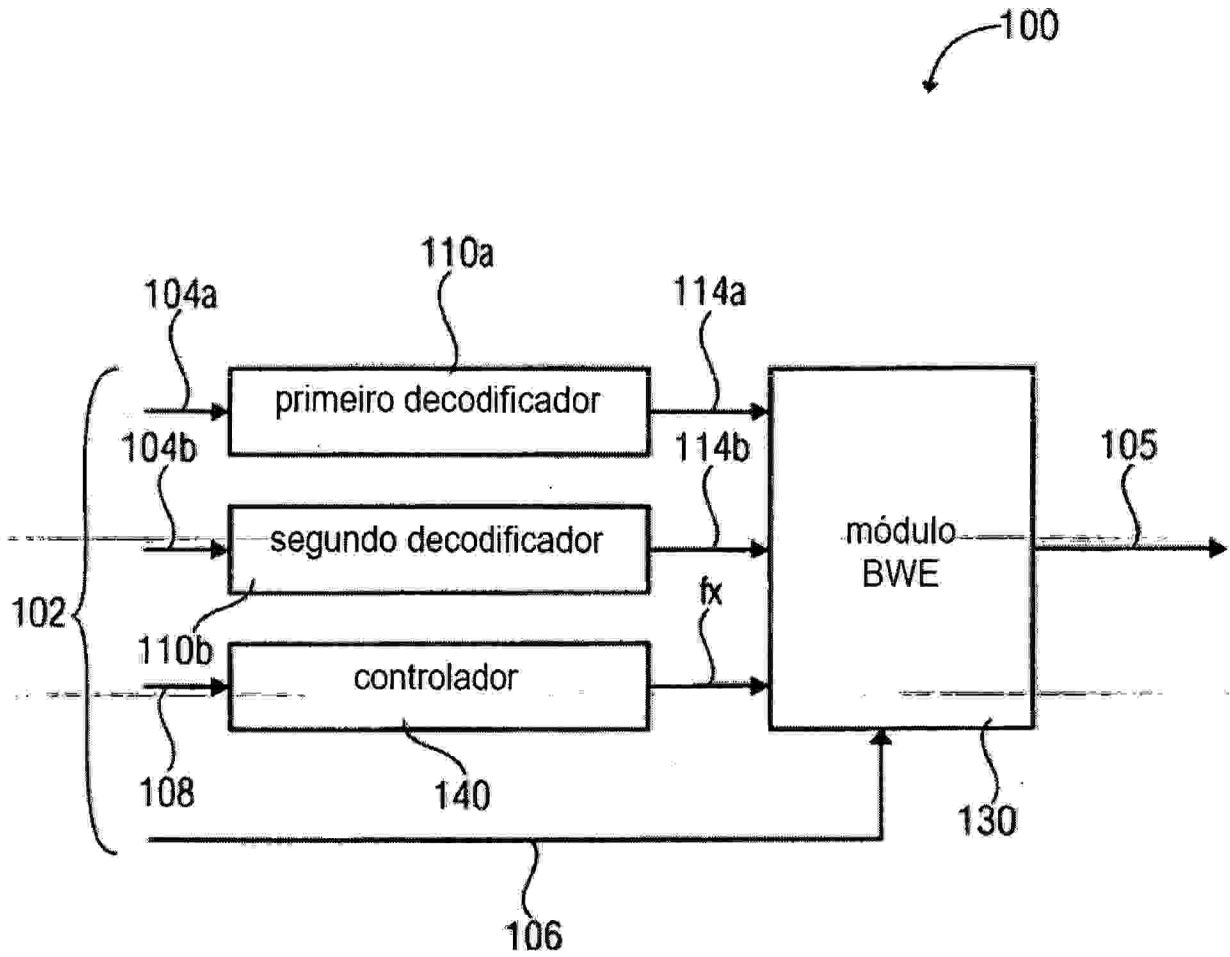


FIG 1

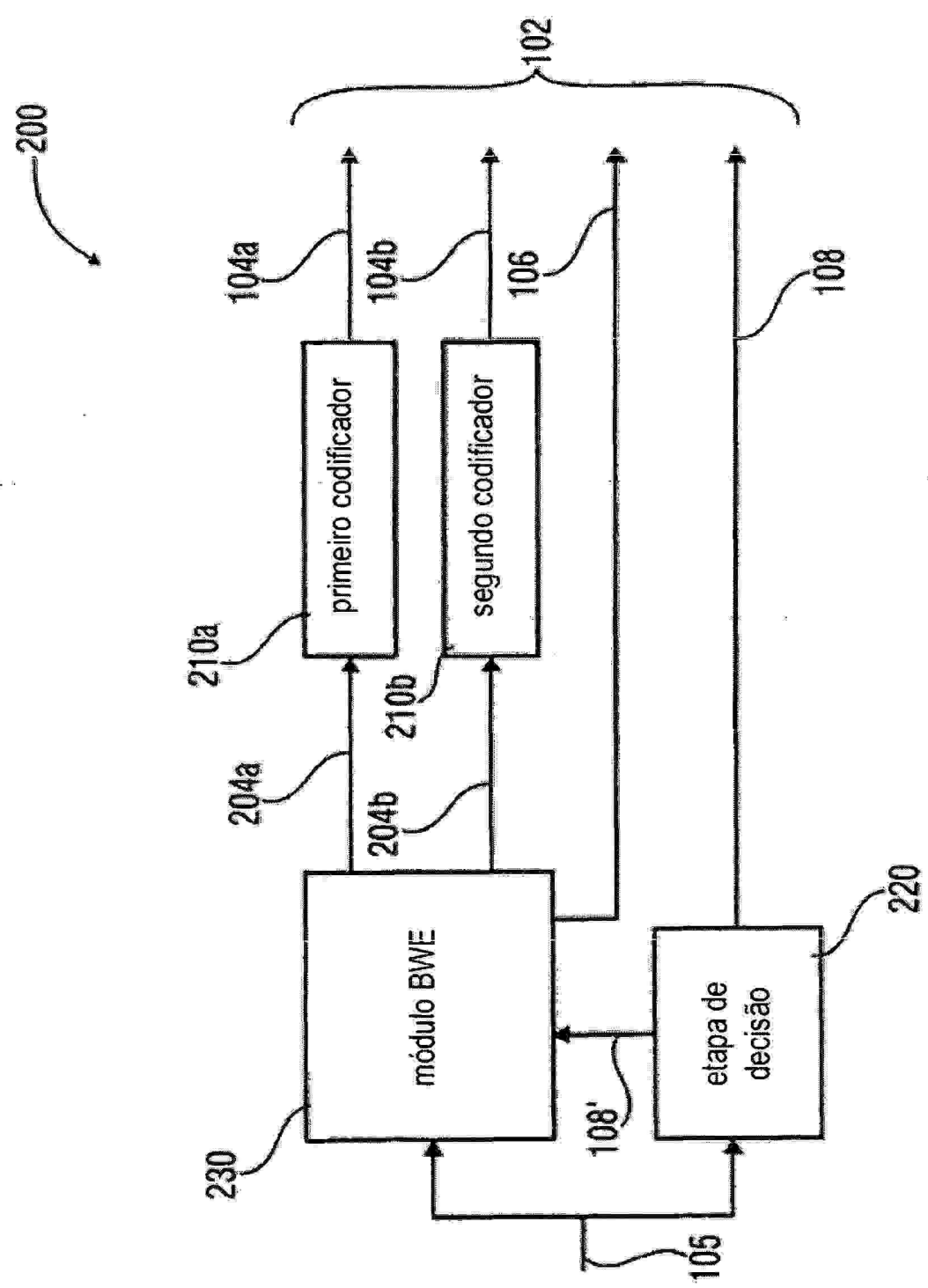


