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

(21) PI 0910811-4 A2

(22) Data do Depósito: 25/06/2009

(43) Data da Publicação Nacional: 03/11/2020

(54) Título: CODIFICADOR DE ÁUDIO, DECODIFICADOR DE ÁUDIO, MÉTODOS PARA CODIFICAR E DECODIFICAR UM SINAL DE ÁUDIO, FLUXO DE ÁUDIO DE PROGRAMA DE COMPUTADOR

(51) Int. Cl.: G10L 19/00; G10L 19/02.

(30) Prioridade Unionista: 08/10/2008 US 61/103,820; 11/07/2008 US 61/079,872.

(71) Depositante(es): FRAUNHOFER-GESELLSCHAFT ZUR FÖRDERUNG DER ANGEWANDTEN FORSCHUG E.V..

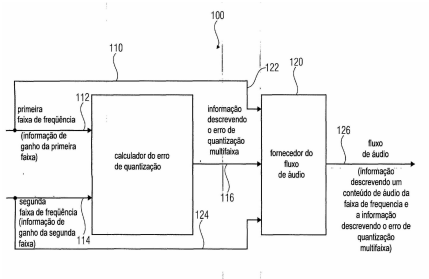
(72) Inventor(es): RETTELBACH, NIKOLAUS; GRILL, BERNHARD; FUCHS, GUILLAUME; MULTRUS, MARKUS; POPP, HARALD; HERRE, JUERGEN; WABNIK, STEFAN.; SCHULLER, GERALD; HIRSCHFELD, JENS; GEYRSBERGER, STEFAN..

(86) Pedido PCT: PCT EP2009004602 de 25/06/2009

(87) Publicação PCT: WO 2010/003556 de 14/01/2010

(85) Data da Fase Nacional: 11/01/2011

(57) Resumo: CODIFICADOR DE ÁUDIO, DECODIFICADOR DE ÁUDIO, MÉTODOS PARA CODIFICAÇÃO E DECODIFICAÇÃO DE 'UM SINAL DE ÁUDIO, E UM PROGRAMA DE COMPUTADOR. Um codificador para fornecer um fluxo de áudio com base em uma representação de domínio transformada de um sinal de áudio de entrada inclui uma calculadora de erro de quantização configurada para determinar um erro de quantização multi-banda em uma pluralidade de faixas de frequência do sinal de áudio de entrada para qual a banda informações de ganho de banda separada está disponível. O codificador inclui também um fornecedor de fluxo de áudio configurado para fornecer o fluxo de áudio de forma que o fluxo de áudio é "composto por uma"informação que descreve um conteúdo de áudio das faixas de frequências e uma informação que descreve q erro de quantização multi-banda. Um decodificador para fornecer uma representação decodificada de um sinal de áudio a partir de um fluxo de áudio codificado que representa os componentes espectrais de bandas de frequência do sinal de áudio que compreende um preenchedor de ruído configurado para introduzir ruído em componentes espectrais de uma pluralidade de faixas de frequência nos quais a informação de ganho à e frequência separada é associada (...).



"CODIFICADOR DE ÁUDIO, DECODIFICADOR DE ÁUDIO, MÉTODOS PARA CODIFICAR E DECODIFICAR UM SINAL DE ÁUDIO, FLUXO DE ÁUDIO DE PROGRAMA DE COMPUTADOR"

HISTÓRICO DA INVENÇÃO

As materializações de acordo com a invenção são relacionadas a um codificador para fornecer um fluxo de áudio com base em um domínio de transformada de uma saída de sinal de áudio. Materializações adicionais de acordo com a invenção são relacionadas a um decodificador para fornecer uma representação decodificada de um sinal de áudio com base em um fluxo de áudio codificado. Materializações adicionais de acordo com a invenção fornecem métodos para codificar um sinal de áudio e para decodificar um sinal de áudio. Materializações adicionais de acordo com a invenção fornecer um fluxo de áudio. Materializações adicionais de acordo com a invenção fornecem programas de computador para codificar um sinal de áudio e decodificar um sinal de áudio.

Em geral, as materializações de acordo com a invenção são relacionadas ao preenchedor de ruído.

O conceito de codificação de áudio frequentemente codifica um sinal de áudio no domínio da frequência. Por exemplo, o conhecido conceito de "codificação avançada de áudio" (AAC) codifica o conteúdo de bins espectrais (ou bins de frequência), considerando um modelo psicoacústico. Para este propósito, a informação de intensidade de diferentes bins espectrais é codificada. Entretanto, a resolução usada para as intensidades de codificação em diferentes bins espectrais é adaptada de acordo com as relevâncias psicoacústicas dos diferentes bins espectrais.

Assim, alguns bins espectrais, considerados como sendo de baixa relevância psicoacústica, são codificados com uma resolução de intensidade muito baixa, tal que alguns dos bins espectrais considerados como sendo de baixa relevância psicoacústica, ou mesmo um devido número dominante são quantizados para zero. Quantizando a intensidade de um bin espectral para zero traz a vantagem que o valor-zero quantizado pode ser codificado de um modo que economize muitos bits, o que ajuda a manter a menor taxa de bits possível. Apesar disso, os bins espectrais quantizados para zero resultam algumas vezes em artefatos audíveis, mesmo que o modelo psicoacústico indique que os bins espectrais são de baixa relevância psicoacústica.

Portanto, existe um desejo de trabalhar com bins espectrais quantizado para zero, ambos em um codificador de áudio e um decodificador de áudio.

Abordagens diferentes são conhecidas por trabalhar com bins espectrais codificados para zero em domínio de transformada de sistemas de codificação de áudio e também em codificadores de fala.

Por exemplo, o MPEG-4 "AAC" (codificação avançada de áudio) usa o conceito de substituição perceptível de ruído (PNS). A substituição perceptível de ruído preenche uma faixa do fator de escala somente com ruído. Os detalhes em relação ao MPEG-4 AAC podem, por exemplo, serem encontradas no Padrão Internacional ISO/IEC 14496-3 (Tecnologia de Informação - Codificação de Objetos Áudio Visuais - Parte 3: Áudio). Além disso, o codificador de fala AMR-WB+ substitui o vetor de quantização de vetores (vetores VQ) quantizados para zero com um

vetor de ruído aleatório, onde cada valor espectral complexo possui uma amplitude constante, mas uma fase aleatória. A amplitude é controlada por um valor de ruído transmitido com o fluxo de bits. Os detalhes relacionados com o codificador de fala AMR-WB+ podem, por exemplo, ser encontrados na especificação técnica intitulada "Third Generation Partnership Project; Technical Specification Group Services and System Aspects; Áudio Codec Processing Functions; Extended Adaptive Multi-Rate-Wide Band (AMR-WB+) Codec; Transcoding Functions (Release Six)", que também é conhecida como "3GPP TS 26.290 V6.3.0 (2005-06) - Technical Specification".

Além disso, o EP 1 395 980 B1 descreve um conceito de codificação de áudio. A publicação descreve um meio no qual as faixas de frequência de informação de um sinal original de áudio, que são audíveis, mas são perceptivelmente menos relevantes, não precisam ser codificados, mas podem ser substituídos por um parâmetro de preenchimento de ruído. Essas faixas de sinal tendo conteúdo, o que é perceptivelmente mais relevante são, em contraste, totalmente codificadas. Os bits de codificação são armazenados desta maneira sem deixar espaços no espectro da frequência do sinal recebido. O parâmetro de preenchimento de ruído, é uma medida do valor do sinal RMS dentro da faixa em questão e é usada no extremidade do ganho por um algoritmo de decodificação para indiciar a quantidade de ruído para injetar na faixa de frequência em questão.

Abordagens adicionais são necessárias para uma inserção não guiada no decodificador, considerando a tonalidade do espectro transmitido.

Entretanto, os conceitos convencionais trazem tipicamente o problema de que tanto consistem em uma baixa resolução com relação a granularidade do preenchedor de ruído, que tipicamente reduz a impressão de audição, ou necessitam de uma grande quantidade de informação adicional que necessita de uma taxa extra de bits.

Em vista do mencionado, há uma necessidade de um conceito melhorado do preenchedor de ruído, que proporciona uma troca melhorada entre a impressão de audição alcançável e a taxa de bit necessária.

RESUMO DA INVENÇÃO

Uma materialização de acordo com a invenção cria um codificador para fornecer um fluxo de áudio na base da representação do domínio da transformada de uma entrada de sinal de áudio. O codificador compreende um erro de cálculo de quantização configurado para determinar um erro quantização multi-faixa sobre uma pluralidade de faixas de frequência (por exemplo, sobre a pluralidade de faixas de fator de escala) da entrada do sinal de áudio, para ganho de informação de cada faixa separada (por exemplo, fatores de escala separados) está disponível. O codificador também inclui um fornecedor de fluxo de áudio configurado para fornecer o fluxo de áudio tal que o fluxo de áudio contenha uma informação descrevendo o conteúdo de áudio das faixas de frequência e uma informação descrevendo um erro de quantização da multi-faixa.

O codificador mencionado acima é baseado na descoberta do uso de uma informação de erro da quantização da multi-faixa traz a possibilidade de obter uma boa impressão de

audição na base de uma informação adicional comparativamente pequena. Em particular, o uso de uma informação de erro da quantização da multi-faixa, que cobre a pluralidade das faixas de frequência pelo qual uma faixa separada recebe informação está disponível, permite codificação adicional de escala de valores de ruído, que são baseados no erro de quantização da multi-faixa, em dependência do ganho de informação de cada faixa está disponível. Portanto, enquanto o ganho de informação da faixa está tipicamente correlacionado com a relevância psicoacústica das faixas de frequência ou com a precisão da quantização aplicada as faixas de frequência, a informação de erro da quantização foi identificada como uma informação adicional, o que permite uma síntese de preenchedor de ruído que fornece uma boa impressão de audição na baixa taxa de custo de bit da informação adicional.

Em uma materialização preferencial, o codificador compreende um quantizador configurado para quantizar componentes espectrais (por exemplo, os coeficientes espectrais) de frequências de faixas diferentes da representação do domínio de transformada usando diferentes precisões de quantização em dependência das relevâncias psicoacústicas das diferentes frequências de faixa para obter os componentes espectrais quantizados, onde as diferentes precisões de quantização são refletidas pelo ganho de informação da faixa. Também, o fornecedor de fluxo de está configurado para fornecer o fluxo de áudio tal que o fluxo de áudio inclui uma informação descrevendo o ganho de informação da faixa (por exemplo, na forma de fatores de escala) de tal modo que o fluxo de áudio também inclui a informação que descreve o erro de quantização da multi-faixa.

Em uma materialização preferencial, o calculador de erro de quantização é configurado para determinar o erro de quantização no domínio quantizado, tal que uma escala, em dependência do ganho de informação da faixa do componente espectral, que é realizado anteriormente a uma quantização inteira do valor, é levada em consideração. Ao considerar o erro de quantização no domínio quantizado a relevância psicoacústica dos bis do espectro é considerada ao calcular o erro de quantização da multi-faixa. Por exemplo, para as frequências de faixa de uma pequena relevância perceptível, a quantização pode ser grosseira, tal que um erro de quantização absoluto (no domínio não quantizado) é grande. Consequentemente, para faixas de espectro de alta relevância psicoacústica, a quantização é boa e o erro de quantização, no domínio não quantizado é pequeno. Para fazer com que os erros de quantização nas faixas de frequência de alta relevância psicoacústica e de baixa relevância psicoacústica comparáveis, para que se obtenha uma informação de erro da multi-faixa significativo, o erro de quantização é calculado no domínio quantizado (ao invés do domínio não quantizado) em uma materialização preferencial.

Em uma materialização preferencial adicional, o codificador é configurado para ajustar um ganho de informação da faixa (por exemplo, um fator de escala) de uma frequência de faixa, que é quantizada para zero (por exemplo, no qual todos os bons do espectro da faixa de frequência são quantizados para zero) para um valor representando um índice entre uma energia da faixa de frequência quantizada para zero e uma energia do erro de quantização da multi-faixa. Ao ajustar o fator de escala de uma

faixa de frequência que é quantizada para zero para um valor bem definido, é possível preencher a faixa quantizada para zero com ruído, tal que a energia do ruído seja pelo menos aproximadamente igual a energia do sinal original da faixa de frequência quantizada para zero. Ao adaptar o fator de escala no codificador, um decodificador pode tratar a faixa de frequência quantizada para zero do mesmo modo que quaisquer outras frequências de faixa não quantizadas para zero, tal que não há necessidade de uma manipulação complicada de exceção (exigindo tipicamente uma sinalização adicional). Preferencialmente, adaptando o ganho de informação de faixa (ex. fator de escala), uma combinação de valor de ganho de faixa e a informação de erro de quantização da multi-faixa permite uma determinação conveniente do preenchedor de ruído.

Em uma materialização preferencial, o calculador de erro de quantização é configurado para determinar o erro de quantização da multi-faixa sobre um pluralidade de frequências de faixa incluindo pelo menos um componente de frequência (ex. bin de frequência) quantizado para um valor não zero enquanto evita faixa de frequência totalmente quantizada para zero. Descobriu-se que uma informação de erro de quantização de multi-faixa é significativa se as faixas de frequência totalmente quantizadas para zero são omitidas do cálculo. Nas faixas de frequência totalmente quantizadas para zero, a quantização normalmente é muito grosseira, para que a informação de erro de quantização obtido de tal faixa de frequência não é tipicamente significativa. Preferencialmente, o erro de quantização em faixas de frequência mais relevantes psicoacústicamente, que não são totalmente

quantizadas para zero, fornece uma informação mais significativa, o que permite que um preenchedor de ruído adaptado à audição humana na lateral do decodificador.

Uma materialização de acordo com a invenção cria um decodificador para fornecer uma representação decodificada de um sinal de áudio na base do fluxo codificado representando os componentes do espectro das faixas de frequência do sinal de áudio. O decodificador inclui um preenchedor de ruído configurado para introduzir ruído nos componentes do espectro (por exemplo, linhas de valores de espectro ou, geralmente, os valores de bin espectral) de uma pluralidade de faixas de frequência para as quais a faixa de frequência separada ganha informação (Por exemplo fatores de escala) está associada com a Base de um valor de intensidade de ruído da multi-faixa.

O decodificador é baseado na descoberta que um único valor de intensidade de ruído da multi-faixa pode ser aplicado para um preenchedor de ruído com bons resultados caso o ganho de informação de faixa separa esteja associado com as diferentes faixas de frequência. Consequentemente, uma escala individual de ruído introduzida em diferentes faixas de frequência é possível na base no ganho de informação da faixa de frequência, tal que, por exemplo, um único valor de intensidade de ruído da multi-faixa proporciona, quando combinado com ganho de informação de faixas de frequência separadas, informação suficiente para introduzir ruído de um modo adaptado a psicoacústica humana. Assim, o conceito descrito acima, permite que se aplique um preenchedor de ruído no domínio quantizado (mas não re-escalado). O ruído adicionado no decodificador pode ser escalado com uma

relevância psicoacústica da faixa sem necessitar de informações adicionais (além da informação adicional, que é, de modo geral, necessária para a escala do conteúdo de áudio sem ruído das faixas de frequência de acordo com a relevância psicoacústica das faixas de frequência).

Em uma materialização preferencial, o preenchedor de ruído é configurado para decidir seletivamente com base por bin de espectro ao invés de introduzir um ruído em bins espectrais individuais de uma faixa de frequência em dependência de possibilidade de introduzir ruído em bins espectrais individuais de uma faixa de frequência em dependência de possibilidade de um respectivo bin espectral individual forem quantizados para zero ou não. Consequentemente, é possível obter uma ótima granularidade do preenchedor do ruído enquanto se mantém a quantidade de informação adicional muito pequena. Realmente, não é necessário transmitir qualquer informação adicional de preenchedor de ruído de faixa de frequência específica, enquanto ainda possui uma granularidade excelente em relação ao preenchedor de ruído. Por exemplo, normalmente é necessário transmitir um fator de ganho de faixa (ex. um fator de escala) para uma faixa de frequência mesmo que somente uma única linha do espectro (ou uma única bin espectral) da faixa de frequência mencionada é quantizada para um valor de intensidade não zero. Assim, como pode ser dito que a informação do fator de escala está disponível para preenchedor de ruído sem custo extra (em termos de taxa de bit) se ao menos uma linha do espectro (ou bin espectral) da faixa de frequência for quantizada para uma intensidade não zero. Entretanto, de acordo com uma descoberta da invenção presente, não é necessário transportar uma

informação de ruído de faixa de frequência específica para obter um preenchedor adequado em tal faixa de frequência na qual pelo menos um valor não zero de intensidade do bin espectral exista. Consequentemente, foi descoberto que bons resultados psicoacústicos podem ser obtidos usando um valor de intensidade de ruído de multi-faixa em combinação com a faixa de frequência específica de ganho de informação de faixa de frequência (ex. fator de escala). Assim, não é necessário gastar bits com uma informação de preenchedor de ruído de uma faixa de frequência específica. Consequentemente, a transmissão de um único valor de intensidade de ruído de multi-faixa é suficiente, uma vez que essa informação de preenchedor de ruído de multi-faixa pode ser combinado com o ganho de informação da faixa de frequência transmitida de qualquer modo para obter a informação de preenchedor de faixa de frequência específica bem adaptada às expectativas da audição humana.

Em outra materialização preferencial, o preenchedor de ruído é configurado para receber a pluralidade dos valores de bins espectrais representando diferentes porções de frequência de sobreposição e não-sobreposição da primeira faixa de frequência de um domínio de frequência da representação do sinal de áudio, e para receber a pluralidade dos valores de bins espectrais representando diferentes porções de frequência de sobreposição e não-sobreposição da segunda faixa de frequência de um domínio de frequência da representação do sinal de áudio. Além disso, o preenchedor de áudio é configurado para substituir um ou mais valores de bin espectral da primeira faixa de frequência da pluralidade das faixas de frequência com um valor inicial de bin

espectral, onde a magnitude do valor inicial de ruído do Bin de espectro é determinado pelo valor da intensidade de ruído da multi-faixa. Além disso, o preenchedor de ruído está configurado para substituir um ou mais valores de Bin do espectro da segunda faixa de frequência com um segundo valor de ruído do Bin espectral tendo a mesma magnitude do primeiro valor de ruído do Bin espectral. O decodificador também inclui um escalímetro configurado para medir valores de bin espectral da primeira faixa de frequência com o primeiro valor de ganho de faixa de frequência para obter os valores do bin espectral da primeira faixa de frequência em escala, e também medir os valores do bin espectral da segunda faixa de frequência com um Segundo valor de ganho de faixa de frequência para obter os valores do bin espectral da segunda faixa de frequência, tal que os valores de bin espectral substituídos, substituídos com os valores do primeiro e segundo bin espectral, são em escala com diferentes valores de ganho de faixa de frequência, e tal que o valor do Bin espectral substituído, com o valor do bin espectral substituído, valores dos bins espectrais não substituídos da primeira faixa de frequência representando um conteúdo de áudio de uma primeira faixa de frequência são em escala com o valor de ganho da faixa de frequência, e tal que o valor do bin espectral substituído, pelo com o valor do bin espectral substituído, valores dos bins espectrais não substituídos da segunda faixa de frequência representando um conteúdo de áudio de uma segunda faixa de frequência são em escala com o valor de ganho da faixa de frequência.

Em uma materialização de acordo com invenção, o

preenchedor de ruído é opcionalmente configurado para modificar seletivamente o valor de ganho da faixa de frequência de uma determinada faixa de frequência usando o valor do offset do ruído caso a faixa de frequência seja quantizada para zero. Consequentemente, o offset do ruído serve para minimizar uma quantidade de bits de informação adicional, deve observado que a codificação dos fatores de escala (scf) em um codificado de áudio AAC é realizada usando uma codificação Huffmann da diferença de fatores de escala subsequentes (scf). As pequenas diferenças obtêm os códigos menores (enquanto diferenças maiores obtêm códigos maiores). O offset do ruído minimize a "diferença média" em uma transição de fatores de escala convencionais (fatores de escala de faixas não quantizadas para zero) para fatores de escala de ruído e retorno, e assim otimiza a demanda de bits para informações adicionais. Isto acontece devido ao fato de que os "fatores de escala de ruído" normalmente são maiores que os fatores de escala de ruído, uma vez que as linhas incluídas não são >= 1, mas correspondem ao erro de quantização média e (onde tipicamente $0<e<0.5$).

