



República Federativa do Brasil

Ministério do Desenvolvimento, Indústria,
Comércio e Serviços

Instituto Nacional da Propriedade Industrial

(11) PI 0910790-8 B1

(22) Data do Depósito: 06/07/2009

(45) Data de Concessão: 07/05/2024

(54) Título: PROVEDOR DO SINAL DE ATIVAÇÃO DE CURVA TEMPORAL, CODIFICADOR DO SINAL DE ÁUDIO, MÉTODO PARA FORNECER UM SINAL DE ATIVAÇÃO DE CURVA TEMPORAL E MÉTODO PARA CODIFICAR UM SINAL DE ÁUDIO

(51) Int.Cl.: G10L 19/10; G10L 21/04.

(30) Prioridade Unionista: 11/07/2008 US 61/079,873.

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

(72) Inventor(es): STEFAN BAYER; SASCHA DISCH; RALF GEIGER; GUILLAUME FUCHS; MAX NEUENDORF; GERALD SCHULLER; BERND EDLER.

(86) Pedido PCT: PCT EP2009004874 de 06/07/2009

(87) Publicação PCT: WO 2010/003618 de 14/01/2010

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

(57) Resumo: PROVEDOR DO SINAL DE ATIVAÇÃO DE CURVA TEMPORAL, CODIFICADOR DO SINAL DE ÁUDIO, MÉTODO PARA FORNECER UM SINAL DE ATIVAÇÃO DE CURVA TEMPORAL, MÉTODO PARA CODIFICAR UM SINAL DE , ÁUDIO E PROGRAMAS DE COMPUTADOR. Um codificador de áudio compreende um controlador da função de janela (504), um mecanismo de janela (502), um mecanismo de curva temporal (506) com uma funcionalidade de verificação da qualidade final, um conversor de tempo/frequência (508), uma etapa TNS (510) ou um codificador do quantizador (512), o controlador da função de janela (504), o mecanismo de curva temporal (506), a etapa TNS (510) ou um analisador de preenchimento de ruído adicional (524) são controlados pelos resultados da análise do sinal obtidos por um analisador de curva temporal (516) ou um classificador de sinais (520). Além disso, um decodificador aplica uma operação de preenchimento de ruído usando uma estimativa de preenchimento de ruído manipulado dependendo de uma característica harmônica ou de fala do sinal de áudio.

“PROVEDOR DO SINAL DE ATIVAÇÃO DE CURVA TEMPORAL,
CODIFICADOR DO SINAL DE ÁUDIO, MÉTODO PARA FORNECER UM SINAL DE
ATIVAÇÃO DE CURVA TEMPORAL E MÉTODO PARA CODIFICAR UM SINAL DE
ÁUDIO”

5 ESPECIFICAÇÃO

A presente invenção está relacionada a uma codificação e decodificação de áudio e, especificamente, à codificação/decodificação do sinal de áudio tendo um conteúdo harmônico ou de fala, que pode estar sujeito a um processamento de

10 curva temporal.

A seguir, será fornecida uma breve introdução no campo de codificação de áudio com curva temporal, conceitos que podem ser aplicados em conjunto com algumas das materializações da invenção.

Recentemente, foram desenvolvidas técnicas para transformar um sinal de áudio em uma representação de domínio de frequência e para codificar de forma eficiente a representação desse domínio de frequência, levando-se em consideração, por exemplo, limites de mascaramento perceptuais. Esse conceito de codificação de sinal de áudio será eficiente principalmente se o comprimento do bloco, para o qual um conjunto de coeficientes espectrais codificados for transmitido por extenso e se apenas um número comparativamente pequeno de coeficientes espectrais estiver bem acima do limite de mascaramento global, enquanto um número maior de coeficientes espectrais estiver próximo ou abaixo do limite de mascaramento global e, portanto, poderá ser negligenciado (ou codificado com um comprimento de código mínimo).

Por exemplo, transformações sobrepostas moduladas

baseadas em cosseno ou seno são, com frequência, usadas nas aplicações para a codificação de origem devido às suas propriedades de compactação de energia. Ou seja, para tons harmônicos com frequências fundamentais constantes (tom), eles concentram a energia do sinal em um número baixo de componentes espectrais (bandas secundárias), o que acarreta a uma representação de sinal eficiente.

Em geral, o tom (fundamental) de um sinal deverá ser compreendido como a frequência dominante mais baixa diferenciável a partir do espectro do sinal. No modelo de fala comum, o tom é a frequência do sinal de excitação modulado pela garganta humana. Se somente uma única frequência fundamental estivesse presente, o espectro seria extremamente simples, compreendendo a frequência fundamental e os sobretons apenas.

Tal espectro poderia ser codificado de forma altamente eficiente. Para sinais com tom variável, contudo, a energia correspondente a cada componente harmônico é distribuída em vários coeficientes de transformação, acarretando, portanto, uma redução da eficiência da codificação.

Para solucionar essa redução da eficiência de codificação, o sinal de áudio a ser codificado é ressampleado efetivamente em uma grade temporal não uniforme. No processamento subsequente, as posições de amostragem obtidas pelo ressampleamento não uniforme serão processadas se representarem valores em uma grade temporal uniforme. Essa operação é normalmente denotada pela frase "curva temporal". Os períodos de amostragem poderão ser selecionados de modo vantajoso dependendo da variação temporal do passo, de forma que uma variação do passo

na versão com curva temporal do sinal de áudio seja menor do que uma variação do passo na versão original do sinal de áudio (antes da curva temporal). Essa variação do passo também poderá ser denotada com a frase "contorno de curva temporal". Após a curva

5 temporal do sinal de áudio, a versão com curva temporal do sinal de áudio será convertida no domínio de frequência. A curva temporal dependente de passo tem o efeito que a representação de domínio de frequência do sinal de áudio com curva temporal geralmente exibe uma compactação de energia em um número muito

10 menor de componentes espectrais do que uma representação de domínio de frequência do sinal de áudio original (sem curva temporal).

No lado do decodificador, a representação de domínio de frequência do sinal de áudio com curva temporal é

15 convertido de volta para o domínio de tempo, de forma que uma representação do domínio de tempo do sinal de áudio com curva temporal esteja disponível no lado do decodificador. No entanto, na representação de domínio de tempo do sinal de áudio com curva temporal reconstruído no decodificador, as variações de passo

20 original do sinal de áudio de entrada do codificador não estão incluídas. De forma correspondente, outra curva temporal por meio da resamostragem da representação do domínio de tempo reconstruído pelo decodificador do sinal de áudio com curva temporal é aplicada. Para obter uma boa reconstrução do sinal de áudio de

25 entrada do codificador no decodificador, é desejável que a curva temporal do decodificador seja pelo menos aproximadamente a operação inversa com relação à curva temporal do decodificador. Para obter uma curva temporal apropriada, é desejável ter

informações disponíveis no decodificador que permitam um ajuste da curva temporal do decodificador.

Conforme é normalmente exigido para transferir tal informação do codificador de sinal de áudio para o
5 decodificador de sinal de áudio, é desejável manter uma taxa de bits pequena exigida para essa transmissão enquanto ainda se permite uma reconstrução confiável das informações exigidas de curva temporal no decodificador.

Na visualização da discussão acima, há um desejo
10 de criar um conceito que permita uma aplicação eficiente da taxa de bits do conceito de curva temporal em um codificador de áudio.

Trata-se de um objeto da invenção para criar
conceitos para aprimorar a impressão de audição fornecida por um
sinal de áudio codificado a partir das informações disponíveis em
15 um codificador de sinal de áudio de curva temporal ou em um decodificador de sinal de áudio de curva temporal.

Esse objeto é alcançado por meio de um provedor
de sinal de ativação de curva temporal ao fornecer um sinal de
ativação de curva temporal a partir de uma representação de um
20 sinal de áudio de acordo com a reivindicação 1, um codificador do
sinal de áudio para codificação de um sinal de áudio de entrada de
acordo com a reivindicação 12, um método para fornecer um sinal de
ativação de curva temporal de acordo com a reivindicação 14, um
método para fornecer uma representação codificada de um sinal de
25 áudio de entrada de acordo com a reivindicação 15 ou um programa
de COMPUTADOR de acordo com a reivindicação 16.

Trata-se de outro objeto da presente invenção
para fornecer um esquema de codificação/decodificação de áudio

aprimorado, que fornece uma qualidade mais elevada ou uma taxa de bits mais baixa.

Esse objeto é alcançado por meio de um codificador de áudio de acordo com a reivindicação 17, 26, 32, 37, um decodificador de áudio de acordo com a reivindicação 20, um método de codificação de áudio de acordo com a reivindicação 23, reivindicação 30, reivindicação 35 ou reivindicação 37, um método de decodificação de acordo com a reivindicação 24 ou um programa de computador de acordo com a reivindicação 25, 31, 36 ou 43.

As materializações de acordo com a invenção estão relacionadas a métodos para um codificador de transformação MDCT com curva temporal. Algumas materializações estão relacionadas a ferramentas somente do codificador. No entanto, outras materializações também estão relacionadas a ferramentas do decodificador.

Uma materialização da invenção cria um provedor do sinal de ativação de curva temporal para fornecer um sinal de ativação de curva temporal a partir de uma representação de um sinal de áudio. O provedor do sinal de ativação de curva temporal compreende um provedor de informações de compactação de energia configurado para fornecer informações de compactação de energia descrevendo uma compactação da energia em uma representação de espectro transformada com curva temporal do sinal de áudio. O provedor do sinal de ativação de curva temporal também compreende um comparador configurado para comparar as informações de compactação de energia com um valor de referência e para fornecer o sinal de ativação de curva temporal dependendo de um resultado da comparação.

Essa materialização se baseia na descoberta de que o uso de uma funcionalidade de curva temporal em um codificador de sinal de áudio normalmente traz consigo uma melhoria no sentido de uma redução da taxa de bits do sinal de áudio codificado se a representação do espectro transformada com curva temporal do sinal de áudio compreender uma distribuição de energia suficientemente compacta de forma que a energia esteja concentrada em uma ou mais regiões espectrais (ou linhas espectrais). Isso se deve ao fato de uma curva temporal bem-sucedida trazer consigo o efeito de diminuir a taxa de bits ao transformar um espectro obscurecido, por exemplo, de uma estrutura de áudio, em um espectro tendo um ou mais picos diferenciáveis e, consequentemente, tendo uma compactação de energia mais alta do que o espectro do sinal de áudio original (sem curva temporal).

Com relação a esse problema, deve-se entender que uma estrutura do sinal de áudio, durante a qual o passo do sinal de áudio varia bastante, compreende um espectro obscurecido. O passo de variação de tempo do sinal de áudio tem o efeito de que uma transformação de domínio tempo para domínio de frequência desempenhada na estrutura do sinal de áudio resulta em uma distribuição obscurecida da energia do sinal sobre a frequência, principalmente na região de frequência mais elevada. De forma correspondente, uma representação do espectro de tal sinal de áudio original (sem curva temporal) compreende uma compactação de energia baixa e normalmente não exhibe picos espectrais em uma parte do espectro de frequência superior ou exhibe apenas picos espectrais relativamente pequenos na parte do espectro de frequência mais elevada. Em contrapartida, se a curva temporal for

bem-sucedida (em termos de fornecer uma melhoria da eficiência de codificação), a curva temporal do sinal de áudio original gerará um sinal de áudio com curva temporal tendo um espectro com picos relativamente mais elevados e mais claros (principalmente na parte

5 de frequência mais elevada do espectro). Isso se deve ao fato de um sinal de áudio ter um passo de variação de tempo sendo transformado em um sinal de áudio com curva temporal tendo uma variação de passo menor ou até um passo aproximadamente constante. Consequentemente, a representação do espectro do sinal de áudio

10 com curva temporal (que pode ser considerada como uma representação do espectro transformada com curva temporal do sinal de áudio) compreende um ou mais picos espectrais claros. Ou seja, o obscurecimento do espectro do sinal de áudio original (tendo um passo temporariamente variável) é reduzido por uma operação bem-

15 = sucedida de curva temporal, de forma que a representação do espectro transformada com curva temporal do sinal de áudio compreenda a compactação de energia mais elevada do que o espectro do sinal de áudio original. Não obstante, a curva temporal nem sempre é bem-sucedida no aprimoramento da eficiência da

20 codificação. Por exemplo, a curva temporal não aprimorará a eficiência de codificação se o sinal de áudio de entrada compreender componentes de ruído grandes ou se o contorno da curva temporal extraído for impreciso.

Tendo em vista esta situação, as informações de

25 compactação de energia fornecidas pelo provedor de informações de compactação de energia é um indicador valioso para determinar se a curva temporal é bem-sucedida em termos de reduzir a taxa de bits.

Uma materialização da invenção cria um provedor

do sinal de ativação de curva temporal para fornecer um sinal de ativação de curva temporal a partir de uma representação de um sinal de áudio. O provedor de ativação de curva temporal compreende dois provedores de representação de curva temporal configurados para fornecer duas representações de curva temporal do mesmo sinal de áudio usando diferentes informações de contorno de curva temporal. Portanto, os provedores de representação de curva temporal poderão ser configurados (estrutural e/ou funcionalmente) da mesma forma e usar o mesmo sinal de áudio, mas informações de contorno de curva temporal diferentes. O provedor do sinal de ativação de curva temporal também compreende dois provedores de informações de compactação de energia configurados para fornecer informações sobre a primeira compactação de energia a partir da primeira representação de curva temporal e para fornecer informações sobre a segunda compactação a partir da segunda representação de curva temporal. Os provedores de informações de compactação de energia poderão ser configurados da mesma forma, mas usam representações de curva temporal diferentes. Além disso, o provedor do sinal de ativação de curva temporal compreende um comparador para comparar as duas informações de compactação de energia diferentes e fornecer o sinal de ativação de curva temporal dependendo de um resultado da comparação.

Em uma materialização favorita, o provedor de informações de compactação de energia é configurado para fornecer uma medição do nivelamento espectral descrevendo a representação do espectro transformada com curva temporal do sinal de áudio como as informações de compactação de energia. Descobriu-se que a curva temporal é bem-sucedida, em termos de reduzir uma taxa de bits, se

ela transformar um espectro de um sinal de áudio de entrada em um espectro de curva temporal menos plano, representando uma versão com curva temporal do sinal de áudio de entrada. Da mesma forma, a medição do nivelamento espectral pode ser usada para determinar, sem desempenhar um processo de codificação espectral completo, se a curva temporal deverá ser ativada ou desativada.

Em uma materialização favorita, o provedor de informações de compactação de energia está configurado para computar um quociente de um significado geométrico do espectro de potência transformado com curva temporal e um significado aritmético do espectro de potência transformado com curva temporal para obter a medição do nivelamento espectral. Descobriu-se que esse quociente é uma medição do nivelamento espectral, que é bem adaptado para descrever as possíveis economias de taxa de bit que podem ser obtidas por uma curva temporal.