FIG 2

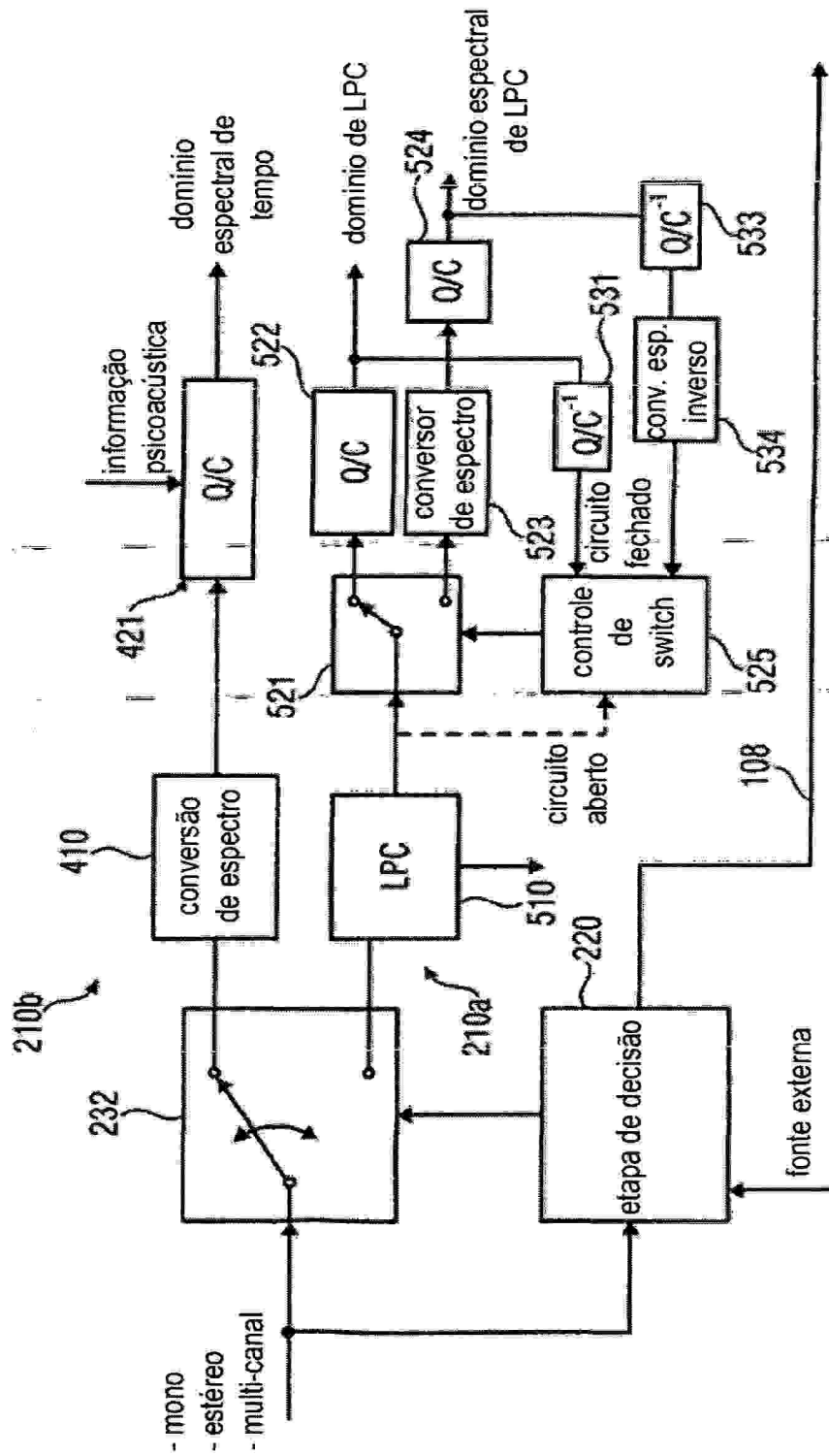


FIG 3  
(CODIFICADOR)