Em uma materialização preferencial, o preenchedor de ruído é configurado para substituir os valores do bin espectral dos bins de espectro quantizados para zero com valores de ruído dos bins espectrais, as magnitudes nas quais os valores de ruído dos bins espectrais são dependentes de um valor de intensidade de ruído de multi-faixa, para obter os valores do Bin o espectro substituídos, somente para faixas de frequência com o menor coeficiente de Bin espectral acima de um índice de Bin de espectro pré-determinado, deixando os valores de bin espectral das faixas

de frequência com o menor coeficiente de Bin espectral acima de um índice de Bin de espectro pré-determinado não afetado. Além disso, o preenchedor de ruído é preferivelmente configurado para modificar seletivamente, para as faixas de frequência com o menor coeficiente de bin espectral, acima do índice de bin do espectro pré-determinado, um valor de ganho de faixa (ex. um valor de fator de escala) para uma determinada faixa de frequência em dependência de um valor de offset de ruído, caso a determinada faixa de frequência seja totalmente quantizada para zero. Preferencialmente, o preenchedor de ruído é somente realizado acima de um índice pré determinado de Bin espectral. Também, o offset de ruído é preferencialmente somente aplicado a faixas quantizadas para zero, e preferencialmente não aplicadas abaixo de um índice pré-determinado de bin espectral. Além disso, o decodificador inclui preferencialmente um escalímetro configurado para aplicar valores de ganho de faixa modificados ou não modificados seletivamente, para valores de Bin espectral substituídos ou não substituídos, para obter informação do espectro em escala, o que representa o sinal de áudio. Usando esta abordagem, o decodificador atinge uma impressão muito equilibrada de audição, que não é severamente reduzida pelo preenchedor de ruído. O preenchedor de ruído é aplicado somente em faixas de frequências superiores (com coeficientes de bin espectral menor acima do índice pré-determinado do bin espectral), uma vez que o preenchedor de ruído nas faixas de frequência menores traria uma redução indesejável nas impressões de audição. Por outro lado, é melhor realizar o preenchedor de ruído nas faixas de frequência superiores. Deve-se observar, que em alguns casos as faixas de

frequência menores (sfb) são melhor quantizadas (que as faixas de fator de escala superiores).

Outra materialização de acordo com a invenção cria um método para fornecer um fluxo de áudio com base em uma representação do domínio da transformada na entrada do sinal de áudio.

Outra materialização de acordo com a invenção cria um método para fornecer uma representação decodificada de um sinal de áudio com base em um fluxo de áudio codificado.

Uma materialização adicional de acordo com a invenção, cria um programa de computador para a realização de um ou mais dos métodos mencionados acima.

Uma materialização adicional de acordo com a invenção cria um fluxo de áudio representando o sinal de áudio. O fluxo de áudio inclui informação do espectro, descrevendo as intensidades de componentes do espectro do sinal de áudio, onde a informação do espectro é quantizada com diferentes precisões de quantização em diferentes faixas de frequência. O fluxo de áudio também inclui uma informação de nível de ruído descrevendo um erro de quantização de multi-faixa sobre uma pluralidade de faixas de frequência, considerando as diferentes precisões de quantização. De acordo com a explicação acima, tal fluxo de áudio permite uma decodificação eficiente do conteúdo de áudio, onde se obtém uma boa troca entre uma impressão de audição e uma taxa de bit necessária.

BREVE DESCRIÇÃO DAS FIGURAS

Fig. 1 Mostra um diagrama esquemático do bloco de um codificador de acordo com uma materialização da invenção;

Fig. 2 Mostra um diagrama esquemático do bloco de um codificador de acordo com outra materialização da invenção;

Figs.3a e 3b Mostra um diagrama esquemático do bloco de uma codificação avançada de áudio (AAC) de acordo com outra materialização da invenção;

Figs. 4a e 4b Mostra uma listagem de algoritmos de pseudo códigos dos programas executados para a codificação de um sinal de áudio;

Fig. 5 Mostra um diagrama esquemático do bloco de um decodificador de acordo com uma materialização da invenção;

Fig. 6 Mostra um diagrama esquemático do bloco de um decodificador de acordo com outra materialização da invenção;

Figs. 7a e 7b Mostram um diagrama esquemático do bloco de um AAC estendido e (codificação de áudio avançada) e decodificador de acordo com uma materialização da invenção;

Fig. 8a Mostra uma representação matemática de uma quantização inversa, que pode ser realizada no decodificador estendido AAC da Fig. 7;

Fig. 8b Mostra uma listagem de um algoritmo de pseudo código para quantização inversa, que pode ser executada pelo decodificador AAC estendido da Fig. 7;

Fig. 8c Mostra um fluxograma de representação de quantização inversa;

Fig. 9 Mostra um diagrama esquemático do bloco de um preenchedor de ruído e escalímetro, que pode ser usado no decodificador AAC estendido da Fig. 7;

Fig. 10a Mostra uma representação de códigos do

pseudo programa de um algoritmo, que pode executado pelo preenchedor do ruído mostrada Fig. 7 ou pelo preenchedor de ruído mostrada na Fig. 9;

Fig. 10b Mostra uma legenda de elementos do código do pseudo programa da Fig. 10a;

Fig. 11 Mostra um fluxograma de um método, que pode ser implementado no preenchedor de ruído da Fig. 7 ou no preenchedor de ruído da Fig. 9;

Fig. 12 Mostra uma ilustração gráfica do método da Fig. 11;

Figs. 13a e 13b Mostra representações de algoritimos do código do pseudo programa, que podem ser realizados pelo preenchedor de ruído da Fig. 7 ou pelo preenchedor de ruído da Fig. 9;

Figs. 14a a 14 Mostram as representações de elementos do fluxo de bits de acordo com uma materialização da invenção; e

Fig. 15 Mostram uma representação gráfica de um fluxo de bits de acordo com outra materialização da invenção.

DESCRIÇÃO DETALHADA

1.CODIFICADOR

1.1. Codificador de acordo com a Fig. 1

A Fig. 1 mostra um diagrama esquemático do bloco de um codificador para fornecer um fluxo de áudio com base na representação do domínio da transformada de uma entrada de sinal de áudio de acordo com uma materialização da invenção.

O codificador 100 da Fig. 1 inclui um calculador de erro de quantização 110 e um fornecedor de fluxo de áudio 120.

O calculador de erro de quantização 110 é configurado para receber uma informação 112 com relação ao ganho de informação da primeira faixa de frequência, quando a informação de ganho da primeira faixa de frequência esteja disponível, e uma informação 114 a respeito de uma segunda faixa de frequência, para quando a informação de ganho da segunda faixa de frequência estiver disponível. O calcular de erro de quantização está configurado para determinar um erro de quantização da multi-faixa sobre uma pluralidade de faixas de frequência na entrada do sinal de áudio, para quando a informação de ganho de faixa de frequncia esteja disponível. Por exemplo, o calculador de erro de quantização 110 está configurado para determinar o erro de quantização da multi-faixa sobre uma primeira faixa de frequência e segunda faixa de frequência usando a informação 112 e 114. Consequentemente, o calculador de erro de quantização 110 está configurado para fornecer a informação 116 descrevendo o erro de quantização da multi-faixa ao fornecedor do fluxo de áudio 120. O fonecedor de fluxo de áudio 120 está configurado para receber também uma informação 122 descrevendo a primeira faixa de frequência e uma informação 124 descrevendo uma segunda faixa de frequência. Além disso, o fornecedor de fluxo de áudio 120 está configurado para fornecer um fluxo de áudio 126, tal que o fluxo de 126 inclua uma representação de informação 116 e também uma representação do conteúdo de áudio da primeira e segunda faixa de frequência.

Consequentemente, o codificador 100 fornece um fluxo de áudio 126, incluindo um conteúdo de informação, que permite uma decodificação eficiente do conteúdo de áudio da faixa de frequência usando um preenchedor de ruído. Em particular, o

fluxo de áudio 126 fornecido pelo codificado traz uma boa troca entre a flexibilidade de decodificação do preenchedor de ruído e a taxa de bits.

1.2. Codificador de acordo com a Fig. 2

1.2.1. Visão Geral do Codificador

No seguinte, um codificador de áudio melhorado de acordo com uma materialização da invenção será descrito, que é baseado no codificador de áudio descrito no Padrão Internacional ISO/IEC 14496-3: 2005(E), Information Technology - Coding of Áudio-Visual Objects - Part 3: Áudio, Sub-part 4: General Áudio Coding (GA) - AAC, Twin VQ, BSAC.

O codificador de áudio 200 de acordo com a Fig. 2, é especificamente baseado no codificador de áudio descrito no ISO/IEC 14496-3: 2005(E), Part 3: Áudio, Sub-part 4, Section 4.1. No entanto, o codificador de áudio 200 não precisa implementar a funcionalidade exata do codificador de áudio da ISO/IEC 14494-3: 2005(E).

O codificador de áudio 200, pode, por exemplo, ser configurado para receber um sinal de tempo de entrada 210 e fornecer, na mesma base, um fluxo de áudio codificado 212. Um sinal de caminho de processamento pode incluir um downsampler opcional 220, um controle de ganho 222 AAC opcional, um banco de filtro de comutação de bloco 224 e um processamento de sinal opcional 226, um codificador AAC estendido 228 e um formatador de carga útil de fluxo de bit 230. No entanto, o codificador 200 inclui tipicamente um modelo psicoacústico 240.

Em um caso muito simples, o codificador 200 somente inclui um banco de filtro de comutação de banco 224, o

codificador AAC estendido 228, o formatador de carga útil de fluxo de bits 230 e o modelo psicoacústico 240, enquanto os outros componentes (em particular, os componentes 220, 222 e 226) devem ser considerados como meramente opcionais.

Em um caso muito simples, o banco de filtro/ comutação de bloco 224 recebe o sinal do tempo de entrada 210. (reduzida opcionalmente pelo downsampler 220, e escalada opcionalmente pelo controlador de ganho AAC 222), e fornece, na base do mesmo, uma representação de domínio de frequência 224ª. A representação de domínio de frequência 224, pode, por exemplo, incluir uma informação descrevendo intensidades (por exemplo, amplitudes e energias) de bins espectrais do sinal de tempo de entrada 210. Por exemplo, o banco de filtro/ comutação de bloco 224, pode ser configurado para realizar uma transformada discreta de cosseno modificada (MDCT) para derivar os valores de domínio de frequência do sinal de tempo de entrada 210. A representação de domínio de frequência 224a pode logicamente se dividir em diferentes faixas de frequência, que também são designadas como "faixas de fatores de escala". Por exemplo, considera-se que o banco de filtro/ comutação de bloco 224, fornece valores do espectro (também designados como valores de bin de frequência) para um grande número de diferentes bins de frequência. A quantidade de bins de frequência é determinada, entre outras, pelo comprimento de uma entrada de janela no banco de filtro 224, e também dependentes na amostragem (e bits) e taxa. Entretanto, as faixas de frequência ou faixas de fatores de escala definem sub-conjuntos dos valores de espectro fornecidos pelo banco de filtro/ comutação de bloco. Os detalhes em relação a definição das faixas

de fator de escala são conhecidos por uma pessoa especialista em ISO/IEC 14496-3: 2005(E), Part 3, Sub-part 4.

O codificador AAC estendido 228 recebe os valores do espectro 224a fornecidos pelo banco de filtro/ comutação de bloco na base do sinal do tempo de entrada 210 (ou uma versão pré-processo) como informação de entrada 228ª. Como pode ser visto na figura 2, a informação de entrada 228a do codificador AAC estendido 228 deve derivar dos valores do espectro 224a usando uma ou mais das etapas do processamento para um processamento otimizado dos espectros 226. Para detalhes com relação as etapas de pré-processamento do processamento espectral 226, é feita uma referência na ISO/IEC 14496-3: 2005(E), e para outros Padrões mencionados.

O codificador AAC estendido 228 é configurado para receber a informação de entrada 228a na forma de valores espectrais para uma pluralidade de bins espectrais e para fornecer, na mesma base, uma representação 228b codificada, quantizada e sem ruído do espectro. Para este propósito, o codificador AAC estendido 228 pode, por exemplo usar a informação derivada de um sinal de entrada de áudio 210 (ou uma versão de pré-processamento) usando o modelo psicoacústico 240. De modo geral, o codificador 228 AAC estendido pode usar informação fornecida pelo modelo psicoacústico 240 para decidir qual a precisão a ser aplicada para a codificação de diferentes faixas de banda (ou faixas de fatores de escala) da informação de entrada espectral 228a. Assim, o codificador 228 AAC estendido pode em geral adaptar sua precisão de quantização para diferentes faixas de frequência das características específicas do sinal de tempo de

entrada 210, e também, o número de bits disponíveis. Assim, o codificador 228 AAC estendido pode, por exemplo, ajustar suas precisões de quantização, para que a informação representando o espectro codificado, quantificado e sem ruídos que inclui uma taxa de bits adequada (ou média).

O formatador de carga útil de fluxo de bits 230é configurado para incluir a informação 228b representando os espectros codificados, quantizados e sem ruídos no fluxo de áudio codificado 212 de acordo com uma sintaxe pré-determinada.

Para maiores detalhes sobre a funcionalidade dos componentes do codificador descritos aqui, há uma referência no ISO/IEC 14496-3: 2005(E) (incluindo o anexo 4.B), e também no ISO/IEC 13818-7: 2003.

Além disso, há uma referência no ISO/IEC 13818-7: 2005, Sub-cláusulas C1 a C9.

Mais além, há uma referência específica com relação a terminologia no ISO/IEC 14496-3: 2005(E), Part 3: Áudio, Sub-part 1: Main.

Além disso, há uma referência específica no ISO/IEC 14496-3: 2005(E), Part 3: Áudio, Sub-part 4: General Áudio Coding (GA) - AAC, Twin VQ, BSAC.

1.2.2. Detalhes do codificador

A seguir, serão descritos os detalhes com relação ao decodificador com referência às Figs. 3a, 3b, 4a e 4b.

As Figs. 3a e 3b mostram um diagrama esquemático do bloco de um codificador 228 AAC estendido de acordo com uma materialização da invenção. O decodificador 228 AAC estendido é designado com 228 e pode substituir o codificador 228 AAC

estendido da Fig. 2. O codificador 228 AAC estendido é configurado para receber, como informação de entrada 228, um vetor de magnitudes das linhas espectrais, onde o vetor das linhas espectrais é designado com uma mdct_line (0..1023). O codificador 228 AAC estendi também recebe uma informação de limitação de codec 228c, que descreve uma energia máxima de erro permitida em um nível MDCT. A informação de limitação do codec 228c é normalmente fornecida individualmente por diferentes faixas de fator de escala e são geradas usando o modelo 240 psicoacústico. A informação de limitação de codec 228, algumas vezes é designada com X_{min} (sb), onde o parâmetro sb indica a dependência da faixa do fator de escala. O codificador 228 AAC estendido também recebe uma informação 228d de quantidade de bits, que descreve a quantidade de bits disponíveis para a codificação do espectro representado pelo vetor 228a das magnitudes dos valores espectrais. Por exemplo, a informação de quantidade de bits 228d pode incluir uma informação média de bits (designada com mean_bits) e uma informação de bits adicionais (designada com more_bits). O codificador 228 AAC estendido também está configurado para receber a informação de faixa de fatores de escala 228e, que descreve, por exemplo, uma quantidade e largura das faixas de fator de escala.

O codificador AAC estendido, inclui um quantizador de valor espectral 310, que é configurado para fornecer um vetor 312 de valores quantizados das linhas espectrais, que também é designado com a x_quant.(0..1023). O quantizador de valor espectral 310 que inclui um escalímetro, também é configurado para fornecer informações do fator de escalas 314, que pode representar um fator de escala para cada faixa de

fator de escala e também como informação do fator de escala comum. Além disso, o quantizador de valor espectral 310 pode ser configurado para fornecer a informação de uso de bits 316, que pode descrever uma quantidade de bits usados para quantizar o vetor 316, que pode descrever uma quantidade usado para quantizar o vetor 228a com precisões diferentes dependendo da relevância psicoacústica de diferentes valores espectrais. Para este propósito, o quantizador de valores espectrais 210 escala os valores do vetor 228a usando diferentes Fatores de escala dependentes de faixas de fatores de escala e quantiza o resultado dos valores espectrais em escala. Tipicamente, os valores espectrais associados com faixas de fator de escala psicoacústicamente importantes serão em escala com grandes fatores de escala, de modo que os valores espectrais em escala de faixas de fator de escala psicoacústicamente importantes cobrem uma ampla gama de valores. Em contraste, valores espectrais de faixas de fator de escala psicoacústicamente menos importantes são escaladas com fatores de escalas menores, de modo que, os valores espectrais em escala de faixas de fator de escala psicoacústicamente menos importantes cobrem somente uma gama menor de valores. Os valores espectrais em escala são então quantizados, para um valor inteiro. Nessa quantização, muitos dos valores espectrais em escala das faixas de fator de escala psicoacústicamente menos importantes são quantizadas para zero, uma vez que os valores espectrais das faixas de fator de escala psicoacústicamente menos importantes são escaladas somente com fatores de escala menores.

Como resultado, pode se dizer que os valores espectrais de faixas de fator de escala psicoacústicamente mais

importantes são quantizadas com maior precisão (uma vez que as linhas de escalas de espectro das faixas de fatores de escala mais relevantes cobrem uma gama maior de valores e, portanto, mais etapas de quantização), enquanto os valores de faixas de fator de escala psicoacústicamente menos importantes são quantizadas com menor precisão de quantização (uma vez que os valores em escala das faixas de fator de escala psicoacústicamente menos importantes cobrem uma gama menor e são , portanto, quantizado com etapas de quantização diferente).

O quantizador de valores espectrais 310 é tipicamente configurada para determinar os fatores de escalamento usando a limitação de codec 228c e a informação de quantidade de bits 228d. Tipicamente, o quantizador de valores espectrais 310 também é configurado para determinar os fatores de escala adequados por si mesmo. Os detalhes em relação a possível implementação do quantizador de valores espectrais 310estão descritos na ISO/IEC 14496-3: 2001, Capítulo 4.B.10. Além disso, a implementação do quantizador de valores espectrais é bem conhecido para uma pessoa especializada no assunto de codificação MPEG4.

O codificador 228 AAC estendido também inclui um calculador de erro de quantização de multi-faixa 330, que é configurado para receber, por exemplo, o vetor 228A das magnitudes dos valores espectrais, o vetor 312 de valores quantizados de linhas espectrais e da informação 314 do fator de escala. O calculador de erro de quantização de multi-faixa 330 é, por exemplo, configurado para determinar o desvio entre uma versão escalada não quantizada de valores espectrais do vetor 228a (por exemplo, escalada usando uma operação de escala não linear e um

fator de escala) e uma versão em escala e quantizada (por exemplo escalada usando uma operação de escalamento não linear e um fator de escala, e quantizada usando uma operação de arredondamento "inteiro") dos valores espectrais. Além disso, o calculador de erro de quantização da multi-faixa 330 pode ser configurado para calcular uma média de erro de quantização sobre uma pluralidade de faixas de fatores de escala. Deve-se notar que o calculador de erro de quantização330 calcula preferencialmente o erro de quantização da multi-faixa em um domínio quantizado (mais precisamente um domínio psicoacústicamente escalado), de modo que o erro de quantização em faixas de fatores de escala psicoacústicamente mais relevantes é enfatizado quando comparado ao erro de quantização em faixas de fator de escala psicoacústicamente menos relevantes. Os detalhes relacionados ao calculador de erro de quantização de multi-faixa será descrito subsequentemente referindo-se as Figs. 4a e 4b.

O codificador 328 AAC estendido também inclui um adaptador de fator de escala 340, que é configurado para receber o vetor 312 de valores quantizados, a informação de fator de escala 314 e também a informação de erro de quantização da multi-faixa 332, fornecida pelo calculador de erro de quantização da multi-faixa 340. O adaptador de fator de escala 340 é configurado para identificar faixas de fatores de escala, que são "quantizados para zero", isto é, faixas de fator de escala para os quais todos os valores espectrais (ou linhas espectrais) são quantizados para zero. Para tal escala de faixas de fatores de escala quantizada totalmente para zero, o adaptador de fator de escala 340 adapta o respectivo fator de escala. Por exemplo, o adaptador de fator de

escala 340 pode ajustar o fator de escala de uma faixa de fator de escala totalmente para zero para um valor, que representa uma taxa entre uma energia residual (antes da quantização) da respectiva faixa de fator de escala e uma energia de erro de quantização da multi-faixa 332. Consequentemente o adaptador de fator de escala 342. Deve-se observar que ambos os fatores de escala fornecidos pelo quantizador de valores espectrais 310 e os fatores de escala adaptados fornecidos pelo adaptador de fator de escala são designados com "scale fator (sb)", "scf[band]", "sf[g][sfb]", "scf[g][sfb]" na literatura e também dentro de sua aplicação. Os detalhes com relação a operação do adaptador de fator de escala 340 são descritos em referência das Figs. 4a e 4b.