Em outra materialização favorita, o provedor de informações de compactação de energia é configurado para enfatizar uma parte de frequência mais elevada da representação do espectro transformada com curva temporal quando comparado a uma parte de frequência mais baixa da representação do espectro transformada com curva temporal para obter as informações de compactação de energia. Esse conceito se baseia na descoberta de que a curva temporal normalmente tem um impacto muito maior no intervalo de frequência mais elevada do que em um intervalo de frequência mais baixa. De forma correspondente, uma avaliação dominante do intervalo de frequência mais elevada é apropriada para determinar a eficiência da curva temporal usando uma medição do nivelamento espectral. Além disso, sinais de áudio típicos exibem um conteúdo

harmônico (compreendendo a harmônica de uma frequência fundamental) que diminui em intensidade com a frequência maior. Uma ênfase de uma parte de frequência mais elevada da representação do espectro transformada com curva temporal quando comparada a uma parte de frequência mais baixa da representação do espectro transformada com curva temporal também ajuda a compensar essa diminuição típica das linhas espectrais com maior frequência. Para resumir, uma consideração enfática da parte de frequência mais elevada do espectro traz consigo uma confiabilidade maior das informações de compactação de energia e, portanto, permite um fornecimento mais confiável do sinal de ativação com curva temporal.

Em outra materialização favorita, o provedor de informações de compactação de energia é configurado para fornecer uma pluralidade de medidas baseadas em banda de nivelamento espectral e para computar uma média da pluralidade das medidas baseadas em banda de nivelamento espectral para obter as informações de compactação de energia. Descobriu-se que a condição das medições de nivelamento espectral baseado em banda traz consigo informações especificamente confiáveis no sentido de a curva temporal ser eficiente para reduzir a taxa de bits de um sinal de áudio codificado. Primeiramente, a codificação da representação do espectro transformada com curva temporal em geral é desempenhada com base na banda, de tal forma que uma combinação das medidas baseadas em banda do nivelamento espectral seja bem adaptada para a codificação e, portanto, represente uma melhoria da taxa de bits que pode ser obtida com uma boa precisão. Além disso, uma COMPUTADOR baseada em banda das medidas de nivelamento

espectral elimina significativamente a dependência das informações de compactação de energia a partir de uma distribuição da harmônica. Por exemplo, mesmo se uma banda de frequência mais elevada compreender uma energia relativamente pequena (menor do
5 que as energias das bandas de frequência inferiores), a banda de frequência mais elevada ainda poderá ser visivelmente relevante. Porém, o impacto positivo de uma curva temporal (no sentido de uma redução do obscurecimento das linhas espectrais) nessa banda de frequência mais elevada seria considerado tão pequeno quanto,
10 simplesmente devido à pequena energia da banda de frequência mais elevada, se a medição do nivelamento espectral não fosse computado com base na banda. Em contrapartida, ao aplicar o cálculo baseado em banda, um impacto positivo da curva temporal poderá ser levado em consideração com um peso apropriado, porque as medições do
-15 - nivelamento espectral baseado em banda são independentes das energias absolutas nas bandas de frequência respectivas.

Em outra materialização favorita, o provedor do sinal de ativação de curva temporal compreende um calculador de valor de referência configurado para computar uma medição do
20 nivelamento espectral descrevendo uma representação de espectro sem curva temporal do sinal de áudio para obter o valor de referência. De forma correspondente, o sinal de ativação de curva temporal pode ser fornecido com base em uma comparação do nivelamento espectral de uma versão sem curva temporal do sinal de
25 áudio de entrada e um nivelamento espectral de uma versão com curva temporal do sinal de áudio de entrada.

Em outra materialização favorita, o provedor de informações de compactação de energia está configurado para

fornecer uma medida da entropia perceptual descrevendo a representação do espectro transformada com curva temporal do sinal de áudio como as informações de compactação de energia. Esse conceito se baseia na descoberta de que a entropia perceptual da
5 representação do espectro transformada com curva temporal é uma boa estimativa de um número de bits (ou uma taxa de bits) necessária para codificar o espectro transformado com curva temporal. De forma correspondente, a medição da entropia perceptual da representação de espectro transformada com curva
10 temporal é uma boa medida do fato de uma redução da taxa de bits pode ser esperada pela curva temporal, mesmo na visualização de que informações adicionais sobre a curva temporal terão de ser codificadas se a curva temporal for usada.

Em outra materialização favorita, o provedor de
15 informações de compactação de energia está configurado para fornecer uma medida de autocorrelação descrevendo uma autocorrelação de uma representação com curva temporal do sinal de áudio como as informações de compactação de energia. Esse conceito se baseia na descoberta de que a eficiência da curva temporal (em
20 termos de reduzir a taxa de bits) pode ser medida (ou pelo menos estimada) a partir de um sinal de domínio de tempo com curva temporal (ou ressampleado sem uniformidade). Descobriu-se que a curva temporal será eficiente de o sinal de domínio de tempo com curva temporal compreender um grau relativamente elevado de
25 periodicidade, que é refletido pela medida de autocorrelação. Em contrapartida, se o sinal de domínio de tempo com curva temporal não compreender uma periodicidade significativa, poderá ser concluído que a curva temporal não é eficiente.

Essa descoberta se baseia no fato de uma curva temporal eficiente transforma uma parte de um sinal sinusoidal de uma frequência variável (que não compreende uma periodicidade) em uma parte de um sinal sinusoidal de frequência aproximadamente constante (que compreende um alto grau de periodicidade). Em contrapartida, se a curva temporal não for capaz de fornecer um sinal de domínio de tempo tendo um alto grau de periodicidade, pode-se esperar que a curva temporal também não forneça uma economia significativa da taxa de bits, o que justificaria sua aplicação.

Em uma materialização preferencial, o provedor de informações de compactação de energia é configurado para determinar uma soma de valores absolutos de uma função normalizada de autocorrelação (em uma pluralidade de valores de retardo) da representação com curva temporal do sinal de áudio para obter as informações de compactação de energia. Descobriu-se que uma determinação computacionalmente complexa dos picos de autocorrelação não é exigida para estimar a eficiência da curva temporal. Em vez disso, descobriu-se que uma avaliação de somatória da autocorrelação em uma (ampla) gama de valores de retardo de autocorrelação também traz consigo resultados muito confiáveis. Isso se deve ao fato de que a curva temporal realmente transforma uma pluralidade de componentes de sinais (p.ex., uma frequência fundamental e sua harmônica) de frequência variável em componentes de sinais periódicos. De forma correspondente, a autocorrelação de um sinal com curva temporal desse tipo exhibe picos em uma pluralidade de valores de retardo de autocorrelação. Portanto, uma formação de somatória é uma forma computacionalmente

eficiente de extrair as informações de compactação de energia a partir da autocorrelação.

Em outra materialização preferencial, o provedor do sinal de ativação de curva temporal compreende um calculador de valor de referência configurado para computar o valor de referência a partir de uma representação espectral sem curva temporal do sinal de áudio ou a partir de uma representação do domínio de tempo sem curva temporal do sinal de áudio. Nesse caso, o comparador normalmente é configurado para formar um valor proporcional usando as informações de compactação de energia descrevendo uma compactação de energia em um espectro transformado de curva temporal do sinal de áudio e do valor de referência. O comparador também é configurado para comparar o valor proporcional com um ou mais valores limite para obter o sinal de ativação de curva temporal. Descobriu-se que a proporção entre as informações de compactação de energia no caso sem curva temporal e as informações de compactação de energia no caso com curva temporal permite uma geração computacionalmente eficiente, mas ainda suficientemente confiável do sinal de ativação de curva temporal.

Outra materialização preferencial da invenção cria um codificador de sinal de áudio para codificar um sinal de áudio de entrada para obter uma representação codificada do sinal de áudio de entrada. O codificador do sinal de áudio compreende um transformador de curva temporal configurado para fornecer uma representação do espectro transformado de curva temporal a partir do sinal de áudio de entrada. O codificador do sinal de áudio também compreende um provedor do sinal de ativação de curva temporal, conforme descrito acima. O provedor do sinal de ativação

de curva temporal é configurado para receber o sinal de áudio de entrada e para fornecer as informações de compactação de energia de forma que as informações de compactação de energia descrevam uma compactação de energia na representação do espectro transformado de curva temporal do sinal de áudio de entrada. O codificador do sinal de áudio compreende ainda um controlador configurado para fornecer seletivamente, dependendo do sinal de ativação de curva temporal, uma parte de contorno de curva temporal inconstante (variável) ou informações de curva temporal, ou uma parte de contorno de tempo constante padrão (não variável) ou informações de curva temporal ao transformador de curva temporal. Dessa forma, é possível aceitar ou rejeitar seletivamente uma parte de contorno de curva temporal inconstante na derivação da representação do sinal de áudio codificado a partir do sinal de áudio de entrada.

Esse conceito se baseia na descoberta de que nem sempre é eficiente introduzir informações de curva temporal em uma representação codificada do sinal de áudio de entrada, porque um número notável de bits é exigido para codificar as informações de curva temporal. Além disso, descobriu-se que as informações de compactação de energia, que são computadas pelo provedor do sinal de ativação de curva temporal, são uma medida computacionalmente eficiente para determinar se é vantajoso fornecer o transformador de curva temporal com a parte de contorno de curva temporal variável (inconstante) descoberta ou um contorno de curva temporal padrão (não variável, constante). Deve-se observar que, quando o transformador de curva temporal compreender uma transformação sobreposta, a uma parte de contorno de curva temporal descoberta

poderá ser usada na COMPUTADOR de dois ou mais blocos de transformação subsequentes. Em especial, descobriu-se que não é necessário codificar totalmente a versão da representação espectral transformada de curva temporal do sinal de áudio de entrada usando a parte de contorno de curva temporal variável descoberta recentemente e a versão da representação espectral transformada de curva temporal do sinal de áudio de entrada usando uma parte de contorno de curva temporal padrão (não variável) para que seja possível tomar uma decisão quanto à curva temporal permitir ou não uma economia na taxa de bits. Em vez disso, descobriu-se que uma avaliação da compactação de energia da representação espectral transformada da curva temporal do sinal de áudio de entrada forma uma base confiável da decisão. De forma correspondente, uma taxa de bits exigida pode ser mantida pequena.

Em outra materialização preferencial, o codificador do sinal de áudio compreende uma interface de saída configurada para incluir seletivamente, dependendo do sinal de ativação de curva temporal, informações de contorno de curva temporal representando um contorno de curva temporal variável descoberto na representação codificada do sinal de áudio. Portanto, uma eficiência elevada da codificação do sinal de áudio pode ser obtida, independentemente de o sinal de entrada ser bem adequado ou não para a curva temporal.

Outra materialização, de acordo com a invenção, cria um método para fornecer um sinal de ativação de curva temporal a partir de um sinal de áudio. O método corresponde à funcionalidade do provedor do sinal de ativação de curva temporal e pode ser complementado por qualquer um dos recursos e

funcionalidades descritas aqui com relação ao provedor do sinal de ativação de curva temporal.

Outra materialização, de acordo com a invenção, cria um método para codificar um sinal de áudio de entrada, para
5 obter uma representação codificada do sinal de áudio de entrada. Esse método pode ser complementado por qualquer um dos recursos e funcionalidades descritas aqui com relação ao codificador do sinal de áudio.

Outra materialização, de acordo com a invenção,
10 cria um programa computacional para desempenhar os métodos mencionados neste documento.

De acordo com um primeiro aspecto da presente invenção, uma análise do sinal de áudio, independentemente de um sinal de áudio ter uma característica harmônica ou uma
-15- característica de fala, é usada vantajosamente para controlar um processamento de preenchimento de ruído no codificador e/ou no decodificador. A análise do sinal de áudio é facilmente obtível em um sistema, no qual uma funcionalidade de curva temporal é usada, uma vez que essa funcionalidade de curva temporal normalmente
20 compreende um mecanismo de monitoramento de passo e/ou um classificador de sinal para diferenciar entre fala, por um lado, e música, por outro lado, e/ou para diferenciar entre fala com voz e fala sem voz. Como essas informações estão disponíveis em um contexto como esse sem nenhum custo extra, as informações
25 disponíveis são usadas vantajosamente para controlar o recurso de preenchimento de ruído de forma que, principalmente para sinais de fala, um preenchimento de ruído entre linhas harmônicas seja reduzido ou, para sinais de fala em particular, até eliminados.

Mesmo em situações em que um conteúdo harmônico consistente for obtido, mas uma fala não será detectada diretamente por um detector de fala, uma redução do preenchimento de ruído resultará, não obstante, em uma qualidade percebida superior. Embora esse

5 recurso seja especialmente útil em um sistema no qual a análise de harmônica/fala é desempenhada de qualquer forma e essas informações estão, portanto, disponíveis sem nenhum custo adicional, o controle do esquema de preenchimento de ruído baseado em uma análise de sinais, independentemente de o sinal ter ou não

10 uma característica harmônica ou de fala também será útil, mesmo quando um analisador de sinais específicos tiver de ser inserido no sistema, já que a qualidade é aprimorada sem aumento da taxa de bits ou, como alternativa, a taxa de bits será diminuída sem ter uma perda na qualidade, já que os bits exigidos para codificar o

15 nível de preenchimento de ruído seja reduzido quando o próprio nível de preenchimento de ruído, que pode ser transmitido a partir de um codificador para um decodificador, for reduzido.

Em outro aspecto da presente invenção, o resultado da análise do sinal, ou seja, se o sinal for um sinal

20 harmônico ou um sinal de fala, é usado para controlar o processamento da função de janela de um codificador de áudio. Descobriu-se que em uma situação na qual um sinal de fala ou um sinal harmônico for iniciado, a possibilidade será alta para que um codificador direto alterne entre janelas longas a janelas

25 curtas. Essas janelas curtas, porém, têm uma resolução de frequência reduzida de modo correspondente que, por outro lado, diminuiriam o ganho de codificação para sinais consistentemente harmônicos e, portanto, aumentariam o número de bits necessários

para codificar tal parte do sinal. Nessa visualização, a presente invenção definida nesse aspecto usa janelas mais longas do que uma janela curta quando um sinal inicial de fala ou harmônico for detectado. Como alternativa, as janelas são selecionadas com um comprimento praticamente semelhante às janelas longas, mas com uma sobreposição mais curta para reduzir efetivamente os pré-ecos. Em geral, a característica do sinal, independentemente do período de um sinal de áudio ter uma característica harmônica ou de fala, é usada para selecionar uma função de janela para esse período.

10 De acordo com outro aspecto da presente invenção, a ferramenta TNS (temporal noise shaping - formatação de ruído temporal) é controlada a partir do fato de o sinal subjacente se basear em uma operação de curva temporal ou estar em um domínio linear. Normalmente, um sinal que tiver sido processado por uma

15 operação de curva temporal terá um conteúdo harmônico consistente. Do contrário, um mecanismo de monitoramento de passo associado a uma etapa de curva temporal não geraria um contorno de passo válido e, na ausência de tal contorno de passo válido, uma funcionalidade de curva temporal teria sido desativada para esse

20 período do sinal de áudio. No entanto, os sinais harmônicos normalmente não serão adequados para serem submetidos ao processamento de TNS. O processamento de TNS é especialmente útil e induz um ganho significativo na taxa de bits/qualidade, quando o sinal processado pela etapa de TNS tem um espectro bastante plano.