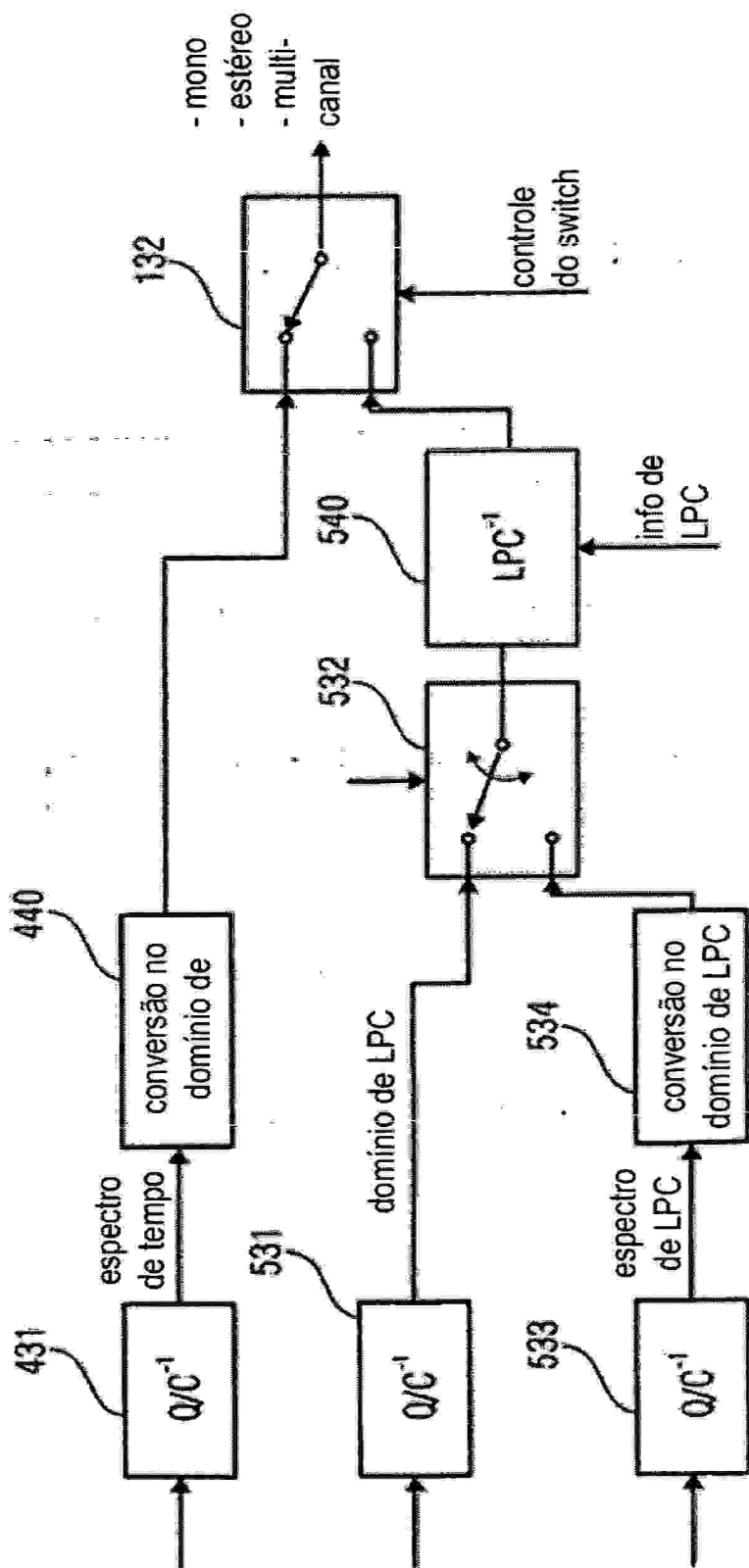


FIG 4  
(DECODIFICADOR)