O codificador 228 AAC estendido também inclui uma codificação sem ruído 350, que em por exemplo, explicada no ISO/IEC 14496-3: 2001, Chapter 4.B.11. Resumindo, a codificação sem ruído 350 recebe o vetor de valores quantizados de linhas espectrais (também designadas como valores quantizados dos espectros") 312, a representação do inteiro 342 dos fatores de escala (mesmo que fornecidos pelo quantizador 310 de valores espectrais, ou adaptados pelo adaptador de fatores de escala 340), e também um parâmetro de preenchimento de ruído 332 (por exemplo, na forma de uma informação de nível) fornecido pelo calculador de erro de quantização de multi-faixa 330.

A codificação 350 sem ruído inclui um coeficiente espectral de codificação 350 a para codificar os valores quantizados 312 das linhas espectrais, e para fornecer valores quantizados e codificados 352 das linhas espectrais. Os detalhes em relação a codificação de coeficiente espectral estão, por

exemplo, descritas nas seções 4.B.11.2, 4.B.11.3, 4.B.11.4 e 4.B.11.6 do ISO/IEC 14496-3: 2001. A codificação sem ruído 350 também inclui o fator de escala de codificação 350b para codificar a representação do inteiro 342 do fator de escala para obter uma informação 354 dos fatores de escala codificados. A codificação sem ruído 350 também inclui uma codificação de parâmetro do preenchedor do ruído 350c para codificar um ou mais parâmetros de preenchedor de ruído 332, para obter um ou mais parâmetros de preenchedor de ruído codificados 356. Consequentemente, o codificador AAC estendido fornece uma informação descrevendo os espectros codificados quantizados como sem ruídos, onde esta informação inclui os valores quantizados e codificados das linhas espectrais, informação do fator de escala codificado e informação do parâmetro de preenchimento de ruído codificado.

A seguir, a funcionalidade do calculador de erro de quantização de multi-faixa 330 e do adaptador de fator de escala 340, que são componentes chave da inventiva do codificador 228 AAC estendido serão descritos, usando como referência as Figs. 4a e 4b. Para este propósito, a Fig. 4a mostra uma listagem de programa de um algoritmo realizado pelo calculador de erro de quantização 330 da multi-faixa e o adaptador de fator de escala 340.

A primeira parte do algoritmo, representada pelas linhas 1 a 12 do pseudo código da Fig. 4a, inclui um cálculo de um erro médio de quantização, que é realizado pelo calculador de erro de quantização de multi-faixa 330. O cálculo do erro médio de quantização é realizado por exemplo, sobre todas as faixas de fator de escala, exceto por aqueles que são quantizados para zero.

Caso uma faixa de fator de escala seja totalmente quantizada para zero (isto é, todas as linhas espectrais da faixa do fator de escala são quantizadas para zero), a faixa de fator de escala mencionada é trocada por um cálculo de erro médio de quantização. Se, no entanto, uma faixa de fator de escala não esteja totalmente quantizado para zero (isto é, inclui pelo menos uma linha espectral, que não é quantizada foi zero), todas as linhas espectrais da faixa de fator de escala mencionada são considerados para um cálculo do erro médio de quantização. O erro médio de quantização é calculado em um domínio quantizado (ou, mais precisamente em um domínio de escala). O cálculo de uma contribuição para o erro médio pode ser visto na linha 7 do pseudo código da Fig. 4a. Em particular, a linha 7 mostra a contribuição de uma única linha espectral para o erro médio, onde a media é realizada sobre todas as linhas espectrais (onde as linhas indicam a quantidade total de linhas consideradas).

Como pode ser visto na linha 7 do pseudo código, a contribuição de uma linha espectral ao erro médio é o valor absoluto ("fabs"- operador) de uma diferença entre a magnitude do valor de escala de uma linha espectral na quantizada e a magnitude do valor de uma linha espectral quantizada. Na magnitude do valor de escala de uma linha espectral não quantizada, a magnitude do valor da "linha" (que pode ser igual a mdtc_line) é escalada não linearmente usando uma potenciação (pow (line, 0.75) = line $^{0.75}$) e usando um fator de escala (ex. um fator de escala 314 fornecido pelo quantizador de valor espectral 310). No cálculo do valor em escala de magnitude da linha espectral da quantizada, o valor "linha" de magnitude da linha espectral pode ser escalado como não

linear usando as funções de potencia mencionadas e escalas usando o fator de escala também mencionado. O resultado deste escalamento linear e não linear pode ser quantizado usando um operador inteiro "(INT)". Usando o calculo como indicado na linha 7 do pseudo código, o impacto diferente na quantização de faixas de frequência mais e menos importantes é considerado.

Seguindo o cálculo de erro de quantização de multi-faixa (média) (avgERROR), a erro médio de quantização pode ser opcionalmente quantizado conforme mostrado nas linhas 13 e 14 do pseudo código. Deve-se observar que a quantização do erro de quantização da multi-faixa conforme mostrado aqui, é especificamente adaptado a taxa esperada de valores e características estatísticas do erro de quantização, de modo que o erro de quantização pode ser representado em um modo de bit-eficiente. Entretanto, outras quantizações da do erro de quantização da multi-faixa pode ser aplicado.

Uma Terceira parte do algoritmo, que é representado nas linhas 15 a 25, podem ser executado pelo adaptador de fator de escala 340. Uma terceira parte do algoritmo serve para ajustar os fatores de escala das faixas de frequência dos fatores de escala, que foram totalmente quantizados para zero, para um valor bem definido, que permite um preenchedor de ruído simples, que traz uma boa impressão de audição. A terceira parte do algoritmo inclui opcionalmente uma quantização inversa do nível de ruído (ex. representado pelo erro de quantização da multi-faixa 332). A terceira parte do algoritmo também inclui o calculo de uma substituição do valor de fatores de escala para faixas de fatores de escala quantizados para zero (enquanto os fatores de escala das

faixas de fatores de escala não serão afetados). Por exemplo, a substituição do valor do fator de escala por uma determinada faixa de fator de escala "(faixa)" é calculada usando a equação mostrada na linha 20 do algoritmo da figura 4A. Nesta equação "(INT)" representa o operador inteiro, o "2.f" representa o número "2" em uma representação de ponto de flutuação "log" designa o operador do logaritmo, "energy" designa uma faixa de fator de escala sob consideração (antes da quantização), "(float)" designa um operador do ponto de flutuação, "sfbWidth" designa a largura de determinadas faixas de fator de escala em termos de linha espectrais (ou bins espectrais, e ""noiseVal designa um valor de ruído descrevendo o erro de quantização da multi-faixa. Consequentemente, a substituição do fator de escala descreve uma taxa entre uma energia de bin por frequência(energy/sfbWidth) de determinadas faixas de fator de escala, sob consideração, e uma energia ($noiseVal^2$) do erro de quantização da multi-faixa.

1.2.3. Conclusão do Codificador

As materializações de acordo com invenção criam um codificador com um novo tipo de cálculo de nível de ruído. O nível de ruído é calculado com base no domínio quantizado do erro médio de quantização.

Ao calcular o erro de quantização no domínio quantizado traz vantagens significantes, por exemplo, devido a relevância psicoacústica de diferentes faixas de frequência (faixas de fatores de escala) são considerados. O erro de quantização. O erro de quantização por linha (isto é, por linha espectral, ou bin espectral) no domínio quantizado está tipicamente na faixa [-0.5; 0.5] (1 nível de quantização) com uma

média de erro absoluto de 0.25 (para valores de entrada normalmente distribuídos que são tipicamente maiores que 1). Usando um codificador, que fornece informação a respeito do erro de quantização da multi-faixa, as vantagens do preenchedor de ruído no domínio quantizado podem ser explorados em um codificador, assim como será descrito subsequentemente

O cálculo de nível de ruído e detecção de substituição de ruído no codificador pode incluir as seguintes etapas:

- Detectar e marcar faixas especiais que podem ser reproduzidas perceptivelmente equivalente no codificador por substituição de ruído. Por exemplo, a tonalidade ou uma medida de monotonia espectral deve ser verificada para este propósito;
- Calcular e quantizar o erro médio de quantização (que pode ser calculado sobre todas as faixas de fatores de escala não quantizados para zero); e
- Calcular o fator de escala (scf) para uma faixa quantizada para zero de modo que os ruídos introduzidos seja compatível com a energia original (decodificador).

Uma quantização de nível de ruído adequada pode ajudar a produzir a quantidade de bits necessária para transportar informação descrevendo o erro de quantização de multi-faixa. Por exemplo o nível de ruído pode ser quantizado em oito níveis de quantização no domínio logarítmico, levando em consideração a percepção humana de volume. Por exemplo, o algoritmo mostrado na Fig. 4b pode ser usado, onde "(INT)" designa um operador inteiro, onde "LD" designa uma operação de logaritmo para uma base de 2 e onde "meanLineError" designa o erro de quantização por linha de

frequência. "min(.,.)" designa um operador de valor minimo, e "max(.,.)" designa um operador de valor máximo.

2. Decodificador

2.1. Decodificador de acordo com a Fig. 5

A Fig. 5 mostra um diagrama esquemático do bloco de um decodificador de acordo com uma materialização da invenção. O decodificador 500 é configurado para receber uma informação de áudio codificada, por exemplo, na forma de um fluxo de áudio codificado 510, e para fornecer, na base do mesmo, uma representação decodificada do sinal de áudio, por exemplo, com base nos componentes espectrais 522 de uma primeira faixa de frequência e componentes espectrais 524 de uma segunda faixa de frequência. O decodificador 500 inclui um preenchedor de ruído 520, que é configurado para receber uma representação 522 de componentes espectrais de uma primeira faixa de frequência, para um ganho de informação de primeira faixa de frequência é associado, e uma representação 524 de componentes espectrais de uma segunda faixa de frequência, para o qual está associado um ganho de informação de segunda faixa de sequência. Além disso, um preenchedor de ruído 520 é configurado para receber uma representação 526 de um valor de intensidade de ruído da multi-faixa. Além disso o preenchedor de ruído é configurado para introduzir ruído nos componentes espectrais (ex. em valores de linhas espectrais ou valores de bins espectrais) de uma pluralidade de faixas de frequência para que o ganho de informação de faixa de frequência separada (por exemplo , na forma de fatores de escala) esteja associado na base do valor de intensidade de ruído de multi-faixa comum 526. Por exemplo, o preenchedor de

ruído 520 é configurado para introduzir ruído nos componentes espectrais 522 da primeira faixa de frequência para obter os componentes espectrais afetados pelo ruído da primeira faixa de frequência, assim como também, para introduzir ruído em componentes espectrais 524 de uma segunda faixa de frequência para obter os componentes espectrais 524 afetados pelo ruído da segunda faixa de frequência.

Ao aplicar o ruído descrito por um único valor de intensidade de ruído de multi-faixa 526 para componentes espectrais de diferentes faixas de frequência nas quais estejam associados diferentes ganhos de informação de faixa de frequência, o ruído pode ser introduzido de maneira muito sintonizada, considerando as diferentes relevâncias psicoacústicas de diferentes faixas de frequência, que é expressada pelo ganho de informação de faixa de frequência. Assim, o decodificador 500 é capaz de realizar um preenchedor de ruído em sintonia com o tempo com base em uma informação adicional muito pequena (bit-efficiency) de preenchedor de ruído.

2.2. Decodificador de acordo com a Fig. 6

2.2.1. Visão Geral do Decodificador

A Fig. 6 mostra um diagrama esquemático do bloco de um decodificador 600 de acordo com uma materialização da invenção.

O decodificador 600 é similar ao decodificador divulgado no ISO/IEC 14496.3: 2005 (E), de modo que a referência é feita a esse Padrão Internacional. O decodificador 600 é configurado para receber um fluxo de áudio codificado 610 e para fornecer, na base de, sinais de tempo de saída 612. O fluxo de

áudio codificado pode incluir algumas ou todas as informações descritas no ISO/IEC 14496.3: 2005 (E), e adicionalmente, compreende informações que descrevem um valor de intensidade de ruído da multi-faixa. O decodificador 600 inclui ainda uma carga de fluxo de bits desformatador 620, que é configurado para extrair do fluxo de áudio codificado 610 uma pluralidade de parâmetros de codificação de áudio, alguns dos quais serão explicados em detalhes a seguir. O decodificador 600 inclui ainda uma "codificação avançada de áudio" estendido (AAC) decodificador 630, a funcionalidade de que serão descritos em detalhes, em referência às Figs. 7a, 7b, 8a a 8c, 9, 10a, 10b, 11, 12, 13a e 13b. O decodificador AAC estendido 630 é configurado para receber uma informação de entrada 630a, que inclui, por exemplo, uma informação de linha espectral quantizada e codificada, uma informação de fator de escala codificado e uma informação de parâmetro de preenchimento de ruído codificado. Por exemplo, a entrada de informação 630A do codificador AAC estendido 630 pode ser idêntico às informações de saída 228b fornecidas pelo Codificador 220a AAC estendido descritos referentes à figura. 2.

O decodificador 630 AAC estendido pode ser configurado para fornecer, a partir da entrada de informação 630a, uma representação 630B de espectros escalados e quantizados inversamente, por exemplo, sob a forma de valores de linha espectral quantizados inversamente de uma pluralidade de bins de frequência (por exemplo, para 1.024 bins de frequência).

Opcionalmente, o decodificador pode incluir 600 decodificadores adicionais de espectro, como, por exemplo, descodificador de espectro TwinVQ e/ou um decodificador de

espectro BSAC, que pode ser utilizado como alternativa para o decodificador de espectro estendido AAC 630 em alguns casos.

O decodificador 600 pode opcionalmente incluir um processamento de espectro 640, que é configurado para processar a saída de informações 630B do decodificador 630 AAC estendido para obter informação de entrada 640A de um banco de filtro/ comutação de bloco 640. O processamento espectral opcional 630 pode incluir um ou mais, ou mesmo todas, as funcionalidades M/S, PNS, a previsão, a intensidade, previsão a longo prazo, dependente de comutação de acoplamento, TNS, dependente de comutação de acoplamento, onde as funcionalidades são descritas em detalhes na norma ISO / IEC 14.493,3: 2005 (E) e os documentos aqui citados. Se, no entanto, o processamento espectral 630 é omitido, as informações de saída 630B do decodificador 630 AAC estendido pode servir diretamente informação de entrada 640A do banco de filtro/ bloco de comutação 640. Assim, o decodificador 630 AAC estendido pode proporcionar, como informação de saída 630B informações, espectros escalados e quantizados inversamente. O banco de filtro/ bloco de comutação 640 usa, como informações de entrada 640A, os (opcionalmente pré-processadas) espectros quantizados inversamente, e prevê, com base no mesmo, um ou mais domínios do tempo de sinais de áudio reconstruídos uma informação de saída 640b. O banco de filtro/ bloco de comutação pode, por exemplo, ser configurado para aplicar o inverso da frequência com que o mapeamento foi realizado no codificador (por exemplo, no banco de filtro/ bloco de comutação 224). Por exemplo, uma transformada de cosseno discreta modificada inversa (IMDCT) pode ser usado pelo Banco de filtro. Por exemplo, o IMDCT pode ser configurado para

suportar tanto um conjunto de 120, 128, 480, 512, 960 ou 1024, ou quatro conjuntos de 32 ou 256 coeficientes espectrais.

Para mais detalhes, é feita referência, por exemplo, a Norma Internacional ISO / IEC 14496-3: 2005 (E). O decodificador 600 pode, opcionalmente, incluir mais um controle de ganho AAC 650, um decodificador de SBR 652 e um acoplamento independentemente de comutação de 654, para obter o sinal da hora de saída 612, do sinal de saída 640b do 640 banco de filtro/ bloco de comutação.

No entanto, o sinal de saída 640b do banco de filtro/ bloco de comutação 640 também pode servir como o tempo do sinal de saída 612, na ausência da funcionalidade de 650, 652, 654.

2.2.2. Detalhes do Codificador AAC Estendido

A seguir, serão descritos detalhes sobre o decodificador AAC estendido, tendo como referência as Figs. 7a e 7b. Figs. 7a e 7b mostram um diagrama esquemático de blocos do decodificador 630 AAC da Fig. 6, em combinação com o desformatador de carga útil de fluxo de dados 620 da Fig. 6.

o desformatador de carga útil de fluxo de dados 620, recebe um fluxo de áudio decodificado 610, que pode por exemplo, incluir um fluxo de dados de áudio incluindo uma elemento sintático entitulado "ac_raw_data_block", que é um codificador de áudio de bloco de dados brutos. No entanto, o formatador de carga útil do fluxo de bits de carga 620 é configurado para fornecer ao decodificador 630 AAC estendido um espectro codificado, quantizado e sem ruído ou uma representação, que compreende uma informação da linha espectral quantizada e codificada aritmeticamente (por

exemplo, designada como ac_spectral_data), uma informação de fator de escala 630ab (por exemplo, designada como scale_fator_data) e uma informação do parâmetro de preenchimento de ruído 630ac. A informação do parâmetro de preenchimento de ruído 630ac compreende, por exemplo, um offset de valor de ruído (designados com noise_offset) e um valor do nível de ruído (designados com noise_level).

Quanto ao decodificador AAC estendido, deve-se observar que o decodificador 630 AAC estendido é muito parecido com o decodificador AAC da norma internacional ISO / IEC 14496-3: 2005 (E), a referência que é feita a descrição detalhada na referida Norma.

O decodificador 630 AAC estendido inclui um decodificador de fator de escala 740 (também designada como ferramenta de decodificação silenciosa do fator de escala), que é configurado para receber as informações do fator de escala 630ab e oferecer a partir desta, uma representação inteira decodificada 742 dos fatores de escala (que também são designados como sf [g] [sfb] ou scf [g] [sfb]). Quanto ao decodificador fator de escala 740, é feita referência a norma ISO / IEC 14496-3: 2005, capítulos 4.6.2 e 4.6.3. Deve-se notar que a representação inteira decodificada 742 dos fatores de escala refletem uma precisão de quantização de diferentes faixas de frequência (também designada por faixas de fator de escala) de um sinal de áudio são quantizadas. os Fatores de escala maiores indicam que as faizas de fator de escala correspondentes foram quantificados com precisão elevada e menores fatores de escala indicam que as faixas de fator de escala correspondentes foram quantizadas com menor precisão.

O decodificador 630 AAC estendido também inclui um decodificador espectral 750, que é configurado para receber a entropia quantizada e codificada (por exemplo, codificados ou codificados aritmeticamente de Huffman) informação de linha espectral 630aa e fornecer, com base no mesmo, os valores quantizados 752 de um ou mais espectros (por exemplo, designada como x_ac_quant ou x_quant). Quanto ao decodificador espectral, é feita referência, por exemplo, a seção 4.6.3 do padrão internacional mencionado. No entanto, implementações alternativas do decodificador espectral podem naturalmente ser aplicados. Por exemplo, o decodificador de Huffman da ISO / IEC 14496-3: 2005 pode ser substituído por um decodificador aritmético 630aa se a informação de linha espectral é aritmeticamente codificada.

O decodificador 630 AAC estendido inclui ainda um quantizador inverso 760, que pode ser um quantizador não uniforme inverso. Por exemplo, o quantizador inverso 760 pode fornecer valores espectrais não-escalados quantizados inversamente 762 (por exemplo, designados com x_ac_invquant ou x_invquant). Por exemplo, o quantizador inverso 760 pode incluir a funcionalidade descrita na norma ISO / IEC 14496-3: 2005, capítulo 4.6.2. Alternativamente, o quantizador inverso 760 pode incluir a funcionalidade descrita com referência as Figs. 8 de 8c.