25 Contudo, quando a aparência do sinal for tonal, ou seja, não plana, com é o caso dos espectros com conteúdo harmônico ou de voz, o ganho na qualidade/taxa de bits fornecido pela ferramenta TNS será reduzido. Portanto, sem a modificação inventiva da ferramenta TNS,

as partes com curva temporal normalmente não teriam TNS processado, mas seriam processadas sem uma filtragem de TNS. Por outro lado, o recurso de formatação de ruído do TNS fornece, não obstante, uma qualidade aprimorada especificamente nas situações em que o sinal está variando em amplitude/potência. Nos casos em que um princípio de um sinal harmônico ou de um sinal de fala estiver presente, e onde o recurso de alternância de bloco estiver implementado para que, em vez desse princípio, janelas longas ou pelo menos janelas maiores do que as janelas curtas sejam mantidas, a ativação do recurso de formatação de ruído temporal para esse período resultará em uma concentração do ruído ao redor do princípio de fala que efetivamente reduz pré-ecos, o que poderá ocorrer antes do princípio da fala devido a uma quantização do período ocorrendo em um processamento subsequente do codificador.

De acordo com outro aspecto da presente invenção, um número variável de linhas é processado por um quantizador/codificador de entropia dentro de um aparato de codificação de áudio para responder pela largura de banda variável, que é introduzida quadro a quadro devido ao desempenho de uma operação de curva temporal com uma característica variável de curva temporal/contorno de curva. Quando a operação de curva temporal resultar na situação de que o tempo do quadro (em termos lineares) incluído em um quadro com curva temporal for aumentado, a largura de banda de uma única linha de frequência será diminuída e, para uma largura de banda total constante, o número de linhas de frequência a serem processadas deverá ser maior com relação a uma situação sem curva temporal. Por outro lado, quando a operação de curva temporal resultar no fato de que o tempo real do sinal de

áudio no domínio de curva temporal for diminuído com relação ao comprimento do bloco do sinal de áudio no domínio linear, a largura de banda de frequência de uma única linha de frequência é aumentada e, portanto, o número de linhas processadas por um

5 codificador original deverá ser menor com relação a uma situação sem curva temporal para ter uma variação de largura de banda reduzida ou, de forma ideal, nenhuma variação de largura de banda.

As materializações preferenciais estão descritas subsequentemente com relação aos respectivos desenhos, nos quais:

10 Fig. 1 mostra um diagrama esquemático de bloco de um provedor do sinal de ativação de curva temporal, de acordo com uma materialização da invenção;

Fig. 2a mostra um diagrama esquemático de bloco de um codificador do sinal de áudio, de acordo com uma

15 materialização da invenção;

Fig. 2b mostra outro diagrama esquemático de bloco de um provedor do sinal de ativação de curva temporal de acordo com uma materialização da invenção;

Fig. 3a mostra uma representação gráfica de um

20 espectro de uma versão sem curva temporal de um sinal de áudio;

Fig. 3b mostra uma representação gráfica de um espectro de uma versão com curva temporal do sinal de áudio;

Fig. 3c mostra uma representação gráfica de um cálculo individual das medidas de nivelamento espectral para

25 diferentes bandas de frequência;

Fig. 3d mostra uma representação gráfica de um cálculo de uma medida de nivelamento espectral considerando somente a parte de frequência mais elevada do espectro;

Fig. 3e mostra uma representação gráfica de um cálculo de uma medida de nivelamento espectral usando uma representação de espectro na qual uma parte de frequência mais elevada é enfatizada em uma parte de frequência inferior;

5 Fig. 3f mostra um diagrama esquemático de bloco de um provedor de informações de compactação de energia, de acordo com outra materialização da invenção;

Fig. 3g mostra uma representação gráfica de um sinal de áudio tendo um passo temporariamente variável no domínio
10 de tempo;

Fig. 3h mostra uma representação gráfica de uma versão com curva temporal (ressampleada desigualmente) do sinal de áudio da Fig. 3g;

Fig. 3i mostra uma representação gráfica de uma
-15- função de autocorrelação do sinal de áudio de acordo com a Fig. 3g;

Fig. 3j mostra uma representação gráfica de uma função de autocorrelação do sinal de áudio de acordo com a Fig. 3h;

20 Fig. 3k mostra um diagrama esquemático de bloco de um provedor de informações de compactação de energia, de acordo com outra materialização da invenção;

Fig. 4a mostra um fluxograma de um método para fornecer um sinal de ativação com curva temporal na base de um
25 sinal de áudio;

Fig. 4b mostra um fluxograma de um método para codificar um sinal de áudio de entrada para obter uma representação codificada do sinal de áudio de entrada, de acordo

com uma materialização da invenção;

Fig. 5a ilustra uma materialização preferencial de um codificador de áudio tendo aspectos inventivos;

Fig. 5b ilustra uma materialização preferencial de um decodificador de áudio tendo aspectos inventivos;

Fig. 6a ilustra uma materialização preferencial do aspecto de preenchimento de ruído da presente invenção;

Fig. 6b ilustra uma tabela definindo a operação de controle desempenhada pelo manipulador do nível de preenchimento de ruído;

Fig. 7a ilustra uma materialização preferencial para desempenhar uma alternância de bloco baseada em curva temporal de acordo com a presente invenção;

Fig. 7b ilustra uma materialização alternativa para influenciar a função de janela;

Fig. 7c ilustra outra materialização alternativa para ilustrar a função de janela baseada nas informações de curva temporal;

Fig. 7d ilustra uma sequência de janelas de um comportamento AAC normal em um princípio com voz;

Fig. 7e ilustra sequências de janelas alternativas obtidas de acordo com uma materialização preferencial da presente invenção;

Fig. 8a ilustra a materialização preferencial de um controle baseado em curva temporal da ferramenta TNS (temporal noise shaping - formatação de ruído temporal);

Fig. 8b ilustra uma tabela definindo procedimentos de controle desempenhados no gerador de sinais de

controle de limite na Fig. 8a;

Fig. 9a-9e ilustra características diferentes de curva temporal e a influência correspondente na largura de banda do sinal de áudio ocorrendo subsequentemente a uma operação de
5 eliminação de curva temporal no decodificador;

Fig. 10a ilustra uma materialização preferencial de um controlador para controlar o número de linhas dentro de um processador de codificação;

Fig. 10b ilustra uma dependência entre o número
10 de linhas a serem descartadas/adicionadas para uma taxa de amostragem;

Fig. 11 ilustra uma comparação entre uma escala
de tempo linear e uma escala de tempo curva;

Fig. 12a ilustra uma implementação no contexto da
15 extensão da largura de banda; e

Fig. 12b ilustra uma tabela mostrando a dependência entre a taxa de amostragem local no domínio com curva temporal e o controle de coeficientes espectrais.

A Fig. 1 mostra um diagrama esquemático de bloco
20 do provedor do sinal de ativação de curva temporal, de acordo com uma materialização da invenção. O provedor do sinal de ativação da curva temporal 100 está configurado para receber uma representação 110 de um sinal de áudio e para fornecer, com base nele, um sinal de ativação de curva temporal 112. O provedor do sinal de ativação
25 de curva temporal 100 compreende um provedor de informações de compactação de energia 120, que é configurado para fornecer informações de compactação de energia 122, descrevendo uma compactação de energia em uma representação de espectro

transformada de curva temporal do sinal de áudio. O provedor do
sinal de ativação de curva temporal 100 compreende ainda mais um
comparador 130 configurado para comparar as informações de
compactação de energia 122 com um valor de referência 132 e para
5 fornecer o sinal de ativação de curva temporal 112 dependendo do
resultado da comparação.

Conforme discutido acima, descobriu-se que as
informações de compactação de energia são informações valiosas que
permitem uma estimativa computacionalmente eficiente,
10 independentemente de uma curva temporal trazer consigo uma
economia de bits ou não. Descobriu-se que a presença de uma
economia de bits está intimamente relacionada à questão de a curva
temporal resultar ou não em uma compactação de energia.

A Fig. 2a mostra um diagrama esquemático de bloco
- 15 - de um codificador de sinal de áudio 200, de acordo com uma
materialização da invenção. O codificador do sinal de áudio 200 é
configurado para receber um sinal de áudio de entrada 210 (também
designado como $a(t)$) e para fornecer, com base nele, uma
representação codificada 212 do sinal de áudio de entrada 210. O
20 codificador do sinal de áudio 200 compreende um transformador de
curva temporal 220, que está configurado para receber o sinal de
áudio de entrada 210. (que poderá ser representado em um domínio de
tempo) e para fornecer, com base nele, uma representação espectral
transformada de curva temporal 222 do sinal de áudio de entrada
25 210.

O codificador do sinal de áudio 200 compreende
também um analisador de curva temporal 284, que está configurado
para analisar o sinal de áudio de entrada 210 e para fornecer, com

base nele, informações de contorno de curva temporal (p.ex., informações de contorno de curva temporal absolutas ou relativas) 286.

O codificador do sinal de áudio 200 compreende
5 também um mecanismo de alternância, por exemplo, na forma de uma alternância controlada 240 para determinar se as informações de contorno de curva temporal 286 descobertas ou as informações de contorno de curva temporal padrão 288 são usadas para outros processamentos. Portanto, o mecanismo de alternância 240 é
10 configurado para fornecer seletivamente, dependendo das informações de ativação de curva temporal, as informações de contorno de curva temporal descobertas 286 ou as informações de contorno de curva temporal padrão 288 como informações de contorno de curva temporal novas 242, para outro processamento, por
15 exemplo, para o transformador de curva temporal 220. Deve-se observar que o transformador de curva temporal 220 poderá, por exemplo, usar as novas informações de contorno de curva temporal 242 (por exemplo, uma nova parte de contorno de curva temporal) e, além disso, as informações de curva temporal obtidas anteriormente
20 (por exemplo, uma ou mais partes de contorno de curva temporal obtidas anteriormente) para a curva temporal de um período de áudio. O pós-processamento de espectro opcional poderá, por exemplo, compreender uma formatação de ruído temporal e/ou uma análise de preenchimento de ruído. O codificador do sinal de áudio
25 200 também compreende um quantizador/codificador 260, que está configurado para receber a representação espectral 222 (processada opcionalmente pelo pós-processamento de espectro 250) e para quantizar e codificar a representação espectral transformada 222.

Para esse fim, o quantizador/codificador 260 poderá ser acoplado a um modelo perceptual 270 e receber informações de relevância perceptuais 272 do modelo perceptual 270 para considerar um mascaramento perceptual e ajustar precisões de quantização em
5 diferentes compartimentos de frequência de acordo com a percepção humana. O codificador do sinal de áudio 200 compreende também uma interface de saída 208 que é configurado para fornecer a representação codificada 212 do sinal de áudio com base na representação espectral quantizada e codificada 262 fornecida pelo
10 quantizador/codificador 260.

O codificador do sinal de áudio 200 compreende também um provedor do sinal de ativação de curva temporal 230, que está configurado para fornecer um sinal de ativação de curva temporal 232. O sinal de ativação de curva temporal poderá, por
15 exemplo, ser usado para controlar o mecanismo de alternância 240 para determinar se as informações de contorno de curva temporal descobertas recentemente 286 ou as informações de contorno de curva temporal padrão 288 são usadas em outras etapas de processamento (por exemplo, pelo transformador da curva temporal
20 220). Além disso, as informações de ativação da curva temporal 232 poderão ser usadas em uma alternância 280 para determinar se as novas informações de contorno de curva temporal selecionadas 242 (selecionadas a partir das informações de contorno de curva temporal recentemente descobertas 286 e das informações de
25 contorno de curva temporal padrão) são incluídas na representação codificada 212 do sinal de áudio de entrada 210. Normalmente, as informações de contorno de curva temporal são incluídas somente na representação codificada 212 do sinal de áudio se as informações

de contorno de curva temporal selecionadas descreverem um contorno de curva temporal inconstante (variável). Além disso, as próprias informações de ativação de curva temporal 232 poderão ser incluídas na representação codificada 212, por exemplo, no formato

5 de um sinalizador de bits indicando uma ativação ou uma desativação da curva temporal.

Para facilitar a compreensão, deve-se observar que o transformador de curva temporal 220 normalmente compreende um mecanismo de janela de análise 220a, um ressampleador ou um

10 "mecanismo de curva temporal" 220b e um transformador de domínio espectral (ou conversor de tempo/frequência) 220c. Dependendo da implementação, no entanto, o mecanismo de curva temporal 220b pode ser substituído - em uma direção de processamento de sinal - antes do mecanismo de janela de análise 220a. Porém, a curva temporal e

15 o domínio de tempo para a transformação de domínio espectral poderão ser combinados em uma única unidade em algumas materializações.

A seguir, detalhes sobre a operação do provedor do sinal de ativação da curva temporal 230 serão descritos. Deve-

20 se observar que o provedor do sinal de ativação da curva temporal 230 poderá ser equivalente a provedor do sinal de ativação da curva temporal 100.

O provedor do sinal de ativação da curva temporal 230 está configurado preferencialmente para receber a

25 representação do sinal de áudio de domínio de tempo 210 (também designado com $a(t)$), as informações de contorno de curva temporal descobertas recentemente 286 e as informações de contorno de curva temporal padrão 288. O provedor do sinal de ativação de curva

temporal 230 também está configurado para obter, usando o sinal de áudio do domínio de tempo 210, as informações de contorno de curva temporal descobertas recentemente 286 e as informações de contorno de curva temporal padrão 288, as informações de compactação de energia descrevendo uma compactação de energia devido às informações de contorno de curva temporal descobertas recentemente 286 e para fornecer o sinal de ativação de curva temporal 232 a partir dessas informações de compactação de energia.

A Fig. 2b mostra um diagrama esquemático de bloco de um provedor do sinal de ativação de curva temporal 234, de acordo com uma materialização da invenção. O provedor do sinal de ativação de curva temporal 234 poderá assumir o papel do provedor do sinal de ativação de curva temporal 230 em algumas materializações. O provedor do sinal de ativação de curva temporal 234 é configurado para receber um sinal de áudio de entrada 210 e duas informações de contorno de curva temporal 286 e 288, e fornecer, com base nele, um sinal de ativação de curva temporal 234p. O sinal de ativação de curva temporal 234p poderá assumir o papel do sinal de ativação de curva temporal 232. O provedor do sinal de ativação de curva temporal compreende dois provedores de representação de curva temporal idênticos 234a, 234g, que são configurados para receber um sinal de áudio de entrada 210 e as informações de contorno de curva temporal 286 e 288, respectivamente, e para fornecer, com base nele, duas representações com curva temporal 234e e 234k, respectivamente. O provedor do sinal de ativação da curva temporal 234 compreende ainda dois fornecedores de informações de compactação de energia idênticos 234f e 234l, que estão configurados para receber as

representações de curva temporal 234e e 234k, respectivamente e, com base nele, fornecer as informações de compactação de energia 234m e 234n, respectivamente. O provedor do sinal de ativação de curva temporal compreende também um comparador 234o, configurado
5 para receber as informações de compactação de energia 234m e 234n, e, a partir dele, fornecer o sinal de ativação de curva temporal 234p.