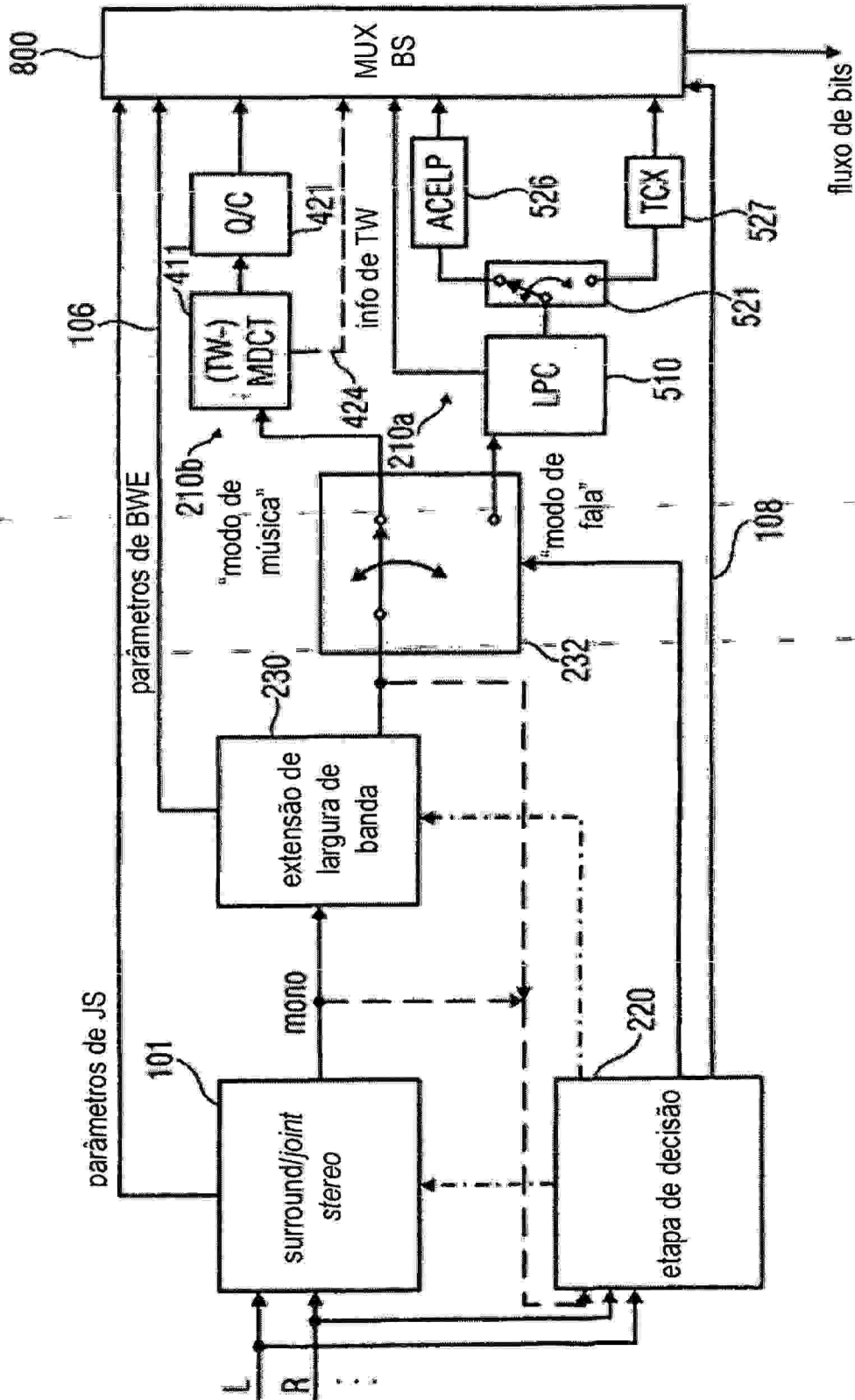


FIG 5  
(CODIFICADOR)