O decodificador 630 AAC estendido inclui também um preenchimento de ruído 770 (designado também como ferramenta de preenchimento de ruído), que recebe a representação inteira decodificada 742 dos fatores de escala do decodificador de fator de escala 740, os valores espectrais não escalados, inversamente quantizados 762 do quantizador inverso 760 e da informação do

parâmetro de preenchimento de ruído 630ac do desformatador de carga útil de fluxo de bits 620. O preenchedor de ruído é configurado para fornecer, com base no mesmo, a representação modificada (tipicamente inteira) 772 dos fatores de escala, que também é designado aqui como sf [g] [sfb] ou scf [g] [sfb]. O preenchimento de ruído 770 também é configurado para fornecer valores espectrais 774 inversamente quantizados e não escalados, também designado como x_ac_invquant ou x_invquant com base em suas informações de entrada. Detalhes a respeito da funcionalidade do preenchedor de ruído será posteriormente descrita, tendo referência às Figs. 9, 10a, 10b, 11, 12, 13a e 13b.

O decodificador 630 AAC estendido também dispõe de um rescalador 780, que é configurado para receber a representação inteira modificada dos fatores de escala 772 e os valores espectrais quantizados inversamente não-escalados 774, e fornecer, com base no mesmo, os valores espectrais quantizados inversamente 782, que também podem ser designados como x_rescal, e que pode servir de saída de informações 630B do decodificador 630 AAC estendido. O rescalador 780 pode, por exemplo, incluir a funcionalidade como descrito na ISO / IEC 14496-3: 2005, capítulo 4.6.2.3.3.

2.2.3. Quantizador Inverso

A seguir, a funcionalidade do quantizador inverso 760 será descrita, em referência a Figs. 8a, 8b e 8c. Fig. 8a mostra uma representação de uma equação para determinar os valores espectrais quantizados inversamente e não escalados 762 dos valores espectrais quantizados 752. Nas equações alternativas da Fig. 8-A, "o sinal (.)" designa um operador de sinal, e ".

"designa um operador de valor absoluto. A Fig. 8b mostra um pseudo código de programa que representa a funcionalidade do quantizador inverso 760. Como pode ser visto, a quantização inversa de acordo com a regra de mapeamento matemático mostrado na figura. 8 é realizada para todos os grupos da janela (designados por sfb variável em execução), para todas as faixas de fator de escala (designados pela execução de uma g variável), e para todas as janelas (designado por índice de execução win) e todas as linhas espectrais (ou bins espectrais) (designada por execução variável do bin). Fig. 8C mostra uma representação do fluxograma do algoritmo da figura. 8b. Para as faixas de fator de escala máxima pré-determinada (designados como max_sfb), os valores espectrais não-escalados quantizados inversamente são obtidos em função da dos valores espectrais não-escalados quantizados. A regra de quantização não-linear inversa é aplicada.

2.2.4 Preenchedor de Ruído

2.2.4.1. Preenchedor de Ruído de acordo com as Figs. 9 a 12

A Fig. 9 mostra um diagrama esquemático de blocos de um preenchedor de ruído 900 de acordo com uma materialização da invenção. O preenchedor de ruído 900 pode, por exemplo, tomar o lugar do preenchedor de ruído 770 descrito em referência às Figs. 7A e 7B.

O preenchedor de ruído 900 recebe a representação inteira decodificada 742 dos fatores de escala, que podem ser considerados como valores de ganho de faixa de frequência. O preenchedor de ruído 900 também recebe os valores espectrais não-escalados quantizados inversamente espectral 762. Além disso, o

preenchedor de ruído 900 recebe a informação do parâmetro do preenchimento de ruído 630ac, por exemplo, incluindo parâmetros de preenchimento de ruído noise_value e noise_offset. O preenchedor de ruído 900 fornece ainda a representação inteira modificada 772 dos fatores de escala e valores espectrais não-escalados quantizados inversamente 774. O preenchedor de ruído 900 dispõe de um detector de linhas espectrais quantizadas para zero 910, que é configurado para determinar se uma linha espectral (ou bin espectral) é quantizado para zero (e possivelmente ainda cumpre requisitos de preenchimento de ruído preenchimento). Para este efeito, o detector de linhas espectrais quantizadas para zero 910 recebe diretamente os espectros quantizados inversamente e não escalados 762 como informação de entrada. O preenchedor de ruído 900 inclui ainda um substituidor seletivo de linhas espectrais 920, que é configurado para substituir seletivamente os valores espectrais das informações de entrada de 762 por valores de substituição de linhas espectrais 922 na dependência do detector de linhas espectrais quantizadas para zero 910. Assim, se o detector de linhas espectrais quantizadas para zero 910 indica que uma determinada informação de entrada de linha espectral 762 deve ser substituída por um valor de substituição, o substituidor seletivo da linha espectral 920 substitui a determinada linha espectral com o valor de substituição da linha espectral 922 para obter as informações de saída 774. Caso contrário, o substituidor seletivo da linha espectral 920 encaminha o valor de determinada linha espectral sem alterações para obter as informações de saída 774. O preenchedor de ruído 900 também dispõe de um modificador seletivo de fatores de escala 930, que é configurado para

modificar seletivamente os fatores de escala das informações de entrada 742. Por exemplo, o modificador seletivo de fator de escala 930 é configurado para aumentar os fatores de escala das faixas de frequência dos fatores de escala, que foram quantificados para zero por um valor pré-determinado, que é designado como "noise_offset". Assim, as informações de saída 772, fatores de escala das faixas de frequência quantizadas para zero são aumentadas quando comparadas aos valores de fator de escala correspondentes na informação de entrada 742. Em contrapartida, os valores de fator de escala correspondentes de faixas de frequência de fatores de escala de frequência, que não são quantizadas a zero, são idênticas nas informações de entrada em 742 e as informações de saída 772.

Para determinar se uma faixa de frequência do fator de escala é quantizada para zero, o preenchedor de ruído 900 também conta com um detector de faixa quantizada para zero 940, que é configurado para controlar o modificador seletivo do fator de escala 930, fornecendo um sinal de "permissão de modificação fator de escala" ou bandeira 942, com base nas informações de entrada 762. Por exemplo, um detector de faixa quantizada para zero 940 pode fornecer um sinal ou bandeira indicando a necessidade de um aumento de um fator de escala para o modificador seletivo do fator de escala 930 se todos os bins de frequência (também designado como bins espectrais) de uma faixa de fator de escala são quantizadas a zero.

Deve-se notar aqui que o modificador seletivo do fator de escala também pode assumir a forma de um substituidor seletivo do fator de escala, que é configurado para ajustar

seguintes fatores de escala de faixas de fator de escala quantizado inteiramente para zero para um valor pré-determinado, independentemente das informações de entrada de 742.

A seguir, será descrito o novo rescalonador 950, que pode assumir a função do rescalonador 780. O rescalonador 950 está configurado para receber a representação inteira modificada 772 dos fatores de escala fornecidos pelo preenchedor de ruído e também para os valores espectrais não-escalados, quantizados inversamente 774 fornecidos pelo preenchedor do ruído. O rescalonador 950 dispõe de um computador de ganho de fatores de escala 960, que é configurado para receber uma representação inteira do Fator de escala por faixa do Fator de escala e de fornecer um valor de ganho por faixa de Fator de escala. Por exemplo, computador de ganho de fator de escala 960 pode ser configurado para calcular um valor de ganho de 962 para uma faixa de frequência i-th, com base em uma representação inteira modificada 772 da faixa do fator de escala i-th. Assim, o computador de ganho de fator de escala 960 fornece valores de ganho individuais para as diferentes faixas do fator de escala. Um rescalonador 950 compreende também um multiplicador de 970, que é configurado para receber os valores de ganho de 962 e os valores espectrais não-escalados, quantizados inversamente 774. Note-se que cada um dos valores espectrais não-escalados, quantizados inversamente 774 está associada a uma faixa de frequência de fator de escala (SFB). Assim, o multiplicador 970 está configurado para escalar cada um dos valores espectrais não-escalados, quantizados inversamente 774 com um valor de ganho correspondente associado com a mesma faixa de Fator de escala. Em outras palavras, todos os

valores espectrais não-escalados, quantizados inversamente 774 associados com uma determinada faixa de Fator escala são dimensionadas com o valor do ganho associado com a determinada faixa de Fator de escala. Consequentemente, os valores espectrais não-escalados, quantizados inversamente 774 associados a diferentes faixas de fator de escala são escalados com valores de ganho tipicamente diferentes associados com as diferentes faixas de fator de escalas.

Assim, os diferentes valores espectrais não-escalados, quantizados inversamente são dimensionados com diferentes valores de ganho, dependendo de quais faixas de Fator escala que estão associados.

Representação do Pseudo Código do Programa

A seguir, será descrita a funcionalidade do preenchedor de ruído 900, com referência as Figs. 10A e 10B, que mostram uma representação pseudo-código do programa (Fig. 10A) e uma legenda correspondente (Fig. 10B). Comentários iniciam com "--".

O algoritmo de preenchimento de ruído representada pela listagem do pseudo-código do programa na figura. 10 compreende uma primeira parte (linhas 1-8) para obter um valor de ruído (noiseVal) a partir de uma representação de nível de ruído (noise_level). Além disso, um offset de ruído (noise_offset) é derivado. A Determinação do valor do ruído do nível de ruído compreende uma escala não linear, onde o valor do ruído é calculado de acordo com

$$noiseVal = 2^{((noise_level-14)/3)}.$$

Além disso, uma mudança de escala do valor do

offset de ruído é realizado de tal forma que o intervalo do valor de compensação do offset de ruído pode assumir valores positivos e negativos.

A segunda parte do algoritmo (linhas 9-29) é responsável por uma substituição seletiva de valores espectrais não-escalados, quantizados inversamente e para a modificação seletiva dos seguintes fatores de escala. Como pode ser visto a partir do pseudo-código do programa, o algoritmo pode ser executado para todos os grupos de janela disponíveis (para-loop das linhas 9-29). Além disso, todas as faixas de fator de escala entre zero e faixa máxima de fator de escala (max_sfb) pode ser processado, embora o tratamento pode ser diferente para diferentes faixas do fator de escala (para o loop entre as linhas 10 e 28). Um aspecto importante é o fato de que é geralmente aceito que uma faixa de fator de escala é quantizada para zero, a menos que se verifique que a faixa do fator de escala não é quantizada para zero (conferir linha 11). No entanto, ao verificar se uma faixa de Fator de escala é quantizada para zero ou não, seja executado somente por faixas de Fator de escala, uma linha de frequência inicial (swb_offset [sfb]) de que está acima de um determinado índice de coeficiente espectral (noiseFillingStartOffset). Uma rotina condicional entre as linhas 13 e 24 só é executado se um índice dos mais baixos coeficientes espectrais da banda Fator sfb escala é maior do que o preenchimento inicial de offset de ruído início offset. Em contraste, para quaisquer faixas de Fator de escala para que um índice dos mais baixos do coeficiente espectral ([sfb] swb_offset) é menor ou igual a um valor predeterminado (noiseFillingStartOffset), presume-se que as faixas de não são

quantizadas a zero, independente da os valores reais espectral da linha (veja as linhas 24, 24b e 24c).

Se, no entanto, o índice dos menores coeficientes espectrais de uma determinada faixa de Fator de escala é maior que o valor pré-determinado (noiseFillingStartOffset), então determinadas faixas de Fator de escala são consideradas como sendo quantificada a zero somente se todas as linhas espectrais de determinadas faixas do Fator escala são quantizadas a zero (a bandeira "band_quantized_to_zero" é redefinida pelo loop entre as linhas 15 e 22, se um único bin espectral da faixa de fator de Fator de escala não é quantizada para zero.

Consequentemente, um Fator de escala de uma determinada faixa de escala Fator é modificada usando o ruído de deslocamento, se a bandeira "band_quantized_to_zero", que inicialmente é definida por padrão (linha 11) não é excluído durante a execução do código do programa entre as linhas 12 e 24. Como mencionado acima, uma redefinição da bandeira só pode ocorrer para as bandas Fator de escala para que um índice dos mais baixos do espectro coeficiente é superior ao valor pré-determinado (noiseFillingStartOffset). Além disso, o algoritmo da Figura. 10A compreende a substituição de valores de linhas espectrais com valores de reposição de linhas espectrais se a linha espectral é quantizada a zero (condição da linha 16 e operação de substituição da linha 17). No entanto, disse que a substituição só é realizada para as bandas Fator de escala para que um índice dos mais baixos do espectro coeficiente é superior ao valor pré-determinado (noiseFillingStartOffset). Para menores faixas de frequências do espectro, a substituição dos valores espectrais quantizada a zero

com os valores de reposição espectral é omitido.

Consequentemente, um Fator de escala de uma determinada faixa de Fator de escala é modificada usando o offset de ruído, se a bandeira "band_quantized_to_zero", que inicialmente é definida por padrão (linha 11) não é excluída durante a execução do código do programa entre as linhas 12 e 24. Como mencionado acima, uma redefinição da bandeira só pode ocorrer para as faixas de Fator de escala para que um índice dos menores coeficientes do espectro seja superior ao valor pré-determinado (noiseFillingStartOffset). Além disso, o algoritmo da Figura. 10A compreende a substituição de valores de linhas espectrais com valores de reposição de linhas espectrais se a linha espectral é quantizada a zero (condição da linha 16 e operação de substituição da linha 17). No entanto, a substituição mencionada só é realizada para as faixas de Fator de escala para que um índice dos menores coeficientes do espectro seja superior ao valor pré-determinado (noiseFillingStartOffset). Para menores faixas de frequências do espectro, a substituição dos valores espectrais quantizada a zero com os valores de reposição espectral é omitida.

Deve-se observar ainda que os valores de substituição podem ser calculados de forma simples em que um sinal aleatório ou pseudo-aleatório é adicionado ao valor do ruído (noiseVal) calculado na primeira parte do algoritmo (conferir linha 17).

Observe que Fig. 10B mostra uma legenda dos símbolos relevantes utilizados no pseudo-código do programa da Fig. 10A para facilitar uma melhor compreensão do pseudo-código do programa.

Importantes aspectos da funcionalidade do offset de ruído são ilustrados na fig. 11. Como pode ser visto, a funcionalidade do produto de preenchimento de ruído, opcionalmente, compreende computação 1110 um valor de ruído em função do nível de ruído. A funcionalidade ddo preenchedor de ruído também inclui a substituição dos 1120 de valores de linhas espectrais de linhas espectrais quantizada para zero, com valores de substituição de linhas espectrais na dependência do valor do ruído para obter valores de substituição de linhas espectrais. No entanto, a substituição de 1120 é realizada apenas para as faixas de Fator de escala com um menor coeficiente espectral acima de um determinado índice espectral coeficiente.

A funcionalidade do preenchimento de ruído também inclui a modificação 1130 uma faixa de Fator de escala na dependência do valor do ruído de deslocamento, se e somente se, a faixa do Fator de escala é quantizada a zero. No entanto, a modificação 1130 é executada em que se formam para as faixas de Fator de escala com um menor coeficiente espectral acima do índice de coeficiente espectral determinado.

O preenchimento do ruído também inclui uma funcionalidade de saida 1140 de faixas de fatores de escala afetada, independente se a faixa de Fator de escala é quantizada a zero, para as faixas de Fator de escala com um menor coeficiente espectral abaixo do índice de coeficiente espectral determinado.

Além disso, o rescalonador inclui uma funcionalidade da aplicação 1150 não modificado ou modificado (o que for disponível) das seguintes faixas de fatores de escala para obter os valores de linhas espectrais escaladas e inversamente

quantizadas não-substituído ou substituído (o que for disponível).

A Fig. 12 mostra uma representação esquemática do conceito descrito, com referência as figs. 10A, 10B e 11. Em particular, as diferentes funcionalidades são representadas na dependência de um Bin inicial de uma faixa de Fator de escala.

2.2.4.2 Preenchimento de Ruído de enchimento de acordo com as Figs. 13A e 13B.

As Figs. 13A e 13B mostram listagens de algoritmos de pseudo-código que podem ser realizados em uma implementação alternativa do offset de ruído 770. A Fig. 13 descreve um algoritmo para obter um valor de ruído (para utilização no preenchimento de ruído) a partir de uma informação de nível de ruído, que pode ser representado pelas informações do parâmetro de preenchimento de ruído 630ac.

Como o erro de quantização médio é aproximadamente 0,25 na maioria do tempo, o intervalo noiseVal [0,0.5] é bastante grande e pode ser otimizado.

A Fig. 13 representa um algoritmo, que pode ser formado pelo preenchedor de ruído 770. O algoritmo da Fig. 13 compreende uma primeira porção da determinação do valor de ruído (designado por "noiseValue" ou "noiseVal" - que é a linha s 1 à 4). A segunda parte do algoritmo é composta por uma modificação seletiva de um Fator de escala (linhas 7 à 9) e uma substituição seletiva de valores de linhas espectrais com valores de substituição de linhas espectrais (linhas 10 à 14).

No entanto, de acordo com o algoritmo da Fig. 13, o Fator de escala (SCF) é modificado usando o offset de ruído (noise_offset) sempre que uma banda é quantizada para zero (ver

linha 7). Não houve diferença entre as faixas de baixa frequência e as faixas de frequência na presente materialização.

Além disso, o ruído é introduzido em linhas espectrais quantizadas para zero apenas para as faixas de frequência mais altas (se a linha estiver acima de um limite pré-determinado "noiseFillingStartOffset").

2.2.5. Conclusão do Decodificador

Para resumir, materializações do decodificador de acordo com a presente invenção pode compreender um ou mais dos seguintes recursos:

- Começando a partir da "linha de início do preenchimento de ruído" (que pode ser um offset fixo ou uma linha representando uma frequência de início substituindo cada 0 com um valor de substituição
- O valor de substituição é o valor de ruído indicado (com um sinal aleatório) no domínio quantizado e então escala esse "valor de substituição" com o fator de escala "scf") transmitido para a banda real do fator de escala; e
- Os valores aleatórios de substituição também podem ser derivados de, por exemplo, uma distribuição de ruído ou um conjunto de valores alternáveis ponderados com um nível de ruído sinalizados.

3. Fluxo Áudio

3.1. Fluxo de Áudio de acordo com as Figs. 14A e 14B

A seguir, um fluxo de áudio de acordo com uma materialização da invenção será descrita. A seguir, a chama "carga de fluxo de bits usac" irá ser descrita. A "carga de fluxo de bits

usac" carrega informações de carga para representar um ou mais canais singulares (payload "single_channel_element ()) e/ou um ou mais pares de canais (channel_pair_element (), como pode ser visto da Fig. 14A. Uma informação de canal singular (single_channel_element ()) compreende, entre outras informações opcionais, um fluxo de canal do domínio de frequência (fd_channel_stream), como pode ser visto na Fig. 14B.

Uma informação de par de canal (channel_pair_element) compreende, além de elementos adicionais, uma pluralidade de, por exemplo, dois fluxos de canal do domínio de frequência (fd_channel_stream), como pode ser visto na Fig. 14C.

O conteúdo de dados do fluxo de canal do domínio de frequência pode, por exemplo, ser dependente se um preenchimento de ruído é usado ou não (que pode ser sinalizado em uma porção de dados sinalizados não mostrados aqui). A seguir, será pressuposto que o preenchimento de ruído é usado. Nesse caso, o fluxo de canal do domínio de frequência compreende, por exemplo, os elementos de dados mostrados na Fig. 14D. Por exemplo, uma informação de ganho global (global_gain), como definido em ISO/IEC 14496-3: 2005 pode estar presente. Além disso, o fluxo de canal do domínio de frequência pode compreender uma informação de offset de ruído (noise_offset) e uma informação do nível de ruído (noise_level), como descrito aqui. A informação do offset de ruído pode, por exemplo, ser codificado usando 3 bits e a informação do nível de ruído pode, por exemplo, ser codificado usando 5 bits.

Além disso, o fluxo de canal do domínio de frequência pode compreender uma informação do fator de escala

codificado (a scale_factor_data ()) e de dados espectrais codificados de forma aritmética (AC_spectral_data ()) como descrito aqui e também como definido na ISO/IEC 14496-3

Opcionalmente, o fluxo de canal do domínio de frequência também compreende dados de modelação do ruído temporal (tns_data) ()), conforme definido na ISO / IEC 14496-3.

Naturalmente, o fluxo de canal do domínio de frequência pode compreender outras informações, se necessário.

3.2. Fluxo de Áudio de acordo com as Figs. 15

Fig. 15 mostra a representação esquemática da sintaxe do fluxo de canal representando um canal individual (individual_channel_stream ()).