Para facilitar a compreensão, deve-se observar que os provedores de representação de curva temporal 234a e 234g
10 normalmente compreendem mecanismos de janela de análise idênticos (opcionais) 234b e 234h, ressamplers idênticos ou mecanismos de curva temporal 234c e 234i, e transformadores de domínio espectral idênticos (opcionais) 234d e 234j.

A seguir, diferentes conceitos para obter as
15 informações de compactação de energia serão abordados. De antemão, será fornecida uma introdução explicando o efeito da curva temporal em um sinal de áudio típico.

A seguir, o efeito da curva temporal em um sinal de áudio será descrito levando-se em consideração as Figs. 3a e
20 3b. A Fig. 3a mostra uma representação gráfica de um espectro de um sinal de áudio. Uma abscissa 301 descreve uma frequência e uma ordenada 302 descreve a intensidade do sinal de áudio. Uma curva 303 descreve uma intensidade do sinal de áudio sem curva temporal como uma função da frequência f .

25 A Fig. 3b mostra uma representação gráfica de um espectro de uma versão com curva temporal do sinal de áudio representado na Fig. 3a. Novamente, uma abscissa 306 descreve uma frequência e uma ordenada 307 descreve a intensidade da versão

curva do sinal de áudio sobre a frequência. Como pode ser visto a partir de uma comparação da representação gráfica das Figs. 3a e 3b, a versão sem curva temporal do sinal de áudio compreende um espectro obscuro, especialmente em uma região de frequência mais elevada. Em contrapartida, a versão com curva temporal do sinal de entrada de áudio compreende um espectro tendo picos espectrais claramente diferenciáveis, mesmo na região de frequência mais elevada. Além disso, uma formatação moderada dos picos espectrais ainda pode ser observada na região espectral inferior da versão com curva temporal do sinal de áudio de entrada.

Deve-se observar que o espectro da versão com curva temporal do sinal de áudio de entrada, que é mostrada na Fig. 3b, pode ser quantizada e codificada, por exemplo, pelo quantizador/codificador 260, com uma taxa de bits menor do que o espectro do sinal de áudio de entrada sem curva mostrado na Fig. 3a. Isso se deve ao fato de que um espectro obscurecido normalmente compreende um número grande de coeficientes espectrais perceptualmente relevantes (ou seja, um número comparativamente pequeno de coeficientes espectrais quantizados para zero ou quantizados para valores pequenos), enquanto um espectro "menos plano" conforme mostrado na Fig. 3 normalmente compreende um número maior de coeficientes espectrais quantizados para zero ou quantizados para valores pequenos. Os coeficientes espectrais quantizados para zero ou quantizados para valores pequenos podem ser codificados com menos bits do que os coeficientes espectrais quantizados para valores maiores, de forma que o espectro da Fig. 3b possa ser codificado usando menos bits do que o espectro da Fig. 3a.

Não obstante, também se deve observar que o uso de uma curva temporal nem sempre resulta em uma melhoria significativa da eficiência de codificação do sinal com curva temporal. De modo correspondente, em alguns casos, o preço, em
5 termos de taxa de bits, exigido para a codificação das informações de curva temporal (p.ex., contorno da curva temporal) poderá exceder as economias, em termos de taxa de bits, para codificar o espectro transformado de curva temporal (quando comparado à codificação do espectro transformado sem curva temporal). Nesse
10 caso, é preferível fornecer a representação codificada do sinal de áudio usando um contorno de curva temporal padrão (sem variação) para controlar a transformação da curva temporal. Consequentemente, a transmissão de quaisquer informações da curva temporal (ou seja, informações de contorno da curva temporal) pode
15 ser omitida (exceto para um sinalizador indicando a desativação da curva temporal), mantendo, portanto, a taxa de bits baixa.

A seguir, diferentes conceitos para um cálculo confiável e computacionalmente eficiente de um sinal de ativação de curva temporal 112, 232, 234p serão descritos levando-se em
20 consideração as Figs. 3c-3k. Contudo, antes disso, o plano de fundo do conceito inventivo será brevemente resumido.

A hipótese básica é que a aplicação da curva temporal em um sinal harmônico com um passo variável torna o passo constante e isso otimiza a codificação de espectros obtidos por
25 uma transformação de tempo e frequência a seguir, porque, em vez do obscurecimento das diferentes harmônicas em vários compartimentos espectrais (veja a Figs. 3a), somente um número limitado de linhas significativas irá permanecer (veja a Fig. 3b).

No entanto, mesmo quando uma variação de passo for detectada, a melhoria no ganho de codificação (ou seja, a quantia de bits salvos) poderá ser negligenciável (p.ex., se um tiver um ruído forte subjacente ao sinal harmônico ou se a variação for tão
5 pequena que o obscurecimento de harmônicas maiores não será um problema) ou poderá ser inferior à quantidade de bits necessária para transferir o contorno de curva temporal ao decodificador ou poderá simplesmente estar errado. Nesses casos, é preferível rejeitar o contorno de curva temporal variável (p.ex., 286)
10 produzido por um codificador de contorno de curva temporal e, em vez disso, usar uma sinalização eficiente de um bit, sinalizando um contorno de curva temporal padrão (sem variação).

O escopo da presente invenção compreende a criação de um método para determinar se uma parte de contorno de
15 curva temporal obtida fornece um ganho de codificação suficiente (por exemplo, um ganho de codificação suficiente para compensar a elevação exigida para a codificação do contorno da curva temporal).

Conforme informado acima, o aspecto mais
20 importante da curva temporal é a compactação da energia espectral em um número menor de linhas (veja as Figs. 3a e 3b). Um bloco desse mostra que uma compactação de energia também corresponde a um espectro mais "não plano" (veja as Figs. 3a e 3b), já que a diferença entre picos e vales do espectro é maior. A energia é
25 concentrada em menos linhas com as linhas entre aquelas que têm menos energia do que antes.

As Figs. 3a e 3b mostram um exemplo esquemático com um espectro sem curva de uma estrutura com forte variação de

harmônica e de passo (Fig. 3a) e o espectro da versão com curva temporal da mesma estrutura (Fig. 3b).

Ao visualizar essa situação, descobriu-se que é vantajoso usar a medida de nivelamento espectral como uma medida possível para a eficiência da curva temporal.

O nivelamento espectral poderá ser calculado, por exemplo, ao dividir a média geométrica do espectro de potência pela média aritmética do espectro de potência. Por exemplo, o nivelamento espectral (também designado resumidamente como "nivelamento") pode ser computado de acordo com a seguinte equação:

$$\text{Flatness} \equiv \frac{\sqrt[N]{\prod_{n=0}^{N-1} x(n)}}{\left(\frac{\sum_{n=0}^{N-1} x(n)}{N} \right)}$$

Acima, $x(n)$ representa a magnitude de um número n de compartimentos. Além disso, acima, N representa um número total de compartimentos espectrais considerados para o cálculo da medida de nivelamento espectral.

Em uma materialização da invenção, o cálculo mencionado acima do "nivelamento", que poderá atuar como informações de compactação de energia, poderá ser desempenhado usando as representações de espectro transformado de curva temporal 234e, 234k, de forma que a relação a seguir possa conter:

$$x(n) = |X|_{tw}(n).$$

Nesse caso, N poderá ser igual ao número de linhas espectrais fornecidas pelo transformador de domínio espectral 234d, 234j e $|X|_{tw}(n)$ é a representação do espectro transformada de curva temporal 234e, 234k.

Embora a medida espectral seja uma quantidade útil para o fornecimento do sinal de ativação de curva temporal, uma desvantagem da medida de nivelamento espectral, como a medida SNR (signal-to-noise ratio - proporção de sinal para ruído) é que se ela for aplicada a todo o espectro, ela enfatizará partes com mais energia. Normalmente, os espectros harmônicos têm uma determinada inclinação espectral, significando que a maior parte da energia está concentrada nos primeiros tons parciais e, em seguida, diminui com a maior frequência, acarretando uma sub-representação das parciais mais elevadas na medida. Isso não é desejado em algumas materializações, já que se deseja aprimorar a qualidade dessas parciais mais elevadas, porque elas são as que mais ficam obscurecidas (veja a Fig. 3a). A seguir, vários conceitos opcionais para a melhoria da relevância da medida de nivelamento espectral serão discutidos.

Em uma materialização de acordo com a invenção, uma abordagem semelhante à então chamada medida "SNR segmental" é escolhida, acarretando uma medida de nivelamento espectral com base na banda. Um cálculo da medida de nivelamento espectral é desempenhado (por exemplo, separadamente) em uma série de bandas e uma média é obtida. As diferentes bandas poderão ter uma largura de banda semelhante. No entanto, preferencialmente, as larguras de banda poderão seguir uma escala perceptual, como bandas críticas ou corresponder, por exemplo, às bandas do fator de escala da então chamada "codificação de áudio avançada", também conhecida como AAC.

O conceito mencionado acima será resumidamente explicado a seguir, levando-se em consideração a Fig. 3c, que

mostra uma representação gráfica de um cálculo individual de medidas de nivelamento espectral para diferentes bandas de frequência 311, 312, 313, que poderão ter uma largura de banda igual ou poderão ter larguras de banda diferentes. Por exemplo,

5 uma primeira medida de nivelamento espectral poderá ser computada para a primeira banda de frequência 311, por exemplo, usando a equação para o "nivelamento" fornecido acima. Nesse cálculo, os compartimentos de frequência da primeira banda de frequência poderão ser considerados (a variável de execução n poderá

10 considerar os índices de compartimento de frequência dos compartimentos de frequência da primeira banda de frequência) e a largura da primeira banda de frequência 311 poderá ser considerada (a variável N poderá considerar a largura em termos dos compartimentos de frequência da primeira banda de frequência). De

15 forma correspondente, a medida de nivelamento da primeira banda de frequência 311 será obtida. De modo semelhante, uma medida de nivelamento poderá ser calculada para a segunda banda de frequência 312, considerando-se os compartimentos de frequência das segundas bandas de frequência 312 e também a largura da

20 segunda banda de frequência. Além disso, as medidas de nivelamento das bandas de frequência adicionais, como a terceira banda de frequência 313, poderão ser computadas da mesma forma.

Como consequência, uma média das medidas de nivelamento para bandas de frequência diferentes 311, 312, 313

25 poderá ser computada e a média poderá atuar como as informações de compactação de energia.

Outra abordagem (para a melhoria da derivação do sinal de ativação de curva temporal) é aplicar a medida de

nivelamento espectral somente acima de uma determinada frequência. Tal abordagem está ilustrada na Fig. 3b. Como pode ser visto somente os compartimentos de frequência em uma parte superior da frequência 316 dos espectros são considerados para um cálculo da medida de nivelamento espectral. Uma parte inferior da frequência do espectro é negligenciada para o cálculo da medida de nivelamento espectral. A maior parte da frequência 316 poderá ser considerada com base na banda de frequência para o cálculo da medida de nivelamento espectral. Como alternativa toda a parte de frequência superior 316 poderá ser considerada em sua completude para o cálculo da medida de nivelamento espectral.

Para resumir as informações acima, pode-se afirmar que a diminuição do nivelamento espectral (causada pela aplicação da curva temporal) poderá ser considerada como uma primeira medida para a eficiência da curva temporal.

Por exemplo, o provedor do sinal de ativação da curva temporal 100, 230, 234 (ou o comparador 130, 234o dele) poderá comparar a medida de nivelamento espectral da representação espectral transformada da curva temporal 234e com uma medida de nivelamento espectral da representação espectral transformada da curva temporal 234k usando informações de contorno padrão da curva temporal e determinar, a partir da dita comparação, se o sinal de ativação da curva temporal deverá estar ativado ou desativado. Por exemplo, a curva temporal é ativada por meio de uma configuração apropriada do sinal de ativação da curva temporal se a curva temporal resultar em uma redução suficiente da medida de nivelamento espectral em comparação a um caso sem curva temporal.

Além das abordagens mencionadas acima, a parte

superior da frequência do espectro pode ser enfatizada (por exemplo, por um escalonamento apropriado) sobre a parte inferior da frequência para o cálculo da medida de nivelamento espectral. A Fig. 3c mostra uma representação gráfica de um espectro transformado da curva temporal na qual uma parte superior da frequência é enfatizada em uma parte inferior da frequência. De forma correspondente, uma sub-representação das parciais mais elevadas no espectro é compensada. Portanto, a medida de nivelamento pode ser computada no espectro escalonado completo no qual os compartimentos de frequência mais elevados são enfatizados nos compartimentos de frequência inferiores, conforme mostrado na Fig. 3e.

Em termos de economias de bits, uma medida típica da eficiência de codificação seria a entropia perceptual, que pode ser definida de tal forma que seja muito bem correlacionada com o número real de bits necessários para codificar um determinado espectro conforme descrito em 3GPP TS 26.403 V7.0.0: 3rd Generation Partnership Project; Technical Specification Group Services and System Aspects; General audio codec audio processing functions; Enhanced aacPlus general audio codec; Encoder specification AAC part: Section 5.6.1.1.3 Relation between bit demand and perceptual entropy. Como consequência, a redução da entropia perceptual é outra medida para a eficiência da curva temporal.

A Fig. 3f mostra um provedor de informações sobre compactação de energia 325, que pode ocupar o lugar do provedor de informações sobre compactação de energia 120, 234f, 2341 e que poderá ser usado nos provedores do sinal de ativação de curva

temporal 100, 290, 234. O provedor de informações de compactação de energia 325 é configurado para receber uma representação do sinal de áudio, por exemplo, no formato de uma representação do espectro transformado de curva temporal 234e, 234k, também
5 designado como $|X|_{tw}$. O provedor de informações de compactação de energia 325 também está configurado para fornecer informações sobre entropia perceptual 326, que poderá ocupar o lugar das informações de compactação de energia 122, 234m, 234n.

O provedor de informações sobre compactação de
10 energia 325 compreende um calculador de fator de formatação 327, que está configurado para receber a representação do espectro transformado de curva temporal 234e, 234k e para fornecer, a partir dele, informações do fator de formatação 328, que poderão estar associadas a uma banda de frequência. O provedor de
15 informações sobre compactação de energia 325 também compreende um calculador de energia de banda de frequência 329, que está configurado para calcular as informações de energia da banda de frequência $en(n)$ (330) a partir da representação do espectro com curva temporal 234e, 234k. O provedor de informações de
20 compactação de energia 325 também compreende uma série de mecanismos de estimativa de linhas 331, que estão configurados para fornecer um número estimado de informações de linhas nl (332) para uma banda de frequência tendo um índice n . Além disso, o provedor de informações sobre compactação de energia 325
25 compreende um calcular de entropia perceptual 333, que está configurado para computar as informações de entropia perceptual 326 a parti das informações sobre energia de banda de frequência 330 e do número estimado de informações de linhas 332. Por exemplo,

o calculador do fator de formatação 327 poderá ser configurado para computar o fator de formatação de acordo com

$$ffac(n) = \sum_{k=kOffset(n)}^{kOffset(n+1)-1} \sqrt{|X(k)|} \quad (1)$$

Na equação acima, $ffac(n)$ designa o fator de formatação para a banda de frequência tendo um índice n da banda de frequência. k designa uma variável em execução, que é executada sobre os índices de compartimentos espectrais da banda do fator de escala (ou banda de frequência) n . $X(k)$ designa um valor espectral (por exemplo, um valor de energia ou um valor de magnitude) do compartimento espectral (ou compartimento de frequência) tendo um índice de compartimento espectral (ou um índice de compartimento de frequência) k .