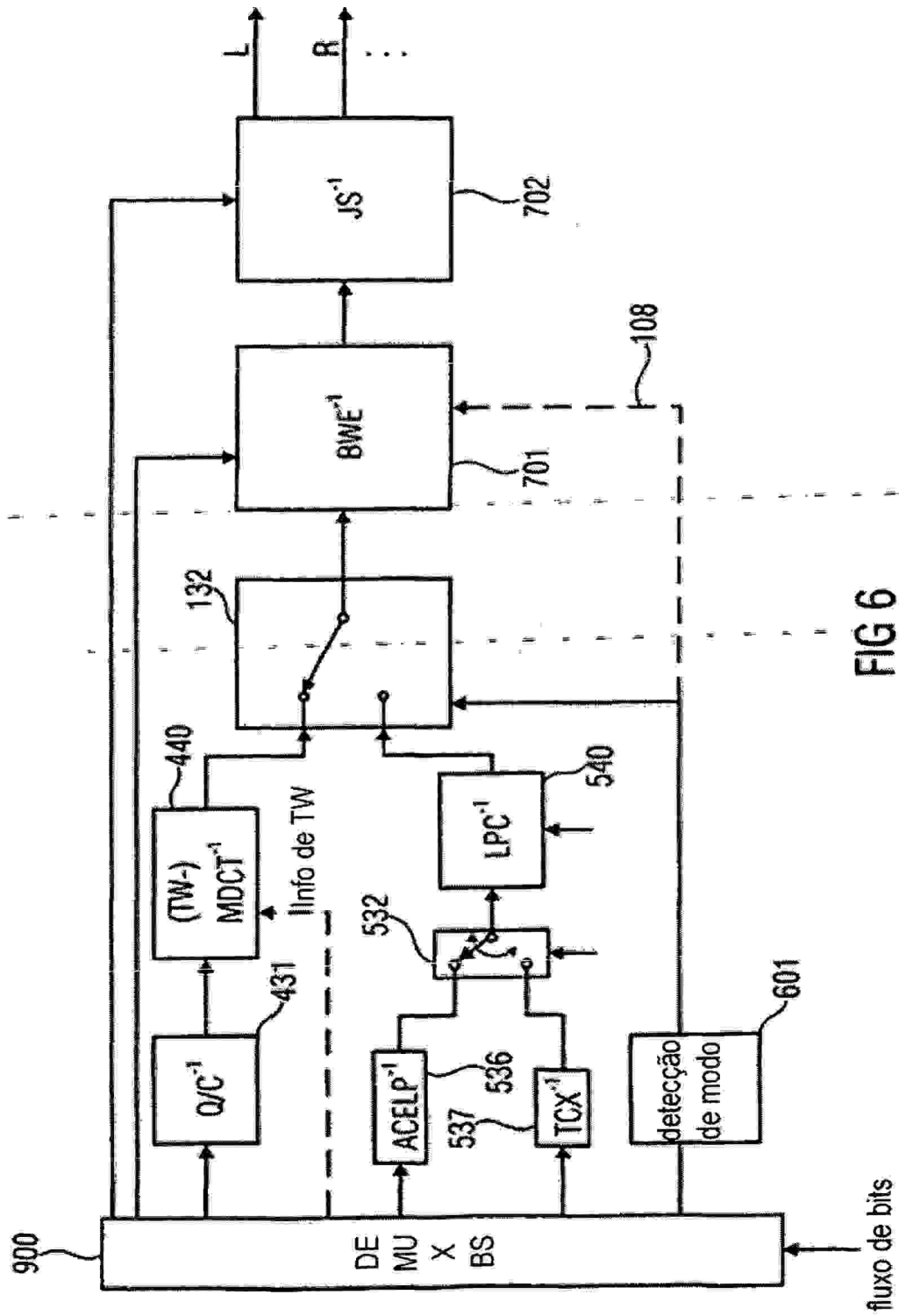


FIG 6  
(DECODIFICADOR)