O fluxo do canal individual pode compreender um ganho de informação global (global_gain) codificado utilizando, por exemplo, 8 bits, a informação de offset de ruído (noise_offset) codificado utilizando, por exemplo, 5 bits e uma informação de nível de ruído (noise_level) codificado utilizando, por exemplo, 3 bits.

O fluxo do canal individual inclui ainda dados da seção (section_data ()), os dados de fator de escala (scale_factor_data ()) e dados espectrais (spectral_data ()).

Além disso, o fluxo do canal individual pode incluir mais informações opcionais, como pode ser visto na Figura. 15.

3.3. Conclusão Audio Stream

Para resumir o supramencionado, em algumas materializações de acordo com a invenção, os seguintes elementos de sintaxe de fluxo de bits são utilizados:

• Valor que indica um offset fator de escala do ruído para otimizar os bits necessários para transmitir os fatores de escala;

• valor que indica o nível de ruído; e/ou;

• valor opcional para escolher entre diferentes formas para a substituição de ruído (ruídos distribuídos uniformemente ao invés de valores constantes ou múltiplos níveis discretos em vez de apenas um).

4. Conclusão

Em codificação de razão de baixos de bits, o ruído de preenchimento pode ser usado para duas finalidades:

• quantização grosseira de valores espectrais em uma codificação de áudio de baixa taxa de bits pode levar a espectros escassos após a quantização inversa, já que muitas linhas espectrais poderiam ter sido quantificadas a zero. Os espectros povoados de forma escassa irão resultar no sinal decodificado soando agudo ou instável(passarinhos). Ao substituir as linhas zeradas com "pequenos" valores no decodificador, é possível mascarar ou reduzir esses artefatos muito óbvios, sem acrescentar novos óbvios artefatos de ruído.

• Se houver partes do sinal como ruído no espectro original, uma representação perceptivamente equivalente dessas partes de sinal ruidosas podem ser reproduzidas no decodificador baseado em apenas poucas informações paramétricas, como a energia da parte do sinal com ruído. As informações paramétricas podem ser transmitidas com menos bits em relação ao número de bits necessários para transmitir a forma de onda codificada.

O novo esquema de codificação de preenchimento de ruído proposto aqui descritos combina de forma eficiente os fins acima em uma única aplicação.

Como comparação, em áudio MPEG-4, a substituição de ruído perceptivo (PNS) é usada apenas para transmitir uma informação parametrizada de partes de sinal com ruído e como reproduzir estas partes de sinal de forma perceptivamente equivalente no decodificador.

Como comparação adicional, em AMR-WB +, vetores de quantização vetorial (VQ-vectors) quantizados a zero são substituídos por um vetor de ruído aleatório, onde cada valor espectral complexo tem amplitude constante, mas fase aleatória. A amplitude é controlada por um valor de ruído transmitido com o fluxo de bits.

No entanto, os conceitos de comparação fornecem desvantagens significativas. PNS só pode ser usado para preencher bandas de fator de escala completas com ruído, enquanto AMR-WB + apenas tenta mascarar artefatos no sinal decodificado resultando em grandes partes do sinal serem quantizadas a zero. Em contraste, o esquema de codificação de preenchimento de ruído proposto combina de forma eficiente os aspectos de preenchimento de ruído em uma única aplicação.

De acordo com um aspecto, a presente invenção compreende uma nova forma de cálculo do nível de ruído. O nível de ruído é calculado no domínio quantizado com base no erro de quantização médio.

O erro de quantização no domínio quantizado difere de outras formas de erro de quantização. O erro de

quantização por linha no domínio quantizado está no intervalo [-0,5, 0,5] (1 nível de quantização) com um erro médio absoluto de 0,25 (para valores de entrada distribuídos normais que são geralmente maiores do que 1).

A seguir, algumas vantagens do preenchimento de ruído no domínio quantizado serão resumidas. A vantagem da adição de ruído no domínio quantizado é o fato de que o ruído adicionado no decodificador é dimensionado, não só com a energia média em uma determinada faixa, mas também com a relevância psicoacústica de uma banda.

Normalmente, as bandas mais perceptivamente relevantes (tonal) serão as bandas quantizadas com mais precisão, ou seja, múltiplos níveis de quantização (valores quantizados maiores que 1) serão usados nestas bandas. Agora, a adição de ruído com um nível de erro de quantização médio nestas bandas só terá influência muito limitada sobre a percepção de uma banda.

Bandas que não são tão relevantes ou com mais ruído, podem ser quantificadas com um número menor de níveis de quantização. Apesar de que muito mais linhas espectrais na banda serão quantizadas a zero, o erro de quantização médio resultante será o mesmo para as bandas quantizadas adequadas (pressupondo um erro de quantização de distribuição normal em ambas as bandas), enquanto o erro relativo na banda pode ser muito maior.

Nestas bandas quantizadas de forma grosseira, o preenchimento de ruído vai ajudar a mascarar perceptivamente os artefatos resultantes de furos espectral devido à quantização grosseira.

Uma reflexão sobre o preenchimento de ruído no

domínio quantizado pode ser alcançada pelo codificador descrito acima e também pelo decodificador descrito acima.

5. Alternativas de Implementação

Dependendo de certos requisitos de implementação, as materializações da invenção podem ser implementadas em hardware ou software. A aplicação pode ser realizada utilizando um meio de armazenamento digital, por exemplo, uma disquete, um DVD, um CD, um ROM, um PROM, uma EPROM, uma EEPROM ou memória flash, tendo sinais de controle lidos eletronicamente, que cooperam (ou são capazes de cooperar) com um sistema de computador programável de tal forma que o método em questão é realizado.

Algumas materializações de acordo com a invenção compreende um portador de dados tendo sinais de controle de leitura eletrônica, que são capazes de cooperar com um sistema de computador programável, de modo que um dos métodos descritos aqui é realizado.

Geralmente, as materializações da presente invenção podem ser implementadas como um produto de programa de computador com um código de programa, o código do programa operativo para realização de um dos métodos, quando o produto programa de computador é executado em um computador. O código de programa pode, por exemplo, ser armazenado em um portador de leitura de máquina.

Outras materializações incluem um programa de computador para a realização de um dos métodos descritos neste documento, armazenado em um portador de leitura de máquina.

Em outras palavras, uma materialização do método inventivo é, portanto, um programa de computador com um código de

programa para executar um dos métodos descritos aqui, quando o programa de computador é executado em um computador.

Uma materialização dos métodos inventivos é, portanto, um portador de dados (ou um meio de armazenamento digital, ou um meio de leitura por computador), que inclui, gravado nele, o programa de computador para a realização de um dos métodos descritos neste documento.

Uma materialização do método inventivo é, portanto, um fluxo de dados ou uma sequência de sinais que representam o programa de computador para a realização de um dos métodos descritos neste documento. O fluxo de dados ou a sequência de sinais podem, por exemplo, ser configurados para serem transferidos através de uma conexão de comunicação de dados, por exemplo através da Internet.

Uma materialização compreende ainda um meio de processamento, por exemplo um computador, ou um dispositivo de lógica programável, configurado ou adaptado para executar um dos métodos descritos neste documento. Al

Uma outra materialização compreende ainda um computador que tenha instalado nele o programa de computador para a realização de um dos métodos descritos neste documento.

REIVINDICAÇÕES

1. **"CODIFICADOR DE ÁUDIO, DECODIFICADOR DE ÁUDIO, MÉTODOS PARA CODIFICAR E DECODIFICAR UM SINAL DE ÁUDIO, FLUXO DE ÁUDIO DE PROGRAMA DE COMPUTADOR", caracterizado** pelo codificador (100; 228) para fornecimento de um fluxo de áudio (126; 212) na base de uma representação de domínio da transformada (112; 114; 228a) de uma entrada de sinal de áudio, compreender:

um calculador de erro de quantização (110; 330) configurado para determinar erro de quantização da multi-banda (116; 332) sobre uma pluralidade de faixas de frequência de sinal de entrada de áudio, para o qual o ganho de informação de faixa separada (228a) está disponível; e

um fornecedor de fluxo de áudio (120; 230) configurado para fornecer fluxo de áudio (126; 212) de modo que o fluxo de áudio inclua uma informação descrevendo o conteúdo de áudio das faixas de frequência e uma informação descrevendo o erro de quantização das multi-bandas

2. Codificador (100; 228) de acordo com a reivindicação 1, **caracterizado** pelo calculador de erro de quantização (110; 330) estar configurado para calcular uma média de e erros de quantização sobre uma pluralidade de faixas de frequência da entrada do sinal de áudio, para onde o ganho de informação de faixa separa está disponível, de modo que a informação de erro de quantização cubra uma pluralidade de faixas de frequência, para o qual o ganho de informação de faixa separada está disponível.

3. Codificador (100; 228) de acordo com reivindicação 1 e 2, **caracterizado** pelo codificador incluir um

quantizador (310) configurado para quantizar os componentes espectrais de diferentes faixas de frequência da representação do domínio da transformada (228a) usando diferentes precisões de quantização em dependência de relavâncias psicoacústicas (228c) de diferentes faixas de frequência, para obter componentes espectrais quantizados, onde as diferentes precisões de quantização são refletidas pelo ganho de informação de faixa; e

onde o fornecedor de fluxo de áudio (212) é configurado para fornecer um fluxo de áudio de modo que o fluxo de áudio inclua uma informação descrevendo o ganho de informação de faixa de modo que o fluxo de áudio inclua também a informação descrevendo o erro de quantização de multi-banda.

4. Codificador (100; 228) de acordo com a reivindicação 3, **caracterizado** pelo quantizador (310) ser configurado para realizar um escalamento do componente espectral em dependência da informação de ganho da faixa e realizar a quantização de um valor íntegro dos componentes espectrais em escala; e

onde o calculador de erro de quantização (330) é configurado para determinar o erro de quantização da multi-banda (332) no domínio quantizado, de modo que o escalamento dos componentes espectrais que é realizado antes da quantização do valor integro, é levado em consideração no erro de quantização de multi-banda.

5. Codificador (100; 228) de acordo com uma das reivindicações 1 a 4, **caracterizado pelo** codificador ser configurado para ajustar um ganho de informação da faixa de uma faixa de frequência, que foi completamente quantizada para zero,

para um valor representando uma taxa entre a energia da faixa de frequência completamente quantizado para zero e a energia do erro de quantização da multi-faixa.

6. Codificador (100; 228) de acordo com as reivindicações 1 a 5, **caracterizado** pelo calculador de erro de quantização (330) ser configurado para determinar o erro de quantização da multi-banda (332) sobre uma pluralidade de faixas de frequência, cada uma incluindo ao menos um componente espectral quantizado para um valor não zero, enquanto evitando faixas de frequência, componentes espectrais que foram completamente quantizados para zero.

7. Decodificador (500, 600), **caracterizado** por fornecer uma representação decodificada (512, 514, 630B) de um sinal de áudio a partir de um fluxo de áudio codificado (510, 610) representa os componentes espectrais das bandas de frequência do sinal de áudio compreender:

um preenchimento de ruído (520, 770) configurado para introduzir ruído em componentes espectrais de uma pluralidade de faixas de frequência, as quais as informações do ganho de faixa de frequência separada estão associadas, com base em um valor comum de ruído de multi-banda(526).

8. Decodificador (500; 600) de acordo com a reivindicação 7, **caracterizado** pelo decodificador incluir um rescalonador (780), que é configurado para receber uma representação de um ganho de informação de frequência separada e valores espectrais fora da escala quantizados inversamente (774), e para fornecer, na base, valores espectrais quantizados inversamente (782).

9. Decodificador (500; 600) de acordo com a reivindicação 7 ou 8, **caracterizado** pelo preenchimento de ruído (520; 770) ser configurado para decidir seletivamente em uma base por bin do espectro, onde para introduzir o ruído em bins de espectro individuais de uma faixa de frequência em dependência de se os bins do espectro serão quantizados para zero ou não.

10. Decodificador (500; 600) de acordo com uma das reivindicações 7 a 9, **caracterizado** pelo preenchimento de ruído (520; 770) ser configurado para receber a pluralidade de valores de bins de espectro (522) representando diferentes porções de frequência de sobreposição e não sobreposição de uma primeira faixa de frequência de uma representação do sinal de áudio do domínio da frequência, e para receber uma pluralidade de valores de bins do espectro (524) representando diferentes porções de frequência de sobreposição e não sobreposição de uma segunda faixa de frequência de uma representação do sinal de áudio do domínio da frequência; para substituir um ou mais, para substituir um ou m ais valores de bins espectrais da primeira faixa de frequência da pluralidade das faixas de frequência com um primeiro valor de bin espectral, uma magnitude no qual é determinado pelo valor de intensidade de ruído da multi-banda (526), e para substituir um ou m ais valores de bins espectrais da segunda faixa de frequência da pluralidade das faixas de frequência com um segundo valor de bin espectral, uma magnitude no qual é determinado pelo valor de intensidade de ruído da multi-banda com a mesma magnitude do primeiro valor de ruído do Bin do espectro; onde o decodificador inclui um escalímetro (780) configurado para escalar os valores do bin espectral da primeira faixa de frequência da pluralidade das

faixas de frequência com o primeiro valor de ganho da faixa de frequência, para obter valores dedo bin espectral em escala da primeira faixa de frequência, e para escalar os valores do bin espectral da segunda faixa de frequência da pluralidade das faixas de frequência com o segundo valor de ganho da faixa de frequência, para obter valores do bin espectral em escala da segunda faixa de frequência, de modo que os valores de bin espectral substituídos, com o primeiro e segundos valores de ruídos de bin espectral, são escalados com diferentes valores de ganho de faixa de frequência, e de modo que o valor de bin espectral substituído, com o primeiro valor de bin espectral substituído, e valores de bin espectral não substituídos da primeira faixa de frequência representando um conteúdo de áudio de uma primeira faixa de frequência são escaladas com o primeiro valor de ganho de faixa de frequência, e que o valor de bin espectral substituído, com o segundo valor de bin espectral substituído, e valores de bin espectral da segunda faixa de frequência representando um conteúdo de áudio da segunda faixa de frequência são escalados com o valor de ganho da segunda faixa de frequência.

11. Decodificador (500; 600) de acordo com uma das reivindicações 7 a 10, **caracterizado** pelo preenchimento de ruído (520; 770) ser configurado para modificar seletivamente um valor de ganho de faixa de frequência de uma determinada frequência usando um valor de offset de ruído caso a determinada faixa de frequência seja quantizada para zero.

12. Decodificador (500; 600) de acordo com reivindicações 7 a 11, **caracterizado** pelo preenchimento de ruído (520; 770) ser configurado para substituir valores de bin

espectral de bins espectrais quantizados para zero com valores de ruídos de bins espectrais, magnitudes na qual os valores de bins espectrais são dependentes de valores de intensidade de ruído da multi-banda (526), para obter os valores de bins espectrais substituídos, somente para faixas de frequência com o menor índice de bin espectral acima de um índice de bin espectral pré-determinado, deixando os valores de bins espectrais de faixas de frequência com menores índices com de bin espectral abaixo de índices de bin espectral pré-determinados não afetados;

onde o preenchimento de ruído é configurado para modificar seletivamente, para faixas de frequência com o menor índice de bin espectral acima do índice de bin espectral pré-determinado, o valor de ganho da faixa de uma determinada faixa de frequência em dependência de um valor de offset de ruído, caso a faixa de frequência seja inteiramente quantizado para zero; e

onde o decodificador inclui um escalímetro adicional (770) configurado para aplicar os valores de ganho de faixas modificadas e não modificadas, para os valores de bin espectrais seletivamente substituídos ou não substituídos, para obter uma informação espectral em escala, que representa o sinal de áudio.

13. Decodificador (500; 600) de acordo com as reivindicações 7 a 12, **caracterizado** pelo decodificador ser configurado para receber um fluxo de áudio (610) incluindo uma representação quantizada de entropia codificada (630aa) dos valores de bins espectral para uma pluralidade de faixas de frequência, onde uma pluralidade de valores do Bin espectral está associada com uma primeira faixa de frequência da pluralidade das

faixas de frequência, e onde a pluralidade de faixas de frequência está associada com uma segunda faixa de frequência da pluralidade das faixas de frequência,

uma representação codificada (630ab) dos valores de ganho de faixa, onde um primeiro ganho de faixa está associado com a primeira faixa de frequência e um segundo valore de ganho de faixa está associado com uma segunda faixa de frequência, e

um representação codificada (630ac) do valor de intensidade de ruído da multi-banda;

onde o decodificar inclui um decodificador espectral (750) configurado para fornecer uma representação codificada e quantizada (752) dos valores de bins espectrais na base de representação quantizada de entropia codificada dos valores de bins espectrais;

onde o decodificador inclui um quantizador inverso (760) configurador para quantizar inversamente a representação decodificada quantizada (752) dos valores de Bin espectral, para obter uma representação decodificada e inversamente quantizada (762) dos valores de Bin espectral;

onde o decodificador inclui um decodificador de fator de escala (740) configurado para decodificar a representação codificada (630ab) do valor de ganhos espectrais, para obter uma representação decodificada (742) do valor dos ganhos espectrais; e

onde o preenchimento de ruído (770) é configurado para substituir seletivamente os valores de bins espectrais inversamente quantizados para zero das múltiplas faixas de frequência com valores de substituição dos valores de Bin espectral de magnitudes idênticas, para obter os valores de Bin

espectral substituídos das faixas de frequência; e

onde o decodificador inclui um escalímetro (780) configurado para escalar um conjunto de todos os valores espectrais de uma primeira faixa de frequência, alguns dos valores espectrais da primeira faixa de frequência são originalmente inversamente quantizados, valores de bin espectrais decodificados fornecidos pelo quantizador inverso e alguns dos valores são valores substituição de bin espectral, com representação decodificada de um fator de escala associado com a primeira faixa de frequência, para obter um conjunto de valores de Bin espectral em escala da primeira faixa de frequência, e para escalar um conjunto de todos os valores de Bin espectral de uma segunda faixa de frequência, dos quais são valores de Bin espectral de segunda faixa de frequências, originalmente quantizados inversamente, os valores de Bin espectrais decodificados fornecidos pelo quantizador inverso e alguns valores de Bin espectral são valores de substituição de Bin espectral, com uma representação codificada de um fator de escala associado com uma segunda faixa de frequência, para obter um conjunto de valores de Bin espectral em escala da segunda faixa de frequência.

14. Método para fornecer um fluxo de áudio (126; 212) com base na representação do domínio da transformada (112; 114;228a) de uma entrada de sinal de áudio, **caracterizado** por compreender:

determinação de um erro de quantização da multi-banda sobre uma pluralidade de faixas de frequência, para os quais os valores de ganho de faixas separadas estejam disponíveis; e

fornecimento de um fluxo de áudio de modo que o

fluxo de áudio inclua uma informação descrevendo um conteúdo de áudio das faixas de frequência descrevendo um erro de quantização de multi-banda.

15. Método para fornecer uma representação decodificada (512; 514: 630b) de um sinal de áudio com base em um fluxo de áudio codificado (510; 610), **caracterizado** por compreender:

introdução ruído em componentes espectrais de uma pluralidade de faixas de frequência, para qual a informação de ganho de faixas de frequência é associada com base em um valor de intensidade de ruído de multi-banda comum.

16. Programa de computador, **caracterizado** para realizar o método de acordo com uma das reivindicações 14 ou 15 quando executado em um computador.

17. Fluxo de áudio (510; 610) representando um sinal de áudio, **caracterizado** por compreender:

informação espectral descrevendo as intensidades de componentes espectrais do sinal de áudio, onde a informação espectral é quantizada com diferentes precisões de quantização e diferentes faixas de frequência; e

um informação de nível de ruído descrevendo um erro de quantização de multi-banda sobre um pluralidade de faixas de frequência, considerando as diferentes precisões de quantização.