O número do estimador de linhas poderá ser configurado para estimar o número de linhas diferentes de zero, designadas com nl , de acordo com a seguinte equação:

$$nl = \frac{ffac(n)}{\left(\frac{en(n)}{kOffset(n+1)-kOffset(n)}\right)^{0.25}} \quad (2)$$

Na equação acima, $en(n)$ designa uma energia na banda de frequência ou na banda do fator de escala tendo o índice n . $kOffset(n+1)-kOffset(n)$ designa uma largura da banda de frequência ou uma banda do fator de escala do índice n em termos de compartimentos de frequência.

Além disso, o calculador de entropia perceptual 332 poderá ser configurado para computar as informações de entropia perceptual $sfbPe$ de acordo com a seguinte equação:

$$sfbPe = nl \cdot \begin{cases} \log_2\left(\frac{en}{thr}\right) & \text{for } \log_2\left(\frac{en}{thr}\right) \geq c1 \\ (c2 + c3 \cdot \log_2\left(\frac{en}{thr}\right)) & \text{for } \log_2\left(\frac{en}{thr}\right) < c1 \end{cases} \quad (3)$$

Acima, as seguintes relações poderão conter:

$$c1 = \log_2(8) \quad c2 = \log_2(2.5) \quad c3 = 1 - c2/c1,$$

(4)

Um total de entropia perceptual pe poderá ser computado como a soma das entropias perceptuais de várias bandas de frequência ou bandas de fator de escala.

Conforme mencionado acima, as informações da entropia perceptual 326 poderão ser usadas como informações sobre compactação de energia.

Para obter maiores detalhes sobre a COMPUTADOR da entropia perceptual, uma referência é feita à seção 5.6.1.1.3 do Padrão Internacional "3GPP TS 26.403 V7.0.0(2006-06)".

10 A seguir, um conceito será descrito para a COMPUTADOR das informações sobre compactação de energia no domínio de tempo.

Outro panorama no TW-MDCT (transformação do cosseno distinto modificado com curva temporal) é a ideia básica para alterar o sinal de uma forma que haja um passo constante ou praticamente constante dentro de um bloco. Se um passo constante for atingido, isso significará que os máximos da autocorrelação de um bloco de processamento aumentarão. Como não é elementar localizar os máximos correspondentes na autocorrelação para o caso com curva temporal e sem curva temporal, a soma dos valores absolutos para a autocorrelação normalizada pode ser usada como uma medida para a melhoria. Um aumento nessa soma corresponde a um aumento na compactação de energia.

25 Esse conceito será explicado com mais detalhes a seguir, considerando-se as Figs. 3g, 3h, 3i, 3j e 3k.

A Fig. 3g mostra uma representação gráfica de um

sinal com curva temporal no domínio de tempo. Uma abscissa 350
 descreve o tempo e uma ordenada 351 descreve um nível $a(t)$ do
 sinal de tempo sem curva temporal. Uma curva 352 descreve a
 evolução temporal do sinal de tempo sem curva temporal. Considera-
 5 se que a frequência do sinal de tempo sem curva temporal descrito
 pela curva 352 aumenta ao longo do tempo, como pode ser visto na
 Fig. 3g.

A Fig. 3h mostra uma representação gráfica de uma
 versão com curva temporal do sinal de tempo da Fig. 3g. Uma
 10 abscissa 355 descreve a curva temporal (por exemplo, em um formato
 normalizado) e uma ordenada 356 descreve o nível da versão com
 curva temporal $a(t_w)$ do sinal $a(t)$. Como pode ser visto na Fig.
 3h, a versão com curva temporal $a(t_w)$ do sinal de tempo sem curva
 temporal $a(t)$ compreende (pelo menos aproximadamente) uma
 15 frequência temporalmente constante no domínio com curva temporal.

Ou seja, a Fig. 3h ilustra o fato de que um sinal
 de tempo de uma frequência temporalmente variável é transformado
 em um sinal de tempo de uma frequência temporalmente constante por
 uma operação apropriada com curva temporal, que poderá compreender
 20 um ressampleamento com curva temporal.

A Fig. 3i mostra uma representação de uma função
 de autocorrelação do sinal de tempo sem curva $a(t)$. Uma abscissa
 360 descreve um retardo na autocorrelação τ e uma ordenada 361
 descreve uma magnitude da função de autocorrelação. As marcações
 25 362 descrevem uma evolução da função de autocorrelação $R_{uw}(\tau)$ como
 uma função do retardo na autocorrelação τ . Como pode ser visto na
 Fig. 3i, a função de autocorrelação R_{uw} do sinal de tempo sem curva
 $a(t)$ compreende um pico para $\tau = 0$ (refletindo a energia do sinal

$a(t)$) e assume valores pequenos para $\tau \neq 0$.

A Fig. 3j mostra uma representação gráfica da função de autocorrelação R_{tw} do sinal de tempo com curva temporal $a(t_w)$. Como pode ser visto na Fig. 3j, a função de autocorrelação R_{tw} compreende um pico para $\tau = 0$ e também compreende picos para outros valores τ_1, τ_2, τ_3 do retardo de autocorrelação τ . Esses picos adicionais para τ_1, τ_2, τ_3 são obtidos pelo efeito da curva temporal para aumentar a periodicidade do sinal de tempo com curva temporal $a(t_w)$. Essa periodicidade é refletida pelos picos adicionais da função de autocorrelação $R_{tw}(\tau)$ em comparação à função de autocorrelação $R_{uw}(\tau)$. Portanto, a presença de picos adicionais (ou de maior intensidade dos picos) da função do sinal de áudio com curva temporal, em comparação à função de autocorrelação do sinal de áudio original, pode ser usada como uma indicação da eficiência (em termos de uma redução da taxa de bits) da curva temporal.

A Fig. 3k mostra um diagrama esquemático de bloco de um provedor de informações de compactação de energia 370 configurado para receber uma representação do domínio de tempo com curva temporal do sinal de áudio, por exemplo, o sinal com curva temporal 234e, 234k (onde a transformação de domínio espectral 234d, 234j e, opcionalmente, o mecanismo de janela de análise 234b e 234h é omitido), e para fornecer, a partir dele, informações sobre compactação de energia 374, que poderão assumir o papel das informações sobre compactação de energia 372. O provedor de informações sobre compactação de energia 370 da Fig. 3k compreende um calculador de autocorrelação 371 configurado para computar a função de autocorrelação $R_{tw}(\tau)$ do sinal com curva temporal $a(t_w)$

em um intervalo predeterminado de valores distintos de τ . O provedor de informações sobre compactação de energia 370 também compreende uma somatória de autocorrelação 372 configurada para somar uma pluralidade de valores da função de autocorrelação $R_{tw}(\tau)$ (por exemplo, em um intervalo predeterminado de valores distintos de τ) e para fornecer a soma obtida como as informações de compactação de energia 122, 234m, 234n.

Portanto, o provedor de informações de compactação de energia 370 permite o fornecimento de informações confiáveis indicando a eficiência da curva temporal sem realmente desempenhar a transformação do domínio espectral da versão de domínio com curva temporal do sinal de áudio de entrada 210.

Portanto, é possível desempenhar uma transformação de domínio espectral da versão com curva temporal do sinal de áudio de entrada 310 somente se for descoberto, com base nas informações sobre compactação de energia 122, 234m, 234n fornecidas pelo provedor de informações sobre compactação de energia 370, que a curva temporal na verdade traz consigo uma maior eficiência de codificação.

Para resumir as informações acima, as materializações de acordo com a invenção criam um conceito para uma verificação final de qualidade. Um contorno de passo resultante (usado em um codificador de sinal de áudio com curva temporal) é avaliado em termos do seu ganho de codificação e se é aceito ou rejeitado. Várias medições referentes à esparsidade do espectro ou ao ganho de codificação poderão ser levadas em consideração nessa decisão, por exemplo, uma medida de nivelamento espectral, uma medida de nivelamento espectral segmental baseada

em banda e/ou uma entropia perceptual.

O uso de diferentes informações de compactação espectral foram discutidos, por exemplo, o uso de uma medida de nivelamento espectral, o uso de uma medida de entropia perceptual e o uso de uma medida de autocorrelação de domínio de tempo. Não obstante, há outras medidas que mostram uma compactação da energia em um espectro com curva temporal.

Todas essas medidas podem ser usadas. Preferencialmente, para todas essas medidas, uma relação entre a medida de um espectro sem curva temporal e com curva temporal é definida, e um limite é estabelecido para essa relação no codificador para determinar se um contorno da curva temporal obtido se beneficia ou não na codificação.

Todas essas medidas poderão ser aplicadas a um período total, onde somente a terceira parte do contorno de passo é nova (em que, por exemplo, três partes do contorno de passo estão associadas ao período total) ou preferencialmente somente para a parte do sinal, para a qual essa nova porção foi obtida, por exemplo, usando uma transformação com uma janela sobreposta baixa centralizada na parte do (respectivo) sinal.

Naturalmente, uma única medida ou uma combinação das medidas mencionadas acima poderá ser usada, conforme desejado.

A Fig. 4a mostra um fluxograma de um método para fornecer um sinal de ativação de curva temporal a partir de um sinal de áudio. O método 400 da Fig. 4a compreende uma etapa 410 de fornecer informações sobre compactação de energia descrevendo uma compactação da energia em uma representação espectral transformada da curva temporal do sinal de áudio. O método 400

compreende ainda uma etapa 420 de comparação das informações sobre compactação de energia com um valor de referência. O método 400 também compreende uma etapa 430 de fornecer o sinal de ativação de curva temporal dependendo do resultado da comparação.

5 O método 400 pode ser complementado por qualquer um dos recursos e funcionalidades descritas neste documento com relação ao fornecimento do sinal de ativação de curva temporal.

 A Fig. 4b mostra um fluxograma de um método para codificar um sinal de áudio de entrada para obter uma
10 representação codificada do sinal de áudio de entrada. O método 450 também compreende uma etapa 460 de fornecimento de uma representação espectral transformada de curva temporal a partir do sinal de áudio de entrada. O método 450 também compreende uma etapa 470 de fornecimento de um sinal de ativação de curva
15 temporal. A etapa 470 poderá, por exemplo, compreender a funcionalidade do método 400. Portanto, as informações de compactação de energia poderão ser fornecidas de forma que as informações sobre compactação de energia descrevam uma compactação de energia na representação do espectro transformado de curva
20 temporal do sinal de áudio de entrada. O método 450 também compreende uma etapa 480 de fornecer seletivamente, dependendo do sinal de ativação da curva temporal, uma descrição da representação espectral transformada da curva temporal do sinal de áudio de entrada usando informações de contorno de curva temporal
25 descobertas recentemente ou a descrição de uma representação espectral transformada sem curva temporal do sinal de áudio de entrada usando informações de contorno padrão de curva temporal (não variáveis) para inclusão na representação codificada do sinal

de áudio de entrada.

O método 450 pode ser complementado por qualquer um dos recursos e funcionalidades abordadas neste documento com relação à codificação do sinal de áudio de entrada.

5 A Fig. 5 ilustra uma materialização preferencial de um codificador de áudio de acordo com a presente invenção, na qual vários aspectos da presente invenção são implementados. Um sinal de áudio é fornecido em uma entrada do codificador 500. Esse sinal de áudio normalmente será um sinal de áudio distinto que foi
10 derivado a partir de um sinal de áudio analógico usando uma taxa de amostragem que também é chamada de taxa de amostragem normal. ~~Essa taxa de amostragem normal é diferente de uma taxa de~~
amostragem local gerada em uma operação de curva temporal e a taxa de amostragem normal do sinal de áudio na entrada 500 é uma taxa
15 de amostragem constante resultando em amostras de áudio separadas por um período constante. O sinal é inserido em um mecanismo de janela de análise 502 que está, nessa materialização, conectado a um controlador da função de janela 504. O mecanismo de janela de análise 502 está conectado a um mecanismo de curva temporal 506.
20 Dependendo da implementação, no entanto, o mecanismo de curva temporal 506 pode ser substituído - em uma direção de processamento de sinal - antes do mecanismo de janela de análise 502. Essa implementação será preferencial quando uma característica de curva temporal for exigida para a janela da
25 análise no bloco 502 e quando a operação de curva temporal tiver de ser desempenhada nas amostras com curva temporal em vez de ser realizada nas amostras sem curva temporal. De forma específica, no contexto da curva temporal baseada em MDCT conforme descrito em

Bernd Edler et al., "Time Warped MDCT", International Patent Application PCT/EP2009/002118. Para outras aplicações de curva temporal conforme descrito em L. Villemoes, "Time Warped Transform Coding of Audio Signals", PCT/EP2006/010246, Int. patent application, November 2005., o posicionamento entre o mecanismo de curva temporal 506 e o mecanismo de janela de análise 502 pode ser definido conforme exigido. Além disso, um conversor de tempo/frequência 508 é fornecido para desempenhar uma conversão de tempo/frequência de um sinal de áudio com curva temporal em uma representação espectral. A representação espectral pode ser inserida em uma etapa TNS (formatação de ruído temporal) 510, que fornece, como uma saída 510a, informações sobre TNS e, como uma saída 510b, valores residuais espectrais. A saída 510b é acoplada a um quantizador e ao bloco do codificador 512 que pode ser controlado por um modelo perceptual 514 para quantizar um sinal de forma que o ruído de quantização seja ocultado abaixo do limite de mascaramento perceptual do sinal de áudio.

Além disso, o codificador ilustrado na Fig. 5a compreende um analisador de curva temporal 516, que poderá ser implementado como um mecanismo de monitoramento de passo, que fornece informações de curva temporal na saída 518. O sinal na linha 518 poderá compreender uma característica de curva temporal, uma característica de passo, um contorno de passo ou informações, independentemente de o sinal analisado pelo mecanismo de análise da curva temporal ser um sinal harmônico ou um sinal não harmônico. O analisador de curva temporal também pode implementar a funcionalidade para diferenciar entre fala com voz e fala sem voz. Entretanto, dependendo da implementação e se um classificador do

5 sinal 520 estiver implementado, a decisão com voz/sem voz também pode ser tomada pelo classificador de sinais 520. Nesse caso, o analisador de curva temporal não tem necessariamente que desempenhar a mesma funcionalidade. A saída do analisador de curva temporal 518 está conectada a pelo menos uma e preferencialmente mais de uma funcionalidade no grupo de funcionalidades compreendendo o controlador de função de janela 504, o mecanismo de curva temporal 506, a etapa TNS 510, o quantizador e o codificador 512, e a interface de saída 522.

10 De forma análoga, uma saída 522 do classificador de sinais 520 pode ser conectada a uma ou mais das funcionalidades de um grupo de funcionalidades compreendendo o controlador da função de janela 504, a etapa TNS 510, um analisador de preenchimento de ruído 524 ou a interface de saída 522. Além disso, 15 a saída do analisador de curva temporal 518 também pode ser conectada ao analisador de preenchimento de ruído 524.