fluxo de bits

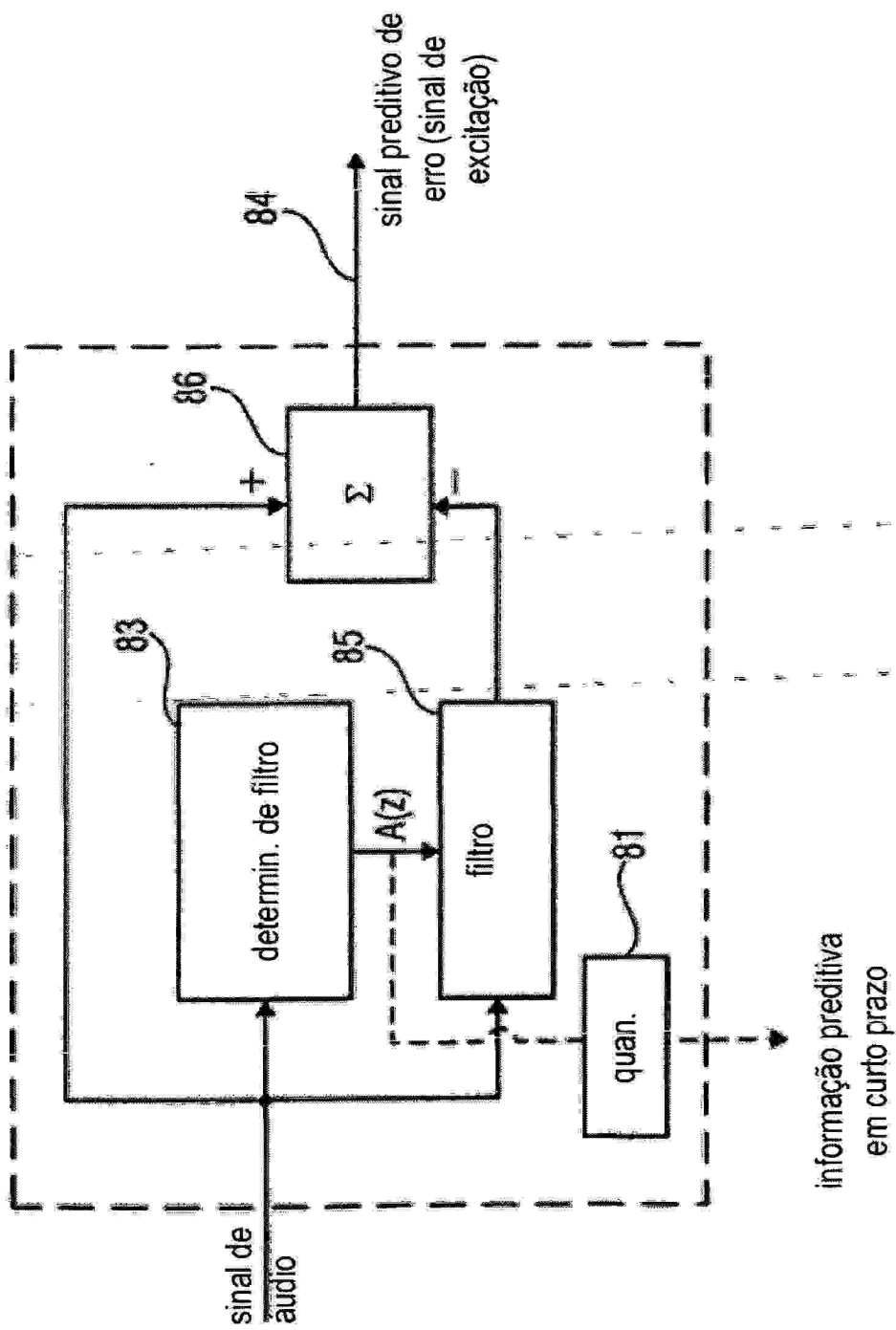


FIG 7

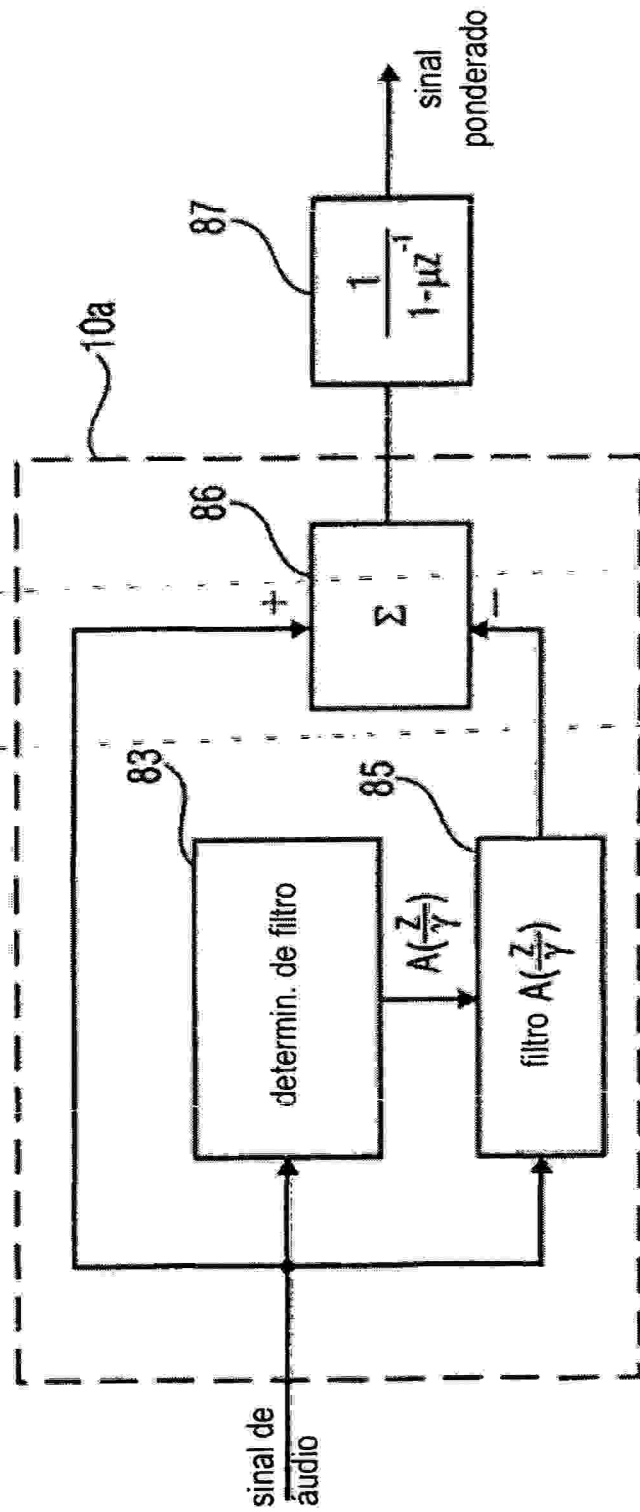