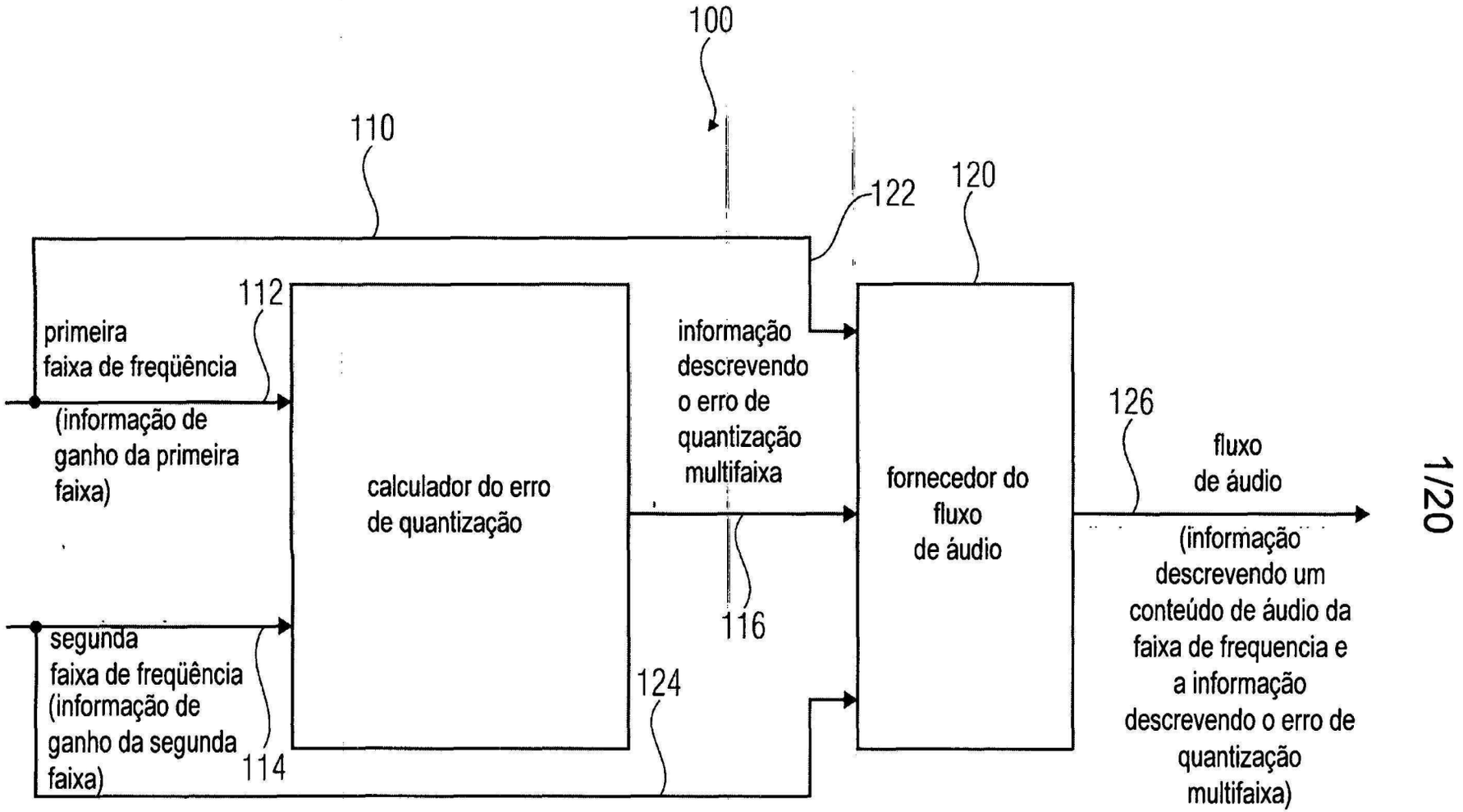
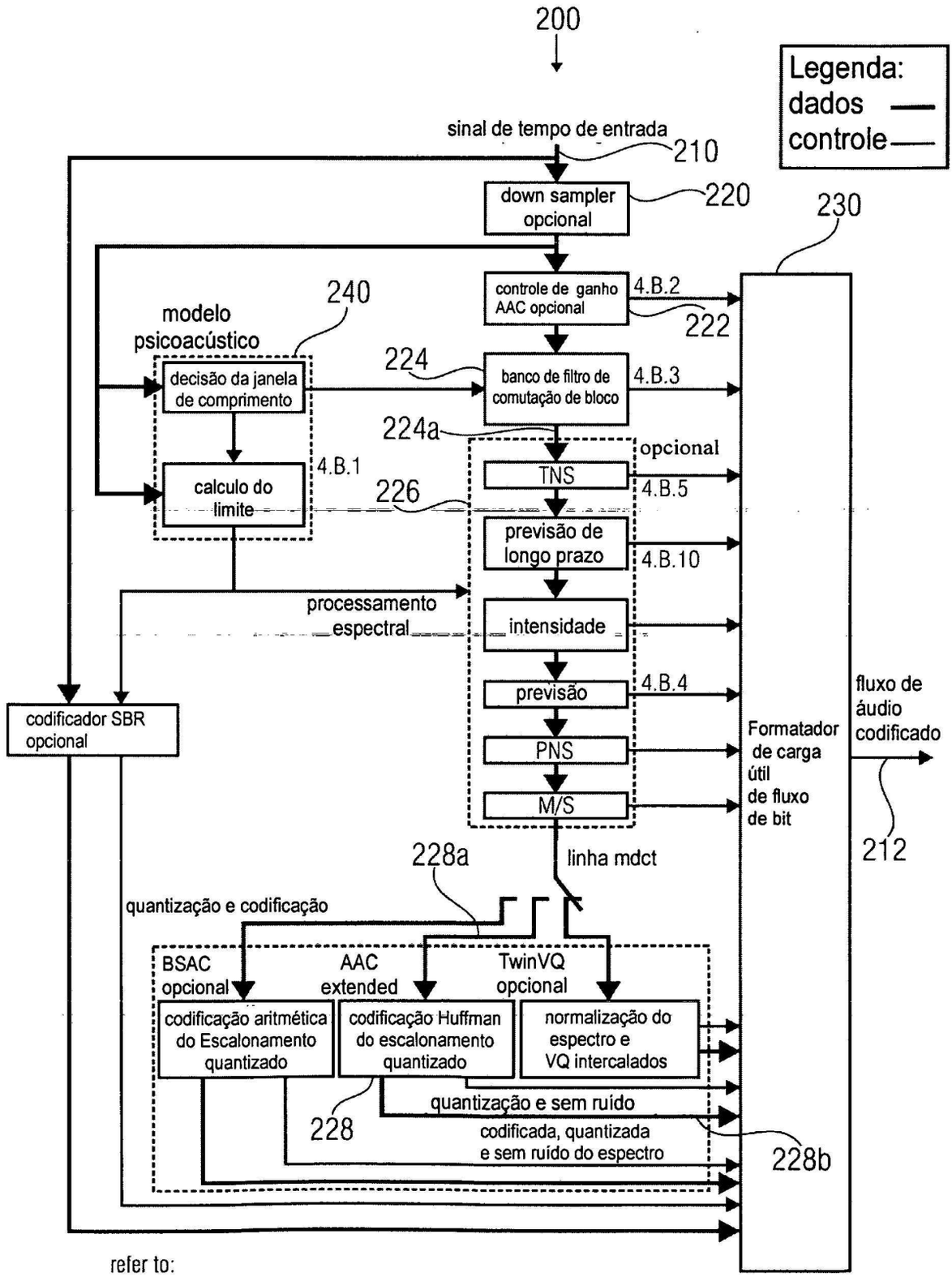

100
110
122
120
112
primeira
faixa de freqüência
(informação de
ganho da primeira
faixa)
calculador do erro
de quantização
informação
descrevendo
o erro de
quantização
multifaixa
fornecedor do
fluxo
de áudio
126
fluxo
de áudio
(informação
descrevendo um
conteúdo de áudio da
faixa de frequencia e
a informação
descrevendo o erro de
quantização
multifaixa)
116
segunda
faixa de freqüência
(informação de
ganho da segunda
faixa)
114
124

FIG 1



refer to:
ISO/IEC 14496-3: 2005 Subport 4, Fig. 4.1

FIG 2

FIG 3A1

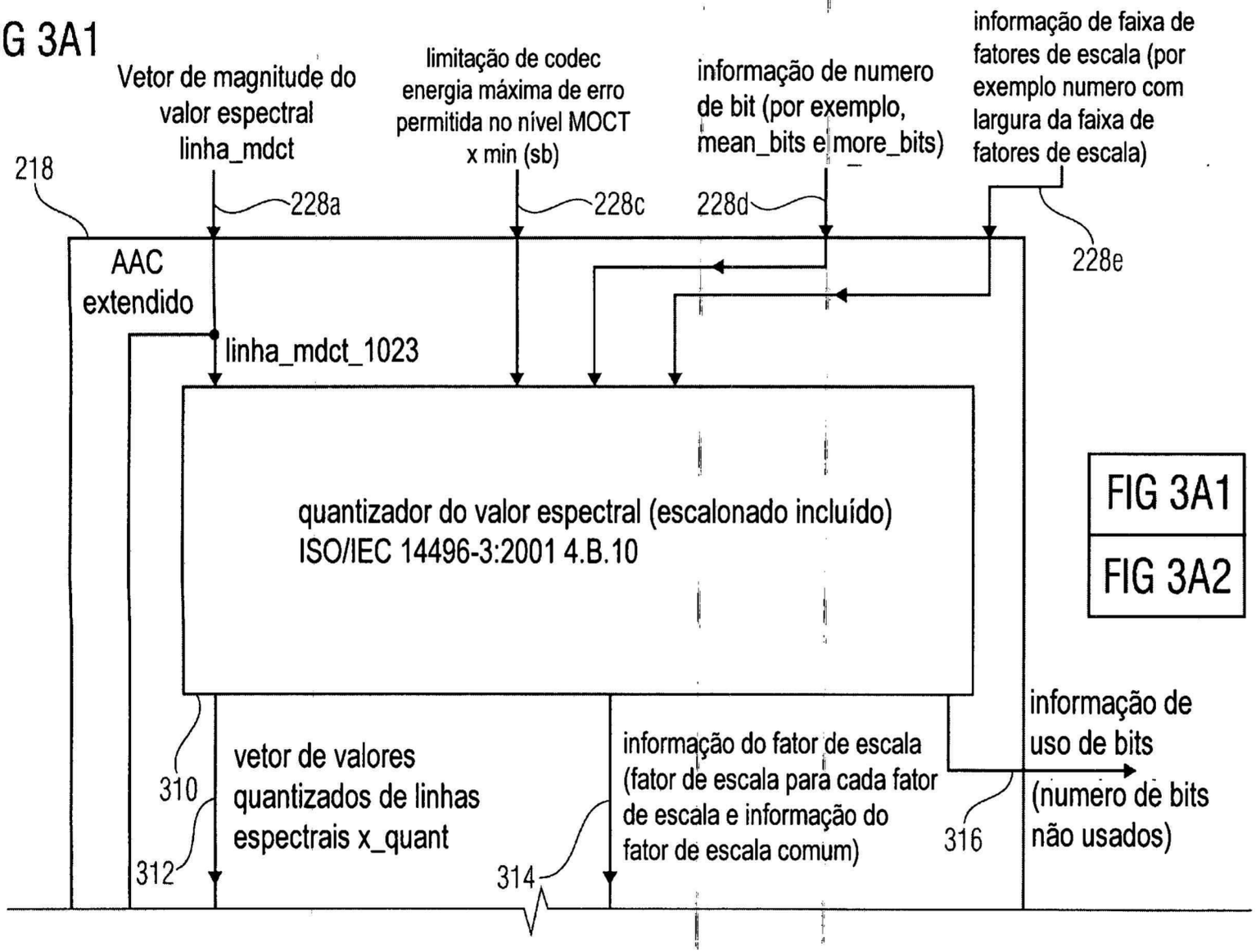


FIG 3A1

FIG 3A2

FIG 3A1

FIG 3A2

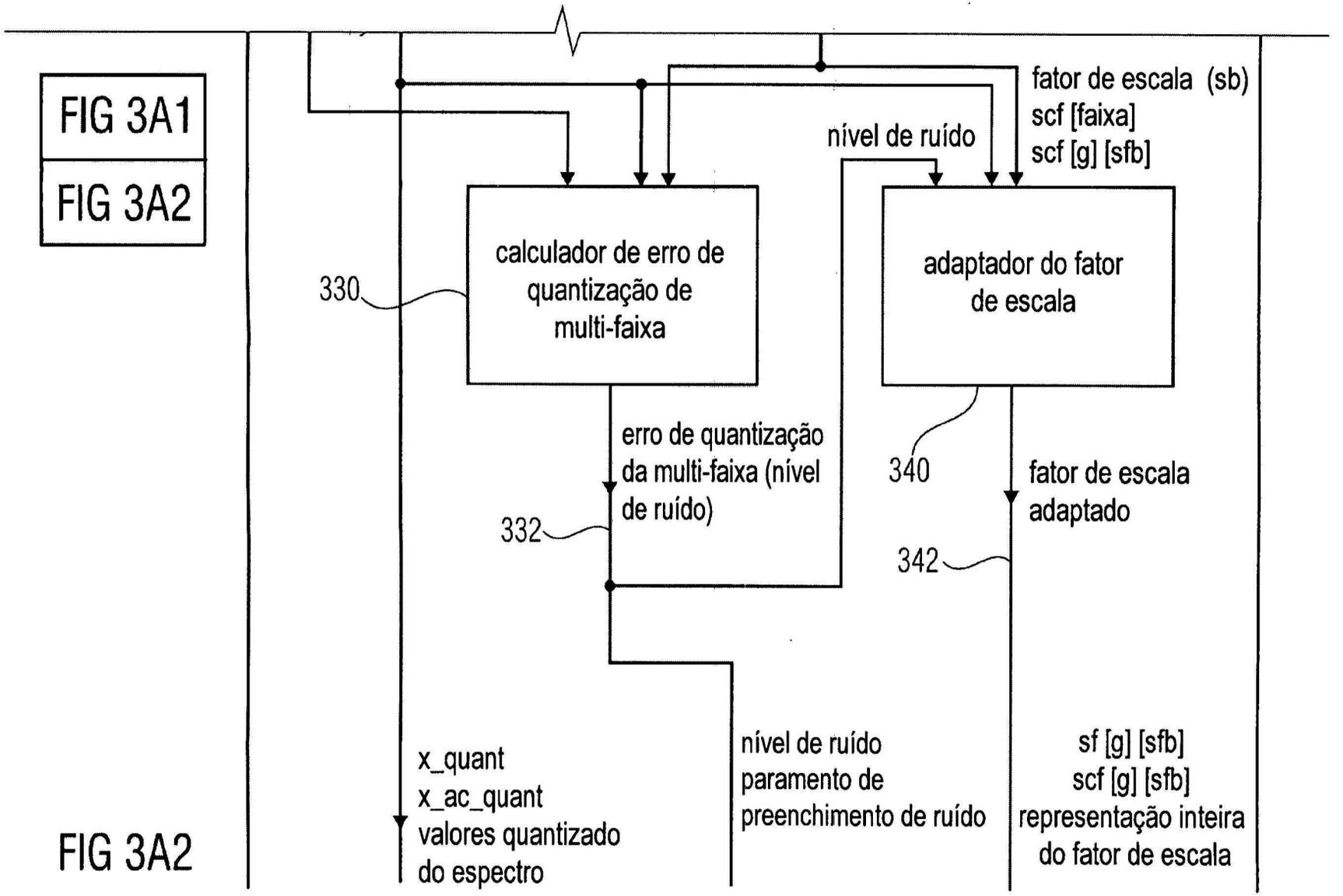


FIG 3A2

FIG 3B

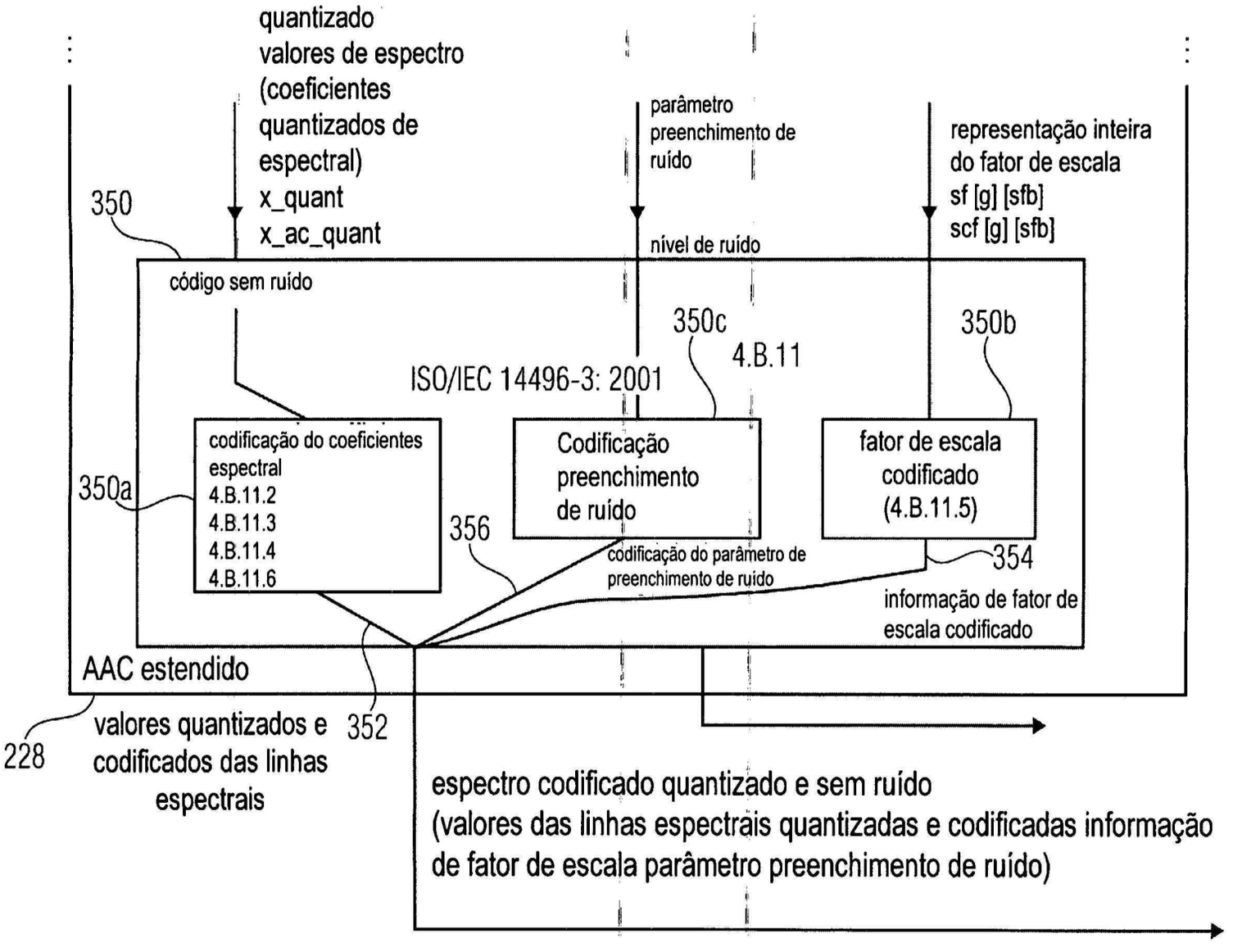

quantizado
valores de espectro
(coeficientes
quantizados de
espectral)
x_quant
x_ac_quant
parâmetro
preenchimento de
ruído
representação inteira
do fator de escala
sf [g] [sfb]
scf [g] [sfb]
350
nível de ruído
código sem ruído
350c
4.B.11
ISO/IEC 14496-3: 2001
350b
codificação do coeficientes
espectral
4.B.11.2
4.B.11.3
4.B.11.4
4.B.11.6
Codificação
preenchimento
de ruído
fator de escala
codificado
(4.B.11.5)
350a
356
codificação do parâmetro de
preenchimento de ruído
354
informação de fator de
escala codificado
AAC estendido
228
valores quantizados e
codificados das linhas
espectrais
352
espectro codificado quantizado e sem ruído
(valores das linhas espectrais quantizadas e codificadas informação
de fator de escala parâmetro preenchimento de ruído)