Embora a Fig. 5a ilustre uma situação na qual o sinal de áudio na entrada 500 do mecanismo de janela de análise seja inserido no analisador de curva temporal 516 e no 20 classificador de sinais 520, os sinais de entrada dessas funcionalidades também poderão ser obtidos a partir da saída do mecanismo de janela de análise 502 e, com relação ao classificador de sinais, podem até mesmo ser obtidos a partir da saída do mecanismo de curva temporal 506, a saída do conversor de 25 tempo/frequência 508 ou a saída da etapa 510 TNS.

Além de uma saída de sinal por meio do codificador do quantizador 512 indicada em 526, a interface de saída 522 recebe informações da TNS 510a, informações do modelo

perceptual 528, que poderão incluir fatores de escala no formato codificado, dados de indicação de curva temporal para informações mais avançadas da curva temporal, como o contorno de passo na linha 518 e informações de classificação de sinais na linha 522.

5 Além disso, o analisador de preenchimento de ruído 524 também pode gerar dados de preenchimento de ruído na saída 530 na interface de saída 522. A interface de saída 522 é configurada para gerar dados de saída de áudio codificados na linha 532 para transmissão para um decodificador ou para armazenar em um dispositivo de
10 armazenamento como um dispositivo de memória. Dependendo da implementação, os dados de saída 532 poderão incluir toda a entrada na interface de saída 522 ou poderão compreender menos informações, desde que as informações não sejam exigidas por um decodificador correspondente, que tem uma funcionalidade reduzida
15 ou desde que as informações já estejam disponíveis no decodificador devido a uma transmissão por meio de um canal de transmissão diferente.

O codificador ilustrado na Fig. 5a poderá ser implementado conforme definido detalhadamente no padrão MPEG-4,
20 além das funcionalidades adicionais ilustradas no codificador inventivo na Fig. 5a representado pelo controlador da função de janela 504, pelo analisador de preenchimento de ruído 524, pelo codificador do quantizador 512 e pela etapa 510 TNS, que têm, em comparação ao padrão MPEG-4, uma funcionalidade avançada. Outra
25 descrição está presente no padrão AAC (padrão internacional 13818-7) ou 3GPP TS 26.403 V7.0.0: projeto de parceria da terceira geração; serviços do grupo de especificação técnica e aspecto do sistema; funções de processamento de áudio codec do áudio geral;

AAC aprimorado mais codec de áudio geral.

A seguir, a Fig. 5b é discutida, ilustrando uma materialização preferencial de um decodificador de áudio para decodificar um sinal de áudio codificado recebido por meio da entrada 540. A interface de entrada 540 está operante para processar o sinal de áudio codificado de forma que os diferentes itens de informação das informações são extraídos do sinal na linha 540. Essas informações compreendem as informações de classificação de sinais 541, as informações de curva temporal 542, dados de preenchimento de ruído 543, fatores de escala 544, dados TNS 545 e informações espectrais codificadas 546. As informações espectrais codificadas são inseridas em um decodificador de entropia 547, que poderá compreender um decodificador Huffman ou um decodificar aritmético, desde que a funcionalidade do codificador no bloco 512 na Fig. 5a seja implementada como um codificador correspondente, como um codificador Huffman ou um codificador aritmético. As informações espectrais decodificadas são inseridas em um requantizador 550, que é conectado a um mecanismo de preenchimento de ruído 552. A saída do mecanismo de preenchimento de ruído 552 é inserida em uma etapa 554 TNS invertida, que recebe adicionalmente os dados TNS na linha 545. Dependendo da implementação, o mecanismo de preenchimento de ruído 552 e a etapa 554 TNS podem ser aplicados em uma ordem diferente de forma que o mecanismo de preenchimento de ruído 552 opere nos dados de saída da etapa 554 TNS em vez de nos dados de entrada TNS. Além disso, um conversor de frequência/tempo 556 é fornecido, o que alimenta um mecanismo de eliminação de curva temporal 558. Na saída da cadeia de processamento de sinais, um mecanismo de janela

de síntese desempenhando preferencialmente um processamento de sobreposição/adição é aplicado conforme indicado em 560. A ordem do mecanismo de eliminação de curva temporal 558 e da etapa de síntese 560 pode ser alterar mas, na materialização preferencial, é preferível desempenhar um algoritmo de codificação/decodificação baseado em MDCT conforme definido no padrão AAC (AAC = codificação de áudio avançado). Então, a operação de cross-fade inerente de um bloco para o próximo devido ao procedimento de sobreposição/adição é usada de forma vantajosa como a última operação nas cadeias de processamento de forma que todos os artefatos de bloqueio sejam evitados com eficiência.

Além disso, um analisador de preenchimento de ruído 562 é fornecido, que está configurado para controlar o mecanismo de preenchimento de ruído 552 e que recebe, como uma entrada, informações de curva temporal 542 e/ou informações de classificação de sinais 541, e as informações sobre o espectro requantizado, como poderá ser o caso.

Preferencialmente, todas as funcionalidades descritas a seguir são aplicadas juntas em um esquema de codificador/decodificador de áudio aprimorado. Não obstante, as funcionalidades descritas a seguir também podem ser aplicadas independentemente entre si, ou seja, de forma que somente uma ou um grupo, mas todas as funcionalidades, sejam implementadas em um determinado esquema de codificador/decodificador.

Subsequentemente, o aspecto de preenchimento de ruído da presente invenção está descrito com detalhes.

Em uma materialização, as informações adicionais fornecidas pela ferramenta de curva temporal/contorno de passo 516

na Fig. 5a são usadas de forma benéfica para controlar outras ferramentas codec e, de modo específico, a ferramenta de preenchimento de ruído implementada pelo analisador de preenchimento de ruído 524 no codificador e/ou implementado pelo analisador de preenchimento de ruído 562 e pelo mecanismo de preenchimento de ruído 552 no decodificador.

Várias ferramentas do codificador dentro da estrutura AAC como uma ferramenta de preenchimento de ruído são controladas pelas informações reunidas pela análise de contorno de passo e/ou por um conhecimento adicional de uma classificação de sinal fornecida pelo classificador de sinais 520.

Um contorno de passo descoberto indica segmentos de sinais com uma estrutura harmônica clara de forma que o preenchimento de ruído entre as linhas harmônicas possa diminuir a qualidade percebida, principalmente em sinais de fala, portanto o nível de ruído é reduzido quando um contorno de passo é encontrado. Do contrário, haveria ruído entre os tons parciais, que tem o mesmo efeito que o ruído de quantização aumentado para um espectro obscurecido. Além disso, a quantidade da redução do nível de ruído pode ser ainda mais aprimorada usando as informações do classificador de sinais, portanto, para os sinais de fala, não haveria nenhum preenchimento de ruído e um preenchimento de ruído moderado seria aplicado a sinais genéricos com uma estrutura harmônica sólida.

Em geral, o mecanismo de preenchimento de ruído 552 para inserir linhas espectrais em um espectro decodificado, onde zeros haviam sido transmitidos de um codificador para um decodificador, ou seja, onde o quantizador 512 na Fig. 5a tem

linhas espectrais quantizadas como zero. Naturalmente, quantizar linhas espectrais em zero reduziu em muito a taxa de bits do sinal transmitido e, na teoria, a eliminação dessas linhas espectrais (pequenas) não é audível, quando essas linhas espectrais estão

5 abaixo do limite de mascaramento perceptual conforme determinado pelo modelo perceptual. Não obstante, descobriu-se que esses "orifícios espectrais", que podem incluir muitas linhas espectrais adjacentes, resultam em um som bastante artificial. Portanto, a ferramenta de preenchimento de ruído é fornecida para inserir

10 linhas espectrais nas posições, onde as linhas haviam sido quantizadas para zero por um quantizador do codificador. Essas linhas espectrais poderão ter uma amplitude ou fase aleatória, e essas linhas espectrais sintetizadas do decodificador são escaladas usando uma medida de preenchimento de ruído determinada

15 no codificador conforme ilustrado na Fig. 5a ou dependendo de uma medida determinada no decodificador conforme ilustrado na Fig. 5b pelo bloco opcional 562. O analisador de preenchimento de ruído 524 na Fig. 5a está, portanto, configurado para estimar uma medida de preenchimento de ruído de uma energia dos valores de áudio

20 quantizados em zero para um período do sinal de áudio.

Em uma materialização da presente invenção, o codificador de áudio para codificar um sinal de áudio na linha 500 compreende o quantizador 512, que está configurado para quantizar valores de áudio, onde o quantizador 512 está configurado também

25 para quantizar valores de áudio em zero abaixo de um limite de quantização. Esse limite de quantização poderá ser a primeira etapa de um quantizador baseado em etapas, que é usado para a decisão, independentemente de um determinado valor de áudio ser

quantizado em zero, ou seja, para um índice de quantização de zero ou é quantizado e um, ou seja, um índice de quantização de um indicando que o valor de áudio está acima desse primeiro limite. Embora o quantizador na Fig. 5a esteja ilustrado conforme o

5 desempenho da quantização dos valores de domínio de frequência o quantizador também pode ser usado para quantizar valores de domínio de tempo em uma materialização alternativa, na qual o preenchimento de ruído é desempenhado no domínio de tempo, em vez do domínio de frequência.

10 O analisador de preenchimento de ruído 524 é implementado como um calculador de preenchimento de ruído para ~~estimar uma medida de preenchimento de ruído de uma energia dos~~ valores de áudio quantizados em zero para um período do sinal de áudio pelo quantizador 512. Além disso, o codificador de áudio

15 compreende um analisador de sinais de áudio 600 ilustrado na Fig. 6a, que está configurado para análise, independentemente de o período do sinal de áudio ter uma característica harmônica ou uma característica de fala. O analisador de sinais 600 pode, por exemplo, compreender o bloco 516 da Fig. 5a ou o bloco 520 da Fig.

20 5a ou pode compreender qualquer outro dispositivo para análise, independentemente de um sinal ser um sinal harmônico ou um sinal de fala. Como o analisador de curva temporal 516 está implementado para sempre procurar um contorno de passo e como a presença de um contorno de passo indica uma estrutura harmônica do sinal, o

25 analisador de sinais 600 na Fig. 6a pode ser implementado como um mecanismo de monitoramento de passo ou um calculador de contorno de curva temporal de um analisador de curva temporal.

O codificador de áudio compreende também um

manipulador do nível de preenchimento de ruído 602 ilustrado na Fig. 6a, que gera uma medida/nível de preenchimento de ruído manipulado a ser gerado para a interface de saída 522 indicada em 530 na Fig. 5a. O manipulador da medida de preenchimento de ruído 602 está configurado para manipular a medida de preenchimento de ruído dependendo da característica harmônica ou de fala do sinal de áudio. O codificador de áudio compreende também a interface de saída 522 para gerar um sinal codificado para transmissão ou armazenamento, o sinal codificado compreendendo a saída da medida de preenchimento de ruído manipulada pelo bloco 602 na linha 530. Esse valor corresponde à saída do valor pelo bloco 562 na implementação do decodificador ilustrada na Fig. 5b.

Conforme indicado na Fig. 5a e na Fig. 5b, a manipulação do nível de preenchimento de ruído pode ser implementada em um codificador ou em um decodificador, ou em ambos os dispositivos juntos. Na implementação em um decodificador, o decodificador para decodificar um sinal de áudio decodificado compreende a interface de entrada 539 para processar o sinal codificado na linha 540 para obter uma medida de preenchimento de ruído, ou seja, os dados de preenchimento de ruído na linha 543 e os dados de áudio codificados na linha 546. O decodificador também compreende um decodificador 547 e um requantizador 550 para gerar dados requantizados.

Além disso, o decodificador compreende um analisador de sinais 600 (Fig. 6a) que poderá ser implementado no analisador de preenchimento de ruído 562 na Fig. 5b para recuperar informações, independentemente de um período dos dados de áudio ter uma característica harmônica ou de fala.

Além disso, o mecanismo de preenchimento de ruído 552 para gerar dados de áudio de preenchimento de ruído, em que o mecanismo de preenchimento de ruído 552 está configurado para gerar dados de preenchimento de ruído em resposta à medida de

5 preenchimento de ruído transmitida por meio do sinal codificado e gerado pela interface de entrada na linha 543 e a característica harmônica ou de fala dos dados de áudio conforme definido pelos analisadores de sinais 516 e/ou 550 no codificador ou conforme definido pelo item 562 no decodificador por meio do processamento

10 e da interpretação das informações de curva temporal 542 indicando, independentemente de um determinado período ter sido ou não submetido a um processamento de curva temporal.

Além disso, o decodificador compreende um processador para processar os dados requantizados e os dados de

15 áudio de preenchimento de ruído para obter um sinal de áudio decodificado. O processador poderá incluir itens 554, 556, 558, 560 na Fig. 5b conforme for o caso. Além disso, dependendo da implementação específica do algoritmo do codificador/decodificador, o processador poderá incluir outros blocos de processamento que

20 são fornecidos, por exemplo, em um codificador de domínio de tempo, como o codificador AMR WB+ ou outros codificadores de fala.

A manipulação de preenchimento de ruído inventiva pode, portanto, ser implementada no codificador somente ao calcular a medida de ruído direta e ao manipular essa medida de

25 ruído com base nas informações harmônicas/de fala e ao transmitir a medida de preenchimento de ruído correta já manipulada que pode, em seguida, ser aplicada por um decodificador de uma forma direta. Como alternativa, a medida de preenchimento de ruído não

manipulada pode ser transmitida de um codificador para um decodificador e, em seguida, o decodificador analisará, independentemente do período real de um sinal de áudio ter apresentado uma curva temporal, ou seja, ter uma característica harmônica ou de fala, de forma que a manipulação real da medida de preenchimento de ruído ocorra no decodificador.

Subsequentemente, a Fig. 6b é abordada para explicar as materializações preferenciais para manipular a estimativa do nível de ruído.

Na primeira materialização, um nível de ruído normal é aplicado quando o sinal não tem uma característica harmônica ou de fala. Esse é o caso quando nenhuma curva temporal é aplicada. Quando, além disso, um classificador de sinais é fornecido, então o classificador de sinais fazendo distinção entre a fala e ausência de fala indicaria ausência de fala para a situação, na qual a curva temporal não foi ativada, ou seja, onde nenhum contorno de passo foi encontrado.

Quando, porém, a curva temporal estava ativa, ou seja, quando um contorno de passo for encontrado, o que indica um conteúdo harmônico, então o nível de preenchimento de ruído seria manipulado para um nível inferior ao do caso normal. Quando um classificador de sinais adicional for fornecido e, em seguida, esse classificador de sinais indicar fala e quando, simultaneamente, as informações de curva temporal indicarem um contorno de passo, então um nível de preenchimento de ruído inferior ou mesmo zero será sinalizado. Portanto, o manipulador do nível de preenchimento 602 da Fig. 6a reduzirá o nível de ruído manipulado para zero ou pelo menos para um valor inferior ao valor

baixo indicado na Fig. 6b. Preferencialmente, o classificador de sinais também tem um detector com voz/sem voz conforme indicado à esquerda da Fig. 6b. No caso da fala com voz, um nível de preenchimento de ruído muito baixo ou zero será
5 sinalizado/aplicado. No entanto, no caso da curva sem voz, onde a indicação de curva temporal não indica um processamento de curva temporal devido ao fato de que nenhum passo foi encontrado, mas onde o classificador de sinais sinalizar o conteúdo de fala, a medida de preenchimento de ruído não será manipulada, mas um nível
10 de preenchimento de ruído normal será aplicado.