FIG 8

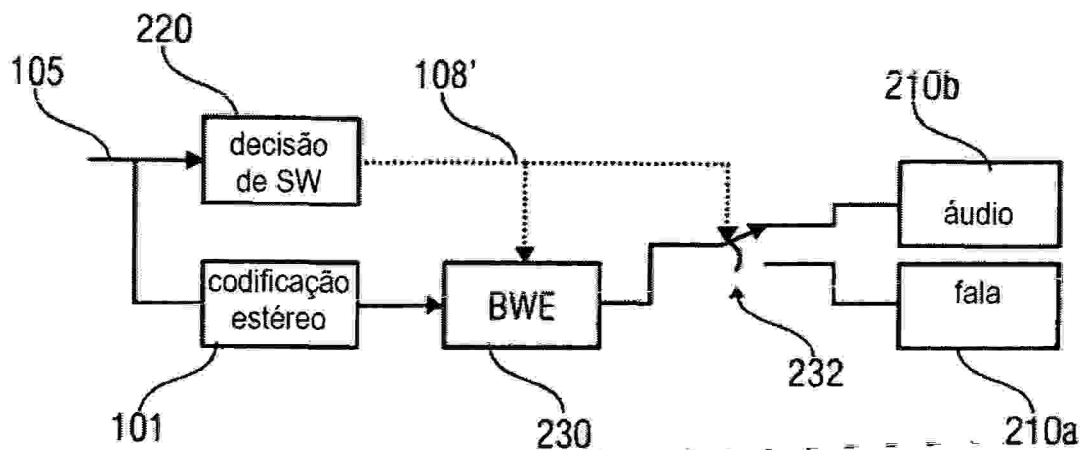


FIG 9A