FIG 4A

line number

Encoder:

```
1   Calculate Mean Quantization error:
2   nLines = 0;
3   avgError = 0;

4   for (band = all scale factor bands) {
5     for(line = all spectral lines in band) {
6       if(band not quantized to zero) {
7         avgError +=fabs ( pow(line, 0.75)*scale factor - (int) pow(line, 0.75)*scale factor)
8         nLines ++;
9       }
10    }
11  }
12  avgError=avgError / nLines;

13  noiseLevel=(int) (14+4*ld(avgError);
14  noiseLevel=max(0, min(7, noiseLevel)

15  Calculate All zero Scale Factor:
16  noiseValue=pow(2.f, ((float)(noiseLevel)-14.f)/4.f)
17  if(noiseLevel>0) {
18  for (band=all scale factor bands){
19    if(band quantized to zero) {
20      scf[band]=(INT) (2.f * log( ((energy)/((float)sfbWidth*noiseVal*noiseVal*noiseVal)))/log(2.f));
21    }
22  }
23  else {
24    scf=don't care
25  }
```

non-quantized, scaled spectral line magnitude value

quantized, scaled spectral line magnitude value

Opcional: quantização nível de ruído

Opcional: quantização inversa nível de ruído

```
Quantização nível de ruído
noiseLevel = (int) (14 + 4*ld(meanLineError)
noiseLevel = Max(0,min(7,noiseLevel)
```

FIG 4B

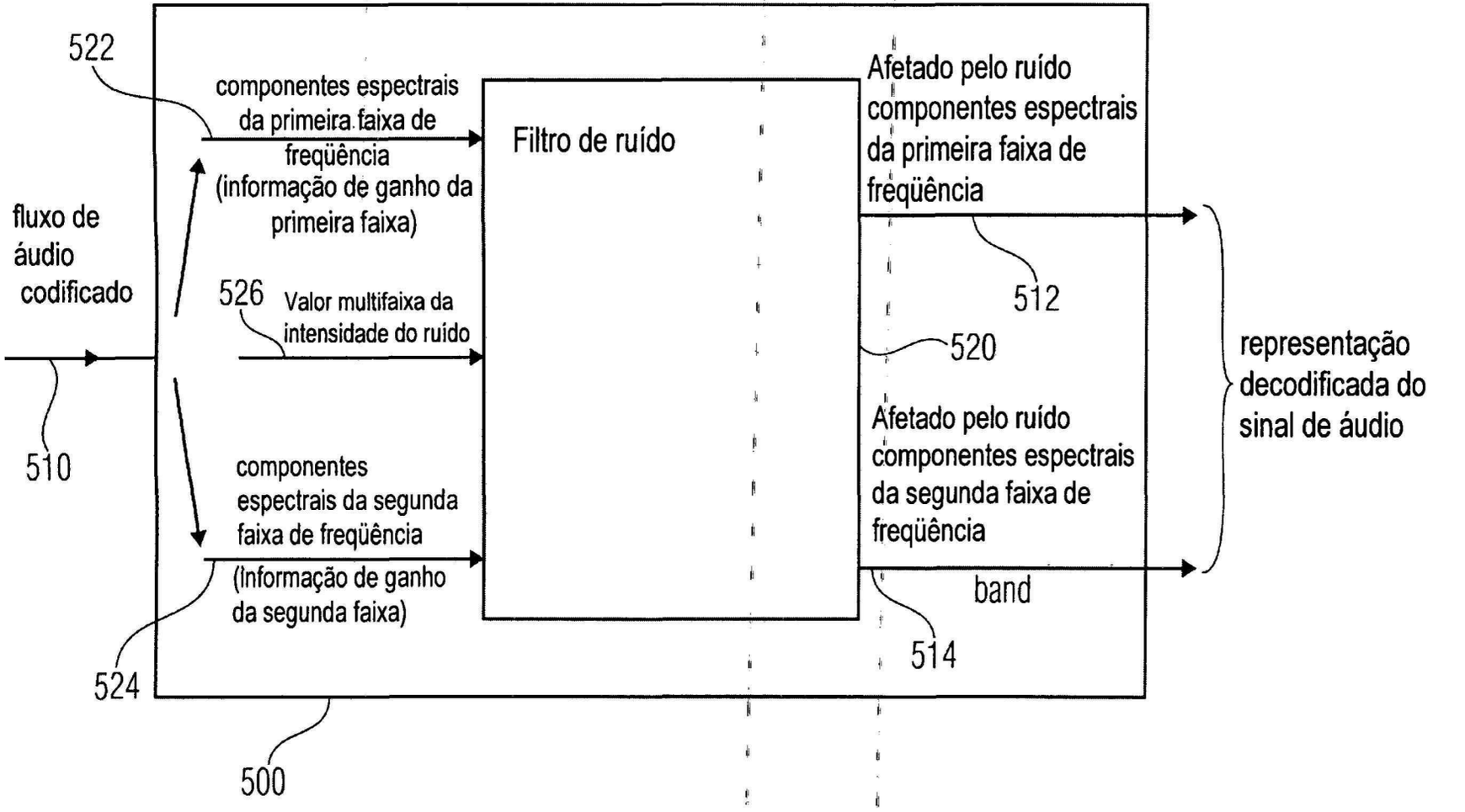

522
componentes espectrais da primeira faixa de freqüência (informação de ganho da primeira faixa)
fluxo de áudio codificado
526
Valor multifaixa da intensidade do ruído
componentes espectrais da segunda faixa de freqüência (Informação de ganho da segunda faixa)
510
524
500
Filtro de ruído
Afetado pelo ruído componentes espectrais da primeira faixa de freqüência
512
520
Afetado pelo ruído componentes espectrais da segunda faixa de freqüência
band
514
representação decodificada do sinal de áudio

FIG 5

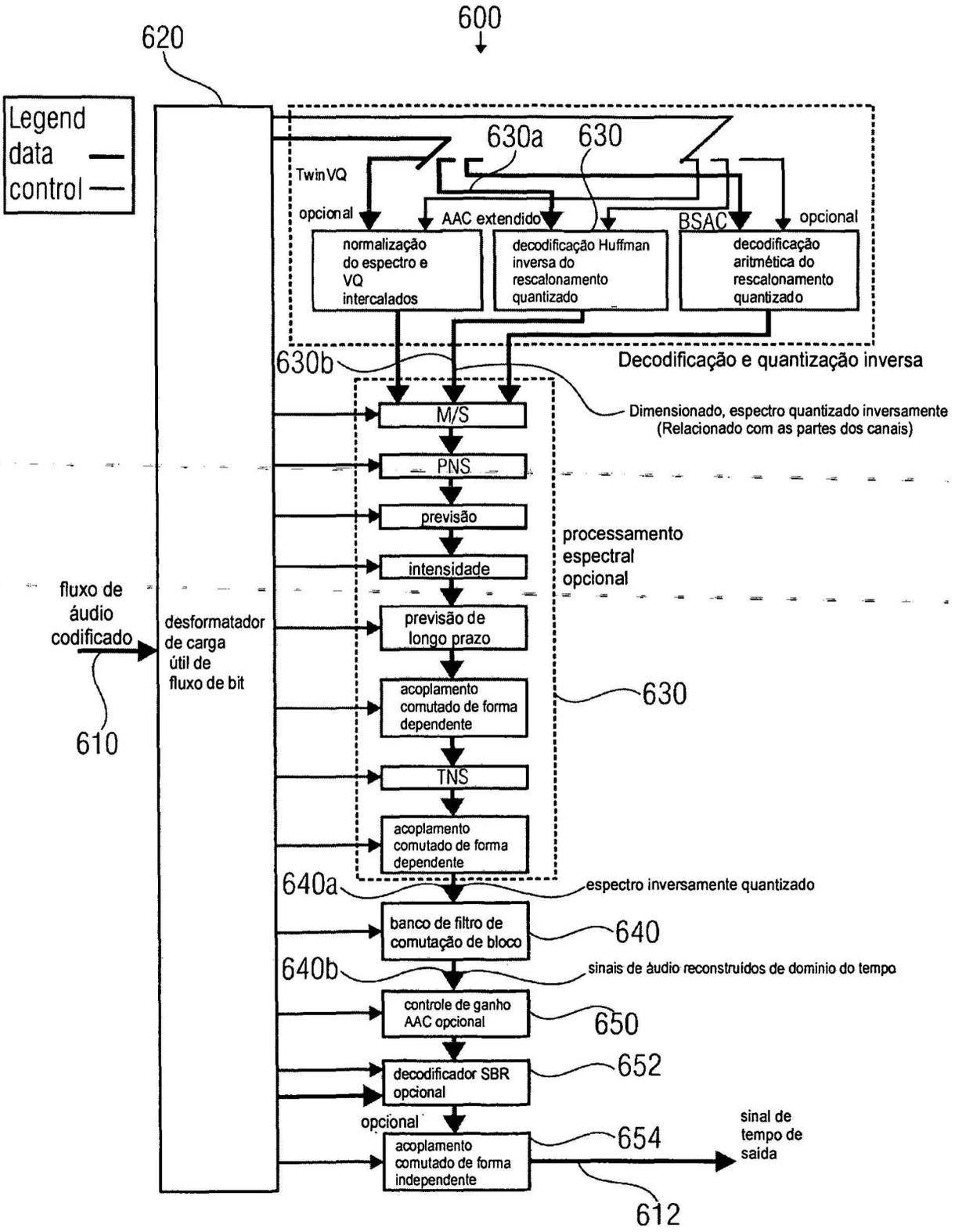

600
620
Legend
data
control
630a
630
TwinVQ
opcional
AAC extendido
BSAC
opcional
normalização do espectro e VQ intercalados
decodificação Huffman inversa do rescalonamento quantizado
decodificação aritmética do rescalonamento quantizado
Decodificação e quantização inversa
630b
Dimensionado, espectro quantizado inversamente (Relacionado com as partes dos canais)
M/S
PNS
previsão
intensidade
processamento espectral opcional
previsão de longo prazo
acoplamento comutado de forma dependente
TNS
acoplamento comutado de forma dependente
630
fluxo de áudio codificado
desformatador de carga útil de fluxo de bit
610
640a
espectro inversamente quantizado
banco de filtro de comutação de bloco
640
640b
sinais de áudio reconstruídos de domínio do tempo
controle de ganho AAC opcional
650
decodificador SBR opcional
652
opcional
acoplamento comutado de forma independente
654
sinal de tempo de saída
612

FIG 6

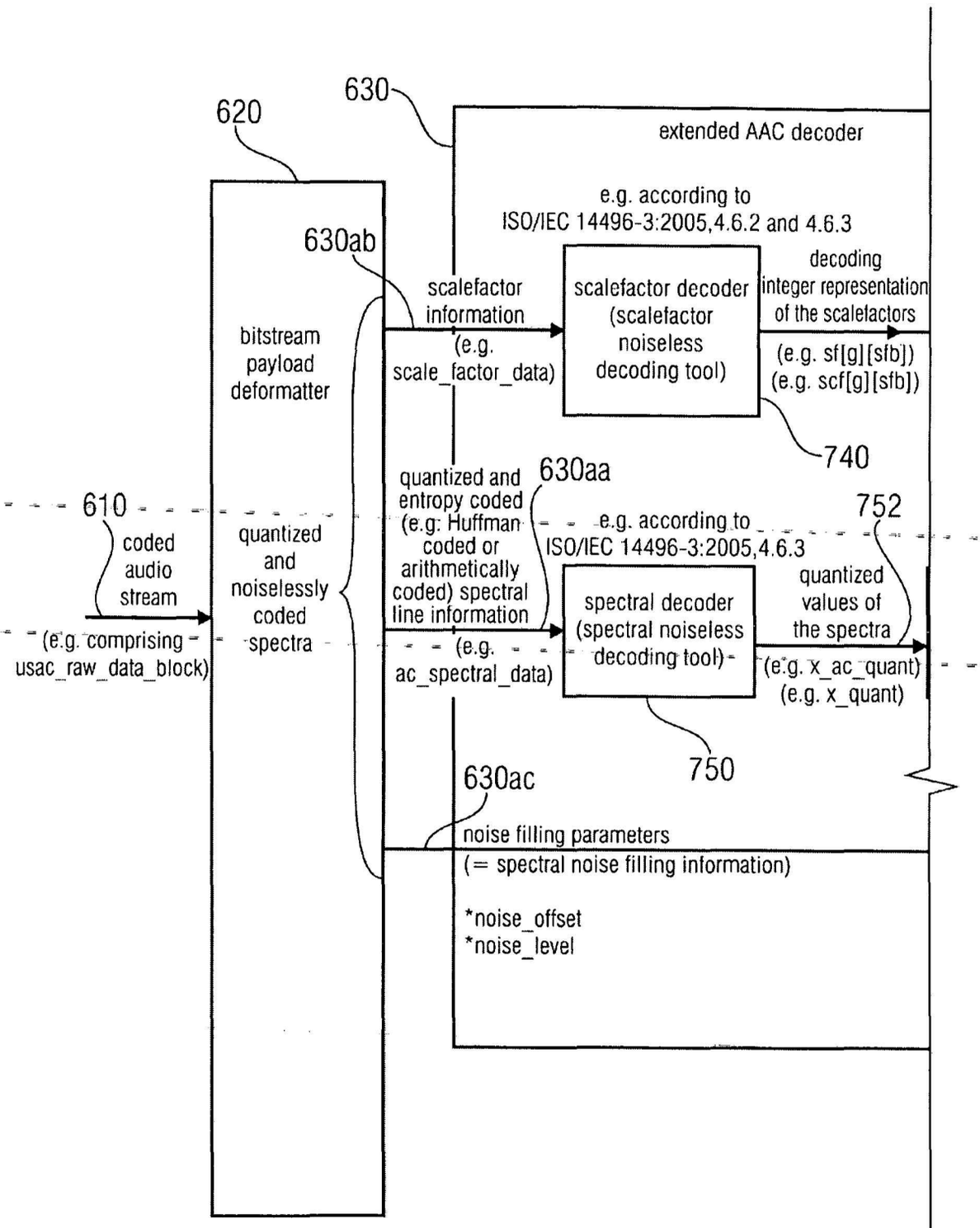


FIG 7A

FIG 7B

FIG 7A

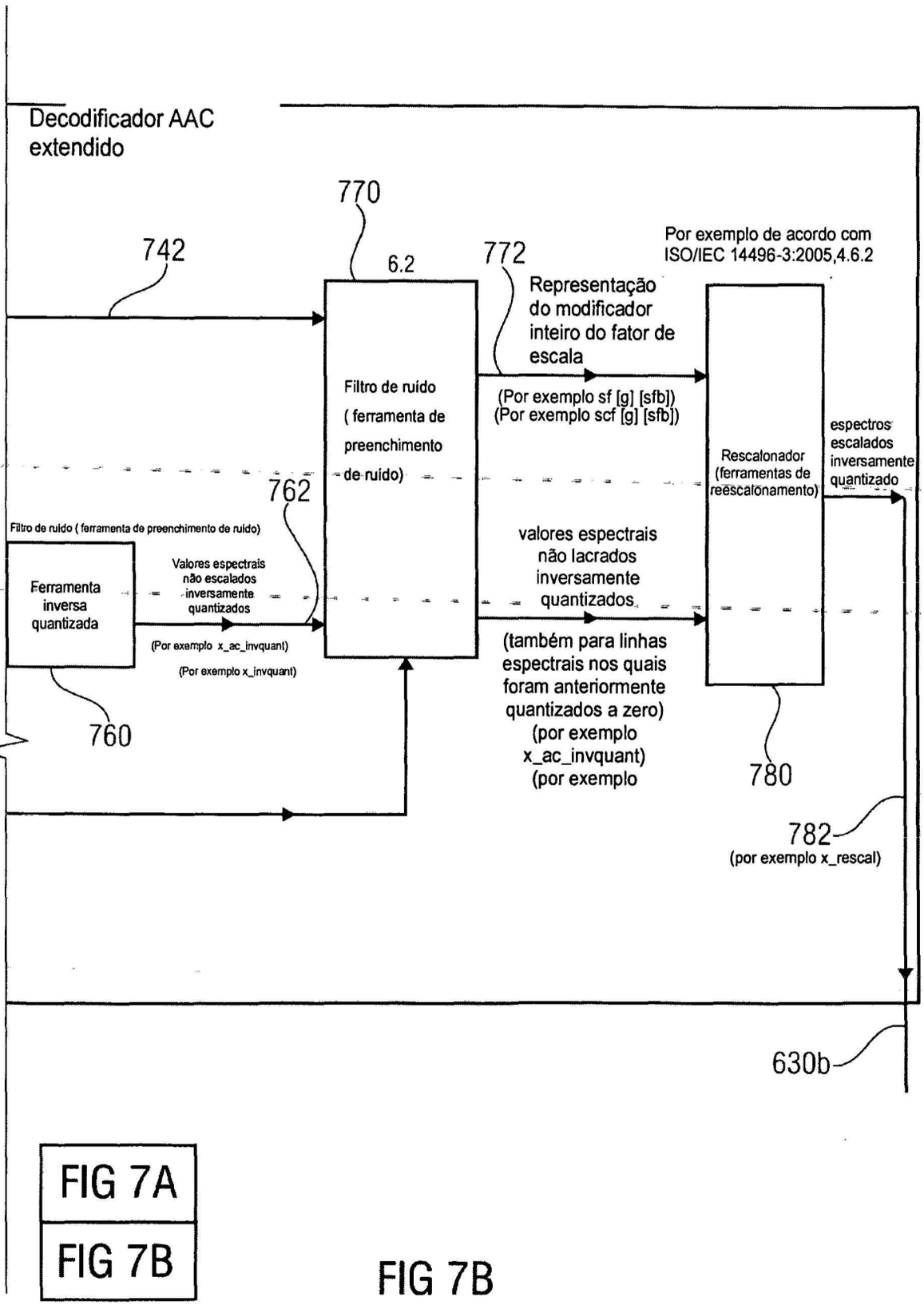


FIG 7A

FIG 7B

FIG 7B

$$x_invquant = Sign(x_quant) \cdot |x_quant|^{\frac{4}{3}}$$

or:

$$x_ac_invquant = Sign(x_ac_quant) \cdot |x_ac_quant|^{\frac{4}{3}}$$

FIG 8A

```
for (g = 0; g< num_window_groups; g++) {
  for (sfb = 0; sfb < max_sfb; sfb++) {
    width = (swb_offset [sfb+1] - swb_offset [sfb]);
    for (win - 0; win < window_group_len[g]; win++) {
      for (bin = 0; bin < width; bin++) {
        x_ac_invquant [g] [sfb] [sfb] [bin] =
sign(x_ac_quant[g] [win] [sfb] [bin])*abs(x_ac_quant[g] [win] [sfb] [bin]) ^ (4/3);
      }
    }
  }
}
```

FIG 8B

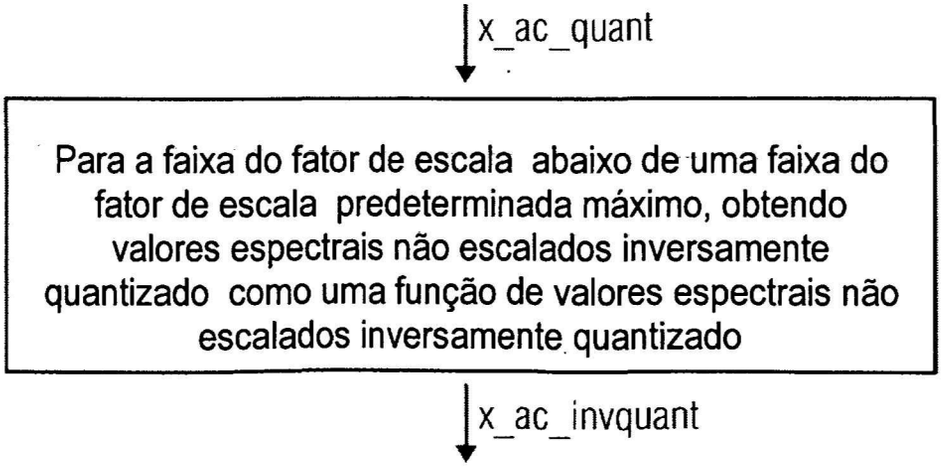


FIG 8C

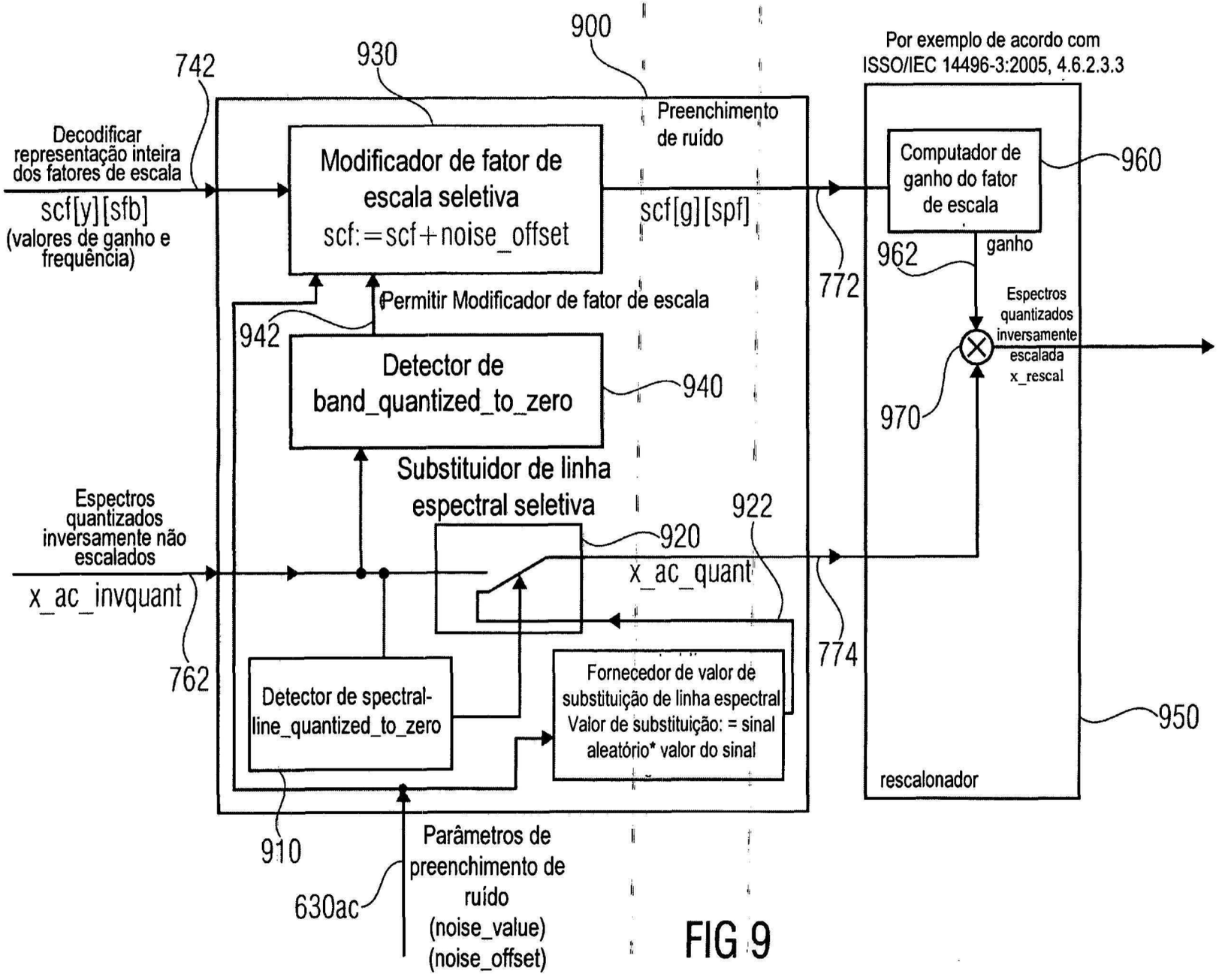

900
930
742
Por exemplo de acordo com
ISSO/IEC 14496-3:2005, 4.6.2.3.3
Preenchimento
de ruído
Decodificar
representação inteira
dos fatores de escala
scf[y][sfb]
(valores de ganho e
frequência)
Modificador de fator de
escala seletiva
scf:=scf+noise_offset
scf[g][spf]
Computador de
ganho do fator
de escala
960
962
ganho
772
Permitir Modificador de fator de escala
942
Detector de
band_quantized_to_zero
940
Espectros
quantizados
inversamente
escalada
x_rescal
970
Substituidor de linha
espectral seletiva
Espectros
quantizados
inversamente não
escalados
x_ac_invquant
920
922
x_ac_quant
762
774
Detector de spectral-
line_quantized_to_zero
Fornecedor de valor de
substituição de linha espectral
Valor de substituição: = sinal
aleatório* valor do sinal
950
rescalonador
910
630ac
Parâmetros de
preenchimento de
ruído
(noise_value)
(noise_offset)

FIG 9

FIG 10A

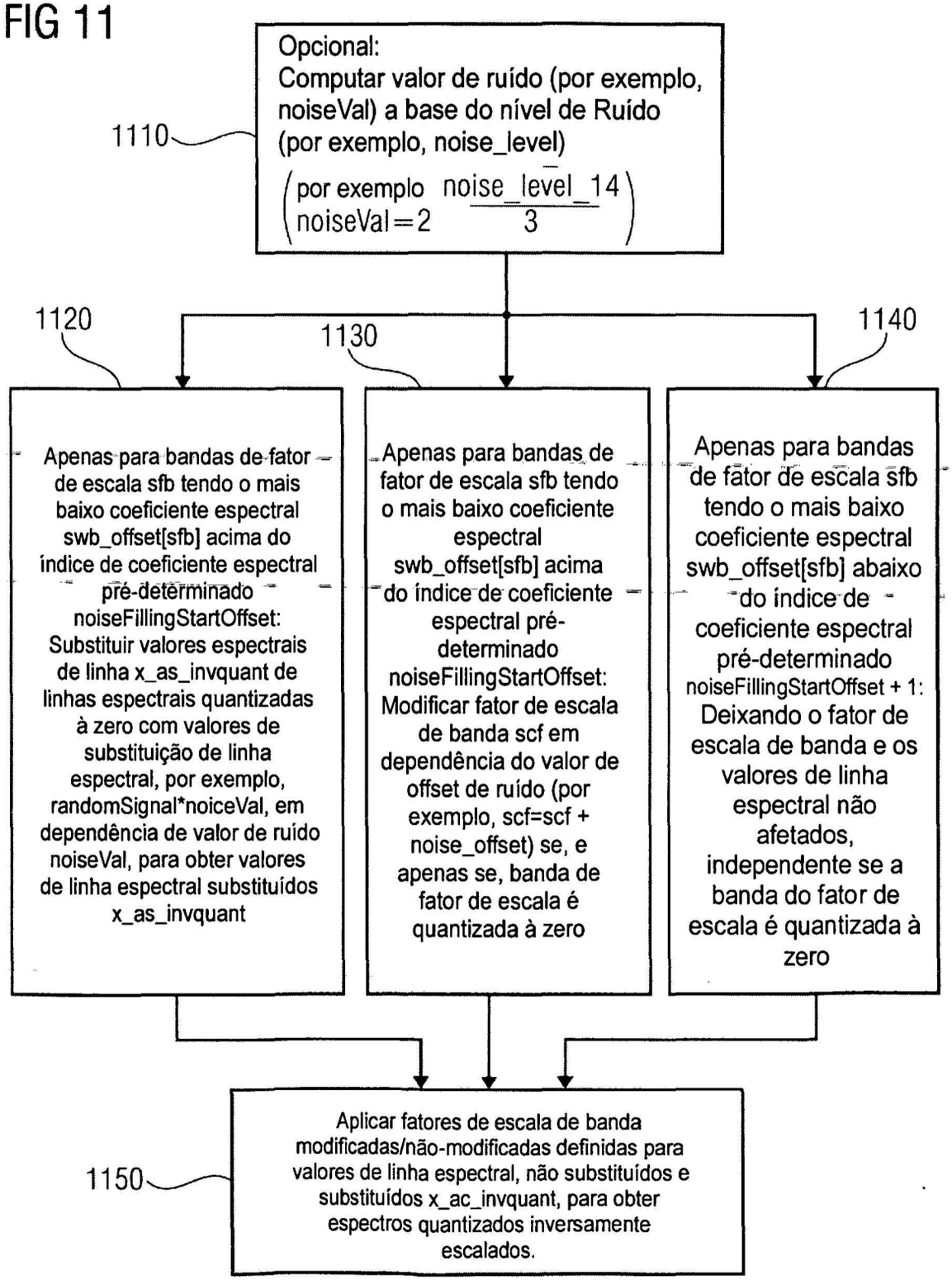
```
1  if (noise_level !=0) {
2    noiseVal = pow(2, (noise_level-14)/3) ;
3    noise_offset = noise_offset - 16;
4  }
5  else {
6    noiseVal = 0;
7    noise_offset = 0;
8  }
9  for (g=0; g<num_window_groups; g++) {
10   for (sfb=0; sfb < max_sfb; sfb++) {
11     band_quantized_to_zero=1;                    -- assume a band is quantized to zero
12     width=(swb_offset [sfb+1] - swb_offset [sfb]);
13     if(swb_offset [sfb] > noiseFillingStartOffset) {
14       for (win=0); win < window_group_len[g]; win++) {
15         for (bin=0; bin < width; bin++) {
16           if (x_ac_invquant [g] [win] [sfb] [bin] ==0) {
17             x_ac_invquand [g] [win] [sfb] [bin] = randomSign () *noiseVal;   -- for scalefactor bands, starting above noiseFillingStartOffset,
18           }                                                                  -- add noise of amplitude noiseVal to spectral lines quantized to zero;
19           else {
20             band_quantized_to_zero=0;          -- for scalefactor bands, starting above noiseFillingStartOffset, if a single
21           }                                    -- bin of a scalefactor band is≠0, then band is not quantized to zero;
22         }
23       }
24     }
24a    else {
24b      band_quantized_to_zero=0;                -- for scalefactor bands, starting below noiseFillingStartOffset, if it is
24c    }                                          -- always assumed thet the band is not quantized to zero;

25     if(band_quantized_to_zero) {
26       scf[g] [sfb] = scf[g] [sfb] + noise_offset;   -- for scalefactor bands quantized to zero, modify band scalefactor in
27     }                                               -- dependance on noise offset value
28   }
29 }
```

Elementos de Dados	
noise_offset	Offset adicional para modificar o fator de escala das bandas quantizadas à zero
noise_level	Integral representando o ruído de quantização para ser adicionado para cada linha espectral quantizada à zero
Elementos de Ajuda	
x_ac_invquant[g][win][sfb][bin]	Coeficiente espectral AAC para grupo g, janela win, banda de fator de escala sfb, coeficiente Bin após quantização inversa
noiseFillingStartOffset[win]	Um offset geral ou frequência de início de preenchimento de ruído. O offset é definido para 20 para curto (window_sequence==EIGHT_SHORT_SEQUENCE) e 160 outro.
noiseVal	O Valor de ruído absoluto que substitui cada Bin quantizado à zero
randomSign	Sinal aleatório (-1,1) multiplicado para noiseVal
band_quantized_to_zero[	Bandeira para sinalizar se um sfb é completamente quantizado à zero
swb_offset[sfb]	Índice do mais baixo coeficiente espectral de banda de fator de escala sfb
num_window_groups	Número de grupos de janelas que dividem um conjunto de fatores de escala
mux_sfb	Número de banda de fatores de escala por grupo
window_group_len	Número de janelas em cada grupo
g	Índice de grupo
win	Índice de janela dentro do grupo
sfb	Índice de banda do fator de escala dentro do grupo
swb	Banda de janela do fator de escala dentro da janela
bin	Índice de coeficiente
num_windows	Número de janelas da sequência real de janela

FIG 10B

FIG 11


Opcional:
Computar valor de ruído (por exemplo, noiseVal) a base do nível de Ruído (por exemplo, noise_level)
por exemplo noiseVal = 2 noise_level_14 / 3
1110
1120
1130
1140
Apenas para bandas de fator de escala sfb tendo o mais baixo coeficiente espectral swb_offset[sfb] acima do índice de coeficiente espectral pré-determinado noiseFillingStartOffset: Substituir valores espectrais de linha x_as_invquant de linhas espectrais quantizadas à zero com valores de substituição de linha espectral, por exemplo, randomSignal*noiceVal, em dependência de valor de ruído noiseVal, para obter valores de linha espectral substituídos x_as_invquant
Apenas para bandas de fator de escala sfb tendo o mais baixo coeficiente espectral swb_offset[sfb] acima do índice de coeficiente espectral pré-determinado noiseFillingStartOffset: Modificar fator de escala de banda scf em dependência do valor de offset de ruído (por exemplo, scf=scf + noise_offset) se, e apenas se, banda de fator de escala é quantizada à zero
Apenas para bandas de fator de escala sfb tendo o mais baixo coeficiente espectral swb_offset[sfb] abaixo do índice de coeficiente espectral pré-determinado noiseFillingStartOffset + 1: Deixando o fator de escala de banda e os valores de linha espectral não afetados, independente se a banda do fator de escala é quantizada à zero
Aplicar fatores de escala de banda modificadas/não-modificadas definidas para valores de linha espectral, não substituídos e substituídos x_ac_invquant, para obter espectros quantizados inversamente escalados.
1150

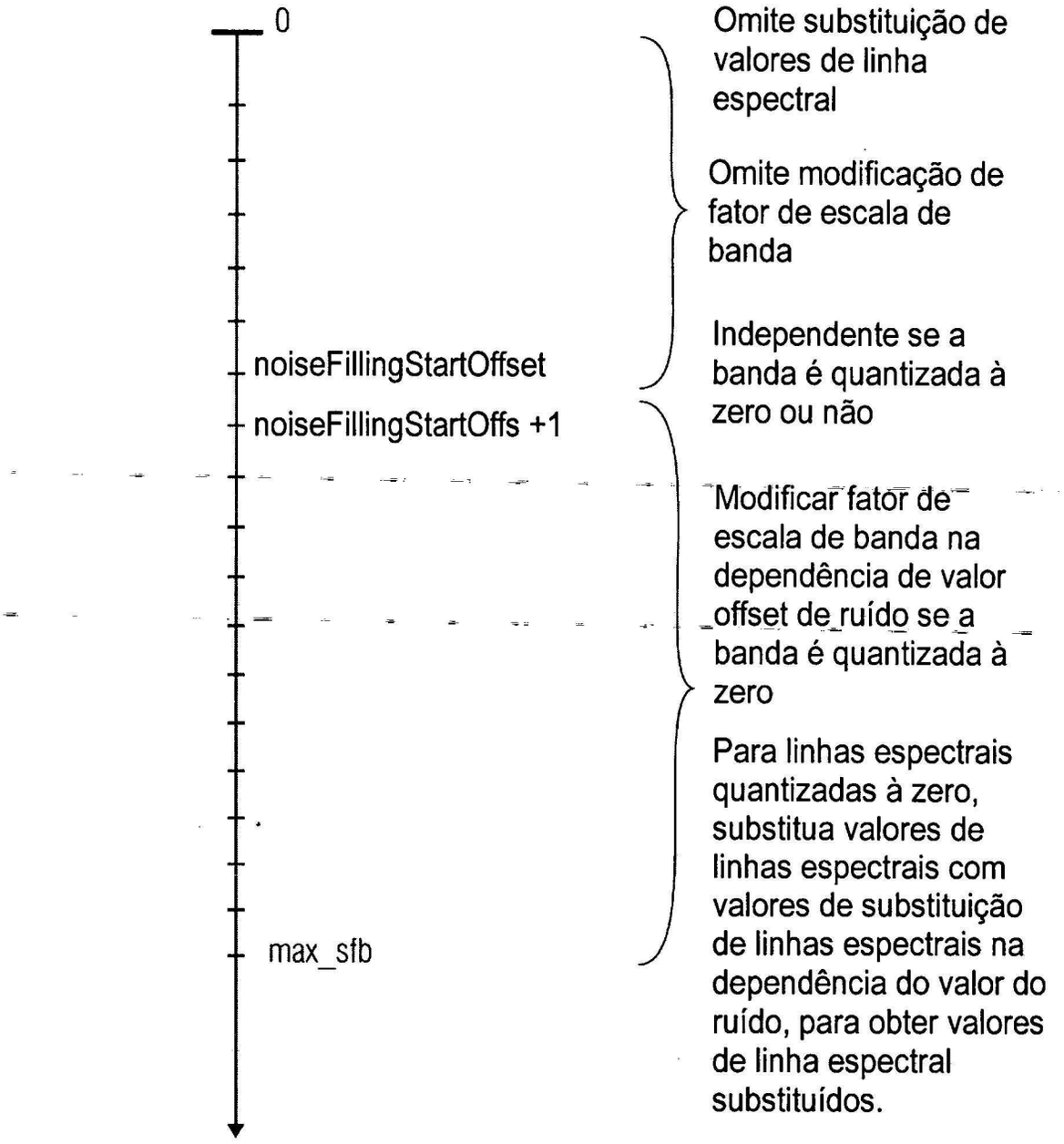

0
noiseFillingStartOffset
noiseFillingStartOffs +1
max_sfb
Omite substituição de valores de linha espectral
Omite modificação de fator de escala de banda
Independente se a banda é quantizada à zero ou não
Modificar fator de escala de banda na dependência de valor offset de ruído se a banda é quantizada à zero
Para linhas espectrais quantizadas à zero, substitua valores de linhas espectrais com valores de substituição de linhas espectrais na dependência do valor do ruído, para obter valores de linha espectral substituídos.
Bin de início da banda de fator de escala

FIG 12

Número de linha	Valor de ruído no decodificador
1	if(noiseLevel !=0)
2	noiseValue=pow(2.f((float)(noiseLevel)-14.f)/4.f)
3	diversamente
4	noiseVal=0

FIG 13A

Decodificador

Número de linha	
	"Zero Replacement Value" Calculation:
1	if(noiseLevel !=0)
2	noiseValue=pow(2.f((float)(noiseLevel)-14.f)/4.f)
3	else
4	noiseVal=0
5	for(band=all scale factor bands) {
6	for(line=all spectral lines in band) {
7	if(band quantized to zero) {
8	scf=scf + noise_Offset {
9	}
10	if(line>noiseFillingStartOffset) {
11	if(quantizedSpec[line]==0) {
12	quantizedSpec[line]=randomSign()*noiseValue;
13	}
14	}
15	}
16	}

FIG 13B

USAC bitstream payload

```
usac_raw_data_block ()
{
  single_channel_element () ; and/or
  channel_pair_element () ;
  optional: additional channel elements
}
```

FIG 14A

```
single_channel_element ()                                    (Tab. 4.4)
{
  fd_channel_stream (*,*, noise Filling)
}
```

FIG 14B

```
channel_pair_element                                         (Tab. 4.5)
{
  fd_channel_stream (*,*, noise Filling)(1st channel); and/or
  fd_channel_stream (*,*, noise Filling)(2nd channel)
}
```

FIG 14C

```
fd_channel_stream ()                                         (Tab. 4.8)
{
  global_gain;                    e.g. 8 bit
  noise_offset;                   e.g. 3 bit
  noise_level;                    e.g. 5 bit

  scale_factor_data ();

  tus_data ();                    optional

  ac_spectral_data ()
}
```

FIG 14D

Sintaxe de individual_channel_stream()

Sintaxe	Nº de bits	Mnemônica
individual_channel_stream(common_window)		
{		
global_gain;	8	uimsbf
noise_Offset	5	uimsbf
noise_Level	3	uimsbf
if (!common_window)		
ics_info();		
section_data();		
scale_factor_data();		
pulse_data_present;	1	uimsbf
if (pulse_data_present) {		
pulse_data();		
}		
tns_data_present;	1	uimsbf
if(tns_data_present){		
tns_data();		
}		
gain_control_data_present;	1	uimsbf
if (gain_control_data_present) {		
gain_control_data();		
}		
spectral_data()		
}		

FIG 15

RESUMO

"CODIFICADOR DE ÁUDIO, DECODIFICADOR DE ÁUDIO, MÉTODOS PARA CODIFICAÇÃO E DECODIFICAÇÃO DE UM SINAL DE ÁUDIO, E UM PROGRAMA DE COMPUTADOR"

Um codificador para fornecer um fluxo de áudio com base em uma representação de domínio transformada de um sinal de áudio de entrada inclui uma calculadora de erro de quantização configurada para determinar um erro de quantização multi-banda em uma pluralidade de faixas de frequência do sinal de áudio de entrada para qual a banda informações de ganho de banda separada está disponível. O codificador inclui também um fornecedor de fluxo de áudio configurado para fornecer o fluxo de áudio de forma que o fluxo de áudio é composto por uma informação que descreve um conteúdo de áudio das faixas de frequências e uma informação que descreve o erro de quantização multi-banda. Um decodificador para fornecer uma representação decodificada de um sinal de áudio a partir de um fluxo de áudio codificado que representa os componentes espectrais de bandas de frequência do sinal de áudio que compreende um preenchedor de ruído configurado para introduzir ruído em componentes espectrais de uma pluralidade de faixas de frequência nos quais a informação de ganho de frequência separada é associada com base em um valor de intensidade de ruído de multi-banda comum.