Preferencialmente, o analisador de sinais de áudio compreende um mecanismo de monitoramento de passo para gerar uma indicação do passo, como um contorno de passo ou um passo absoluto de um período do sinal de áudio. Em seguida, o
- 15 - manipulador é configurado para reduzir a medida de preenchimento de ruído quando um passo é encontrado, e não para reduzir a medida de preenchimento de ruído quando um passo não é encontrado.

Conforme indicado na Fig. 6a, um analisador de sinais 600 não estará, quando aplicado ao decodificador, desempenhando uma análise de sinais real como um mecanismo de
20 monitoramento de passo ou um detector com voz/sem voz, mas o analisador de sinais verificará o sinal de áudio codificado para extrair informações da curva temporal ou informações de classificação de sinais. Portanto, o analisador de sinais 600
25 poderá ser implementado dentro da interface de entrada 539 no decodificador da Fig. 5b.

Uma outra materialização da presente invenção será discutida subsequentemente com relação às Figs. 7a-7e.

Para ataques de fala nos quais uma parte de fala com voz começa após uma parte de sinal silencioso relativo, o algoritmo de alternância de bloco poderá classificá-la como um ataque e poderá selecionar blocos curtos para esse determinado período, com uma perda do ganho de codificação no segmento do sinal que tenham uma estrutura harmônica clara. Portanto, a classificação com voz/sem voz do mecanismo de monitoramento de passo será usada para detectar ataques com voz e impedir que o algoritmo de alternância de blocos indique um ataque transiente ao redor do ataque localizado. Esse recurso também poderá ser acoplado com o classificador de sinais para impedir a alternância de blocos nos sinais de fala e permiti-los a todos os outros sinais. Além disso, um controle mais preciso da alternância de blocos poderá ser implementado não só ao permitir ou impedir a detecção de ataques, mas também ao usar um limite variável para a detecção de ataques com base no ataque de voz e nas informações de classificação de sinais. Além disso, as informações podem ser usadas para detectar ataques como os ataques de voz mencionados acima, mas em vez de alternar para blocos curtos, use janelas extensas com sobreposições curtas, que continuam sendo a resolução espectral preferível, mas diminuem a região do tempo onde os pré e pós-ecos poderão surgir. A Fig. 7d mostra o comportamento típico sem a adaptação, a Fig. 7e mostra duas possibilidades de adaptação diferentes (janelas de prevenção e de sobreposição baixa).

Um codificador de áudio de acordo com uma materialização da presente invenção opera para gerar um sinal de áudio como a saída do sinal por meio da interface de saída 522 da Fig. 5a. O codificador de áudio compreende um analisador do sinal

de áudio, como o analisador de curva temporal 516 ou um classificador de sinais 520 da Fig. 5a. Em geral, o analisador do sinal de áudio analisa se um período do sinal de áudio tem uma característica harmônica ou de fala. Para esse fim, o

5 classificador de sinais 520 da Fig. 5a poderá incluir um detector com voz/sem voz 520a ou um detector de fala/ausência de fala 520b. Embora não seja mostrado na Fig. 7a, um analisador de curva temporal como o analisador de curva temporal 516 da Fig. 5a, que pode incluir um mecanismo de monitoramento de passo, também pode

10 ser fornecido em vez dos itens 520a e 520b ou além dessas funcionalidades. Além disso, o codificador de áudio compreende o controlador da função de janela 504 para selecionar uma função de janela dependendo de uma característica harmônica ou de fala do sinal de áudio conforme determinado pelo analisador do sinal de

15 áudio. O mecanismo de janela 502 reflete, em seguida, o sinal de áudio ou, dependendo da determinada implementação, o sinal de áudio com curva temporal usando a função de janela selecionada para obter um período refletido. Esse período de janela é, em seguida, processado mais uma vez por um processador para obter um

20 sinal de áudio codificado. O processador pode compreender os itens 508, 510, 512 ilustrados na Fig. 5a ou mais ou menos funcionalidades dos codificadores de áudio bastante conhecidos, como codificadores de áudio baseados em transformação ou codificadores de áudio baseados em domínio de tempo que

25 compreendem um filtro LPC como codificadores de fala e, especificamente, codificadores de fala implementados de acordo com o padrão AMR-WB+.

Em uma materialização preferencial, o controlador

da função de janela 504 compreende um detector transiente 700 para detectar um transiente no sinal de áudio, em que o controlador da função de janela está configurado para alternar de uma função de janela para um bloco extenso a uma função de janela para um bloco

5 curto quando um transiente for detectado e uma característica harmônica ou de fala não for localizada pelo analisador de sinais de áudio. Quando, porém, um transiente for detectado e uma característica harmônica ou de fala for encontrada pelo analisador de sinais de áudio, então o controlador da função de janela 504

10 não corresponderá à função de janela para o bloco curto. As saídas da função de janela indicando uma janela longa quando nenhum transiente for obtido e uma janela curta quando um transiente for detectado pelo detector transiente estão ilustradas como 701 e 702 na Fig. 7a. Esse procedimento normal, conforme desempenhado pelo

15 codificador AAC bastante conhecido, está ilustrado na Fig. 7d. Na posição do ataque de voz, o detector transiente 700 detecta um aumento de energia de um período para o próximo período e, portanto, alterna de uma janela longa 710 para janelas curtas 712. Para acomodar essa alternância, uma janela de interrupção longa

20 714 é usada, que tem uma primeira parte de sobreposição 714a, uma parte sem pseudônimo 714b, uma segunda parte de sobreposição mais curta 714c e uma parte zero estendendo-se entre o ponto 716 e o ponto no eixo temporal indicado pelas 2048 amostras. Em seguida, a sequência de janelas curtas indicadas em 712 é desempenhada e,

25 então, é encerrada por uma janela de início longo 718 tendo uma extensa parte sobreposta 718a se sobrepondo à janela longa a seguir não ilustrada na Fig. 7d. Além disso, essa janela tem uma parte sem pseudônimo 718b, uma parte de sobreposição curta 718c e

uma parte zero estendendo-se entre o ponto 720 no eixo temporal até o ponto 2048. Essa parte é a parte zero.

Normalmente, a alternância nas janelas curtas é útil para evitar pré-ecos que ocorreriam em um período antes do
5 evento transiente que é a posição do ataque de voz ou, em geral, o início da fala ou de um sinal tendo um conteúdo harmônico. Geralmente, um sinal tem um conteúdo harmônico quando um mecanismo de monitoramento de passo determina que o sinal tem um passo. Como alternativa, há outras medidas de harmonicidade, como uma medida
10 de tonalidade acima de um determinado nível mínimo juntamente com uma característica de que picos proeminentes estejam em uma relação harmônica entre si. Há uma pluralidade de outras técnicas para determinar se um sinal é harmônico ou não.

Uma desvantagem das janelas curtas é que a
15 resolução de frequência diminui, já que a resolução de tempo aumenta. Visando à codificação de alta qualidade da fala e, especificamente, às partes de fala de voz ou partes que têm um conteúdo harmônico sólido, uma resolução com boa frequência é desejada. Portanto, o analisador do sinal de áudio ilustrado em
20 516, 520 ou 520a, 520b está operante para gerar um sinal desativado para o detector transiente 700 de forma que uma alternância para janelas curtas seja impedida quando um segmento de fala de voz ou um segmento de sinal tendo uma característica harmônica sólida for detectado. Isso assegura que, para codificar
25 tais partes do sinal, uma resolução de alta frequência é mantida. Trata-se de uma troca entre pré-ecos por um lado, e codificação de alta qualidade e de alta resolução do passo para o sinal de fala ou o passo para um sinal sem fala harmônico por outro lado.

Descobriu-se que é muito mais problemático quando o espectro harmônico não é codificado com precisão em comparação aos pré-ecos que ocorreriam. Para diminuir ainda mais os pré-ecos, um processamento TNS é preferível para tal situação, que será
5 abordado em relação às Figs. 8a e 8b.

Em uma materialização alternativa ilustrada na Fig. 7b, o analisador de sinais de áudio compreende um detector de voz/sem voz e/ou fala/ausência de fala 520a, 520b. Contudo, o detector transiente 700 incluído no controlador da função de
10 janela não está totalmente ativado/desativado como na Fig. 7a, mas o limite incluído no detector transiente é controlado usando um sinal de controle de limite 704. Nessa materialização, o detector transiente 700 é configurado para determinar uma característica quantitativa do sinal de áudio e para comparar a característica
15 quantitativa para o limite controlável, em que um transiente é detectado quando a característica quantitativa tem uma relação predeterminada com o limite controlável. A característica quantitativa pode ser um número indicando o aumento de energia de um bloco para o bloco seguinte, e o limite pode ser um determinado
20 aumento de energia limite. Quando o aumento de energia de um bloco para o seguinte for maior do que o aumento de energia limite, então um transiente será detectado de forma que, nesse caso, a relação predeterminada seja uma relação "maior que". Em outras materializações, a relação predeterminada também pode ser uma
25 relação "menor que", por exemplo, quando a característica quantitativa for um aumento de energia invertido. Na materialização da Fig. 7b, o limite controlável é controlado de forma que a probabilidade para uma alternância em uma função de

janela para um bloco curto seja reduzida quando o analisador de sinais de áudio tiver encontrado uma característica harmônica ou de fala. Na materialização de aumento de energia, o sinal de controle limite 704 resultará em um aumento do limite de forma que
5 as alternâncias nos blocos curtos ocorram somente quando o aumento de energia de um bloco para o seguinte for um aumento de energia particularmente elevado.

Em uma materialização alternativa, o sinal de saída do detector de voz/sem voz 520a ou o detector de
10 fala/ausência de fala 520b também pode ser usado para controlar o controlar da função de janela 504 de tal forma que, em vez de alternar em um bloco curto no ataque de fala, a alternância é feita em uma função de janela, que é mais longa do que a função de janela para o bloco curto. Essa função de janela assegura uma
15 resolução de frequência mais elevada do que uma função de janela curta, mas tem um comprimento mais curto do que a função de janela longa de forma que uma boa abrangência entre os pré-ecos por um lado e uma resolução de frequência suficiente por outro lado seja obtida. Em uma materialização alternativa, uma alternância em uma
20 função de janela tendo uma sobreposição menor poderá ser desempenhada conforme indicado pela linha hachurada na Fig. 7e em 706. A função de janela 706 tem um comprimento de 2048 amostras como o bloco longo, mas essa janela tem uma parte zero 708 e uma parte sem pseudônimo 710 de forma que um comprimento de
25 sobreposição curta 712 da janela 706 a uma janela correspondente 707 seja obtida. A função de janela 707, novamente, tem uma parte zero à esquerda da região 712 e uma parte sem pseudônimo à direita da região 712 em analogia à função de janela 710. Essa

materialização com baixa sobreposição resulta efetivamente em um menor tempo para a redução de pré-ecos devido à parte zero da janela 706 e 707, mas, por outro lado, tem um comprimento suficiente devido à parte sobreposta 714 e à parte sem pseudônimo
5 710 de forma que uma resolução com frequência suficientemente razoável seja mantida.

Em uma implementação MDCT preferencial conforme implementado pelo codificador AAC, manter uma determinada sobreposição fornece a vantagem extra de que, no decodificador, um
10 processamento de sobreposição/adição possa ser desempenhado, o que significa que um tipo de cross-fading entre os blocos é desempenhado. Isso evita efetivamente o bloqueio dos artefatos. Além disso, esse recurso de sobreposição/adição fornece a característica de cross-fading sem aumentar a taxa de bits, ou
15 seja, um cross-fade sampleado criticamente é obtido. Em janelas regulares longas ou curtas, a parte sobreposta é uma sobreposição de 50% conforme indicado pela parte de sobreposição 714. Na materialização em que a função de janela tem 2048 amostras de comprimento, a parte sobreposta é 50%, ou seja, 1024 amostras. A
20 função de janela tendo uma sobreposição mais curta que deve ser usada para refletir efetivamente um ataque de fala ou um ataque de um sinal harmônico é preferencialmente menor do que 50% e tem, na materialização da Fig. 7e, apenas 128 amostras, que é 1/16 de todo o comprimento da janela. Preferencialmente, as partes sobrepostas
25 entre 1/4 e 1/32 de todo o comprimento da função de janela são usadas.

A Fig. 7c ilustra essa materialização, na qual um detector exemplar de voz/sem voz 520a controla um seletor de

formatação de janela incluído no controlador da função de janela 504 para selecionar uma formatação de janela com uma sobreposição curta conforme indicado em 749 ou uma formatação de janela com uma sobreposição longa conforme indicado em 750. A seleção de um de
5 ambos os formatos é implementada quando o detector de voz/sem voz 500a emite um sinal detectado de voz em 751, no qual o sinal de áudio usado para análise pode ser o sinal de áudio na entrada 500 na Fig. 5a ou um sinal de áudio préprocessado, como um sinal de áudio com curva temporal ou um sinal de áudio que foi submetido a
10 qualquer outra funcionalidade de préprocessamento. Preferencialmente, o seletor de formatação de janela 504 na Fig. 7c que está incluído no controlador da função de janela 504 na Fig. 5a usa somente o sinal 751, quando um detector transiente incluído no controlador da função de janela detectaria um transiente e
15 comandaria uma alternância a partir de uma função de janela longa para uma função de janela curta conforme abordado com relação à Fig. 7a.

Preferencialmente, a materialização de alternância da função de janela é combinada com uma materialização de formatação de ruído temporal abordada em relação às Figs. 8a e
20 8b. Porém, a materialização TNS (temporal noise shaping - formatação de ruído temporal) também pode ser implementada sem a materialização de alternância de bloco.

A propriedade de compactação de energia espectral
25 do MDCT com curva temporal também influencia a ferramenta de formatação de ruído temporal (TNS), já que o ganho de TNS tende a diminuir para períodos com curva temporal, principalmente para alguns sinais de fala. Não obstante, é desejável ativar a TNS,

p.ex., para reduzir pré-ecos em ataques ou compensações de voz (cf. adaptação de alternância de bloco), onde nenhuma alternância de bloco é desejada, mas, ainda assim, o envelope temporal do sinal de fala exibe mudanças rápidas. Normalmente, um codificador usa

5 alguma medida para ver se a aplicação do TNS é produtiva para um determinado período, p.ex., o ganho de previsão do filtro de TNS quando aplicado ao espectro. Então, um limite de ganho de TNS variável é preferível, que será menor para segmentos com um contorno de passo ativo, de forma que seja assegurado que o TNS

10 esteja ativo com mais frequência para partes de sinal crítico, como ataques de voz. Assim como ocorre com outras ferramentas, isso também poderá ser complementado levando-se em consideração a classificação de sinais.

O codificador de áudio de acordo com essa

15 materialização para gerar um sinal de áudio compreende um mecanismo de curva temporal controlável como um mecanismo de curva temporal 506 para a curva temporal do sinal de áudio para obter um sinal de áudio com curva temporal. Além disso, um conversor de tempo/frequência 508 para converter pelo menos uma parte do sinal

20 de áudio com curva temporal em uma representação espectral é fornecido. O conversor de tempo/frequência 508 implementa preferencialmente uma transformação MDCT conforme informado a partir do codificador AAC, mas o conversor de tempo/frequência também pode desempenhar qualquer tipo de transformação, como uma

25 transformação DCT, DST, DFT, FFT ou MDST, ou pode abranger um banco de filtragem como um banco de filtragem QMF.

Além disso, o codificador compreende uma etapa de formatação de ruído temporal 510 para desempenhar uma filtragem

predictiva sobre a frequência da representação espectral de acordo com a instrução de controle de formatação de ruído temporal, em que a filtragem predictiva não é desempenhada, quando não há a instrução de controle de formatação de ruído temporal.

5 Além disso, o codificador compreende um controlador de formatação de ruído temporal para gerar a instrução de controle de formatação de ruído temporal com base na representação espectral.

De forma específica, o controlador de formatação
10 de ruído temporal é configurado para aumentar a probabilidade de desempenhar a filtragem predictiva sobre a frequência quando a representação espectral se baseia em um sinal de tempo com curva temporal ou para diminuir a probabilidade para desempenhar a filtragem predictiva sobre frequência, quando uma representação
15 espectral não se basear em um sinal de tempo com curva temporal. As especificações do controlador de formatação de ruído temporal são abordadas em relação à Fig. 8.

O codificador de áudio compreende também um processador para processar também um resultado da filtragem
20 predictiva sobre a frequência para obter o sinal codificado. Em uma materialização, o processador compreende a etapa 512 do codificador do quantizado ilustrada na Fig. 5a.

Uma etapa TNS 510 ilustrada na Fig. 5a está ilustrada detalhadamente na Fig. 8. Preferencialmente, o
25 controlador de formatação de ruído temporal incluído na etapa 510 compreende um calculador de ganho TNS 800, um mecanismo de decisão de TNS conectado subsequentemente 802 e um gerador de sinais de controle limite 804. Dependendo de um sinal do analisador de curva

temporal 516 ou do classificador de sinais 520 ou ambos, o gerador de sinais de controle limite 804 gera um sinal de controle limite 806 para o mecanismo de decisão de TNS. O mecanismo de decisão de TNS 802 tem um limite controlável, que é aumentado ou diminuído de acordo com o sinal de controle limite 806. O limite no mecanismo de decisão de TNS 802 é, nessa materialização, um limite de ganho de TNS. Quando o ganho de TNS realmente calculado gerado pelo bloco 800 excede o limite, então a instrução de controle de TNS requer um processamento de TNS como saída, enquanto no outro caso, quando o ganho de TNS está abaixo do limite de ganho de TNS, nenhuma instrução de TNS é gerada ou um sinal é gerado, o que indica que o processamento de TNS não é útil e não é desempenhado nesse período específico.

O calculador de ganho de TNS 800 recebe, como uma entrada, a representação espectral derivada do sinal com curva temporal. Normalmente, um sinal com curva temporal terá um ganho de TNS inferior, mas por outro lado, um processamento de TNS devido ao recurso de formatação temporal no domínio de tempo é benéfico na situação específica, se houver um sinal de voz/harmônico que foi submetido a uma operação de curva temporal. Por outro lado, o processamento de TNS não é útil em situações nas quais o ganho TNS é baixo, o que significa que o sinal residual de TNS na linha 510b tem a mesma energia ou uma energia superior que o sinal anterior à etapa TNS 510. Em uma situação na qual a energia do sinal residual de TNS na linha 510d é levemente inferior do que a energia anterior à etapa TNS 510, o processamento de TNS também poderá não ser vantajoso, já que a redução de bits devido à energia um pouco menor no sinal, que é

usado eficientemente pela etapa do quantizador/codificador de entropia 512, ser inferior ao aumento de bits introduzido pela transmissão necessária das informações da TNS indicadas em 510a na Fig. 5a. Embora uma materialização seja alternada automaticamente
5 no processamento de TNS para todos os períodos, nos quais um sinal com curva temporal é inserido e indicado pelas informações de passo do bloco 516 ou pelas informações do classificador de sinais do bloco 520, uma materialização preferencial também mantém a possibilidade de desativar o processamento de TNS, Mas somente
10 quando o ganho for realmente baixo ou pelo menos inferior ao caso normal, quando nenhum sinal harmônico/de fala for processado.

A Fig. 8b ilustra uma implementação na qual três configurações de limite diferentes são implementadas pelo gerador de sinais de controle limite 804/mecanismo de decisão de TNS 802.
-15- Quando não há um contorno de passo e quando um classificador de sinais indicar uma fala sem voz ou ausência de fala total, então o limite de decisão de TNS será definido como estando em um estado normal exigindo um ganho de TNS relativamente alto para ativar a TNS. Quando, porém, um contorno de passo for detectado, mas o
20 classificador de sinais indicar que não há fala ou o detector de voz/sem voz detectar uma fala sem voz, então o limite de decisão de TNS será definido para um nível inferior, o que significa que mesmo quando ganhos de TNS comparativamente menores forem calculados pelo bloco 800 na Fig. 8a, o processamento de TNS será
25 ativado ainda assim.

Em uma situação na qual um contorno de passo ativo é detectado e na qual a fala com voz é encontrada, então o limite de decisão de TNS é definido para o mesmo valor inferior ou

para um estado ainda mais inferior de forma que os ganhos de TNS pequenos sejam suficientes para ativar um processamento de TNS.

Em uma materialização, o controlador de ganho de TNS 800 será configurado para estimar um ganho em taxa de bits ou qualidade quando o sinal de áudio for submetido a uma filtragem preditiva sobre a frequência. Um mecanismo de decisão de TNS 802 compara o ganho estimado a um limite de decisão e as informações de controle de TNS em favor da filtragem preditiva são geradas pelo bloco 802, quando o ganho estimado estiver em uma relação predeterminada com o limite de decisão, onde essa relação predeterminada pode ser uma relação "maior que", mas também pode ser uma relação "menor que" para um ganho de TNS invertido, por exemplo. Conforme abordado, o controlador de formatação de ruído temporal também é configurado para variar o limite de decisão preferencialmente usando o sinal de controle de limite 806 de forma que, para o mesmo ganho estimado, a filtragem preditiva seja ativada quando a representação espectral estiver baseada no sinal de áudio com curva temporal e não seja ativado quando a representação espectral não estiver baseada no sinal de áudio com curva temporal.

Normalmente, a fala com voz exibirá um contorno de passo e uma fala sem voz, como fricativas ou sibilantes, não exibirá um contorno de passo. No entanto, há sinais sem fala com forte conteúdo harmônico e, portanto, têm um contorno de passo, embora o detector de fala não detecte a fala. Além disso, há determinada fala sobre música ou música em sinais de fala que são especificados pelo analisador de sinais de áudio (516 da Fig. 5a, por exemplo) como tendo um conteúdo harmônico, mas que não são

1

Imitação de largura de banda/operação de iteração com baixa

25

variação da frequência de sampleagem local, uma função do contorno de curva temporal local, acarretando artefatos audíveis. Os artefatos podem ser reduzidos ao selecionar adaptativamente o número de linhas - como uma função do contorno de curva temporal local e sua taxa de sampleagem média obtida - a serem codificadas no codificador central dependendo da frequência de sampleagem local, de forma que a largura de banda média constante seja obtida após a nova curva temporal no decodificador para todos os períodos. Outro benefício é a economia de bits no codificador.

10 O codificador de áudio, de acordo com essa materialização, compreende o mecanismo de curva temporal 506 para a curva temporal de um sinal de áudio usando uma característica de curva temporal variável. Além disso, um conversor de tempo/frequência 508 para converter um sinal de áudio com curva
15 temporal em uma representação espectral tendo um número de coeficientes espectrais fornecido. Também, um processador para processar um número variável de coeficientes espectrais para gerar o sinal de áudio codificado é usado, no qual esse processador compreendendo o quantizador/bloco do codificador 512 da Fig. 5a é
20 configurado para configurar um número de coeficientes espectrais para um período do sinal de áudio com base na característica de curva temporal para o período, de forma que uma variação da largura de banda representada pelo número processado de coeficientes de frequência entre os períodos seja reduzida ou
25 eliminada.

O processador implementado pelo bloco 512 poderá compreender um controlador 1000 para controlar o número de linhas, em que o resultado do controlador é 1000 de forma que, com relação

a um número de linhas definidas para o caso de um período sendo codificado sem nenhuma curva temporal, um determinado número variável de linhas seja adicionado ou descartado na extremidade superior do espectro. Dependendo da implementação, o controlador
5 1000 pode receber informações de contorno de passo em um determinado período 1001 e/ou uma frequência e sampleagem média local no período indicado em 1002.

Nas Figs. 9(a) a 9(e), as figuras corretas ilustram uma determinada situação de largura de banda para certos
10 contornos de passo em uma estrutura, em que os contornos de passo sobre o período são ilustrados nas respectivas figuras esquerdas para a curva temporal e são ilustradas nas figuras intermediárias após a curva temporal, em que uma característica de passo significativamente constante é obtida. Trata-se do alvo da
15 funcionalidade de curva temporal de forma que, após a curva temporal, a característica de passo seja o mais constante possível.

A largura de banda 900 ilustra a largura de banda que é obtida quando um determinado número de linhas gerado por um conversor de tempo/frequência 508 ou gerado por uma etapa TNS 510
20 da Fig. 5a é obtido, e quanto uma operação de curva temporal não é desempenhada, ou seja, quando o mecanismo de curva temporal 506 foi desativado, conforme indicado pela linha hachurada 507. Quando, porém, um contorno de curva temporal não constante é obtido e quando esse contorno de curva temporal for trazido para um passo
25 mais elevado induzindo a um aumento na taxa de sampleagem (Fig. 9(a), (c)), a largura de banda do espectro diminuirá com relação a uma situação normal sem curva temporal. Isso significa que o número de linhas a ser transmitido para esse período terá de ser

aumentado para equilibrar essa perda de largura de banda.

Como alternativa, trazer o passo para um passo constante inferior ilustrado na Fig. 9(b) ou Fig. 9(d) resulta em uma diminuição da taxa de sampleagem. A diminuição da taxa de sampleagem resulta em um aumento da largura de banda do espectro dessa estrutura com relação à escala linear e esse aumento da largura da banda deve ser balanceado usando uma exclusão ou um descarte de um determinado número de linhas com relação ao valor do número de linhas para a situação normal sem curva temporal.

A Fig. 9(e) ilustra um caso especial, no qual um contorno de passo é trazido para um nível intermediário, de forma que a frequência de sampleagem média dentro da estrutura seja, em vez de desempenhar a operação de curva temporal, a mesma que a frequência de sampleagem sem nenhuma curva temporal. Portanto, a largura de banda do sinal não será afetada e o número direto de linhas a ser usado para o caso normal sem curva temporal poderá ser processado, embora a operação de curva temporal seja desempenhada. Na Fig. 9, torna-se claro que desempenhar uma operação de curva temporal não influencia necessariamente na largura de banda, mas a influência da largura de banda depende do contorno de passo e da forma como a curva temporal é desempenhada em um período. Portanto, é preferível usar, como o valor de controle, uma taxa de sampleagem local ou média. A determinação dessa taxa de sampleagem local está ilustrada na Fig. 11. A parte superior na Fig. 11 ilustra o período com os valores de sampleagem equidistantes. Um período inclui, por exemplo, sete valores de sampleagem indicados por T_n na estrutura superior. A estrutura inferior mostra o resultado de uma operação de curva temporal na

qual ocorreu um aumento na taxa de sampleagem. Isso significa que a duração do período com curva temporal é menor do que a duração do período sem curva temporal. Porém, como a duração do período com curva temporal a ser introduzida no conversor de tempo/frequência é fixa, o caso de um aumento da taxa de sampleagem faz com que uma parte adicional do sinal de tempo que não pertence ao período indicado por T_n seja introduzida no período com curva temporal conforme indicado pelas linhas 1100. Portanto, um período de curva temporal compreende um período do sinal de áudio indicado por T_{lin} que é maior do que o período T_n . Na visualização dessa situação, a distância eficiente entre duas linhas de frequência ou a largura de banda de frequência de uma única linha no domínio linear (que é o valor inverso para a resolução) diminuiu e o número de linhas N_n definidas para um caso sem curva temporal quando multiplicado pelos resultados da distância da frequência reduzida em uma largura de banda menor, ou seja, uma diminuição da largura de banda.

O outro caso, não ilustrado na Fig. 11, em que uma diminuição da taxa de sampleagem é desempenhada pelo mecanismo de curva temporal, o período efetivo de uma estrutura no domínio com curva temporal é menor do que o período do domínio sem curva temporal, de forma que a largura de banda da frequência de uma única linha ou a distância entre duas linhas de frequência tenha aumentado. Agora, multiplicar esse Δf aumentado pelo número N_n de linhas para o caso normal resultará em uma largura de banda maior devido à resolução de frequência diminuída/distância de frequência aumentada entre dois coeficientes de frequência adjacentes.

A Fig. 11 ilustra também como uma taxa de

sampleagem média f_{SR} é calculada. Para esse fim, a distância do período entre duas amostras com curva temporal é determinada e o valor invertido é obtido, sendo definido como a taxa de sampleagem local entre duas amostras com curva temporal. Tal valor pode ser
5 calculado entre cada par de amostras adjacentes e o valor médio aritmético pode ser calculado e esse valor finalmente resulta na taxa de sampleagem local média, que é preferivelmente usado para ser inserido no controlador 1000 da Fig. 10a.

A Fig. 10b ilustra uma estrutura indicando
10 quantas linhas devem ser adicionadas ou descartadas dependendo da frequência de amostragem local, em que a frequência de sampleagem f_N para o caso sem curva, juntamente com o número de linhas N_N para o caso sem curva temporal define a largura de banda pretendida, que deve ser mantido constante o máximo possível para uma
15 sequência de estruturas com curva temporal ou para uma sequência de períodos com ou sem curva temporal.

A Fig. 12b ilustra a dependência entre os diferentes parâmetros discutidos em relação à Fig. 9, Fig. 10b e Fig. 11. Basicamente, quanto à taxa de sampleagem, ou seja, a taxa
20 média de sampleagem f_{SR} diminuir com relação ao caso sem curva temporal, as linhas terão de ser excluídas, embora as linhas terão de ser adicionadas quando a taxa de sampleagem aumentar com relação à taxa de sampleagem normal f_N para o caso sem curva temporal, de forma que as variações da largura de banda entre os
25 períodos sejam reduzidas ou preferencialmente até eliminadas o máximo possível.

A largura de banda resultante pelo número de linhas N_N e a taxa de sampleagem f_N define preferencialmente a

frequência cruzada 1200 para um codificador de áudio que, além de um codificador de áudio central original, tem um codificador de extensão de largura de banda (codificador BWE). Conforme conhecido na arte, o codificador de extensão de largura de banda codifica
5 somente um espectro com uma taxa de bits elevada até que a frequência cruzada e codifica o espectro da banda elevada, ou seja, entre a frequência cruzada 1200 e a frequência