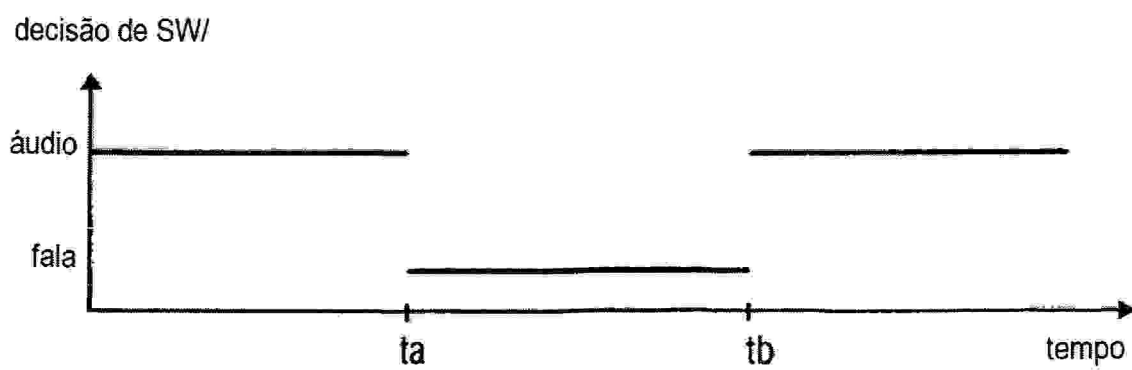


FIG 9B

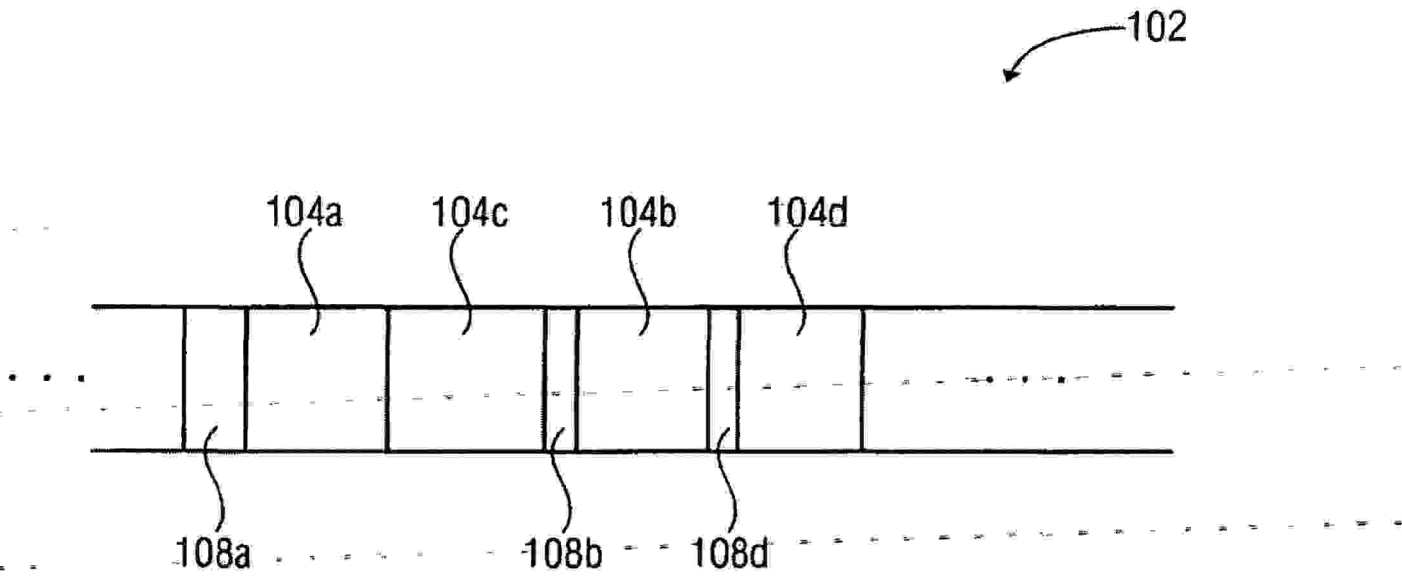


FIG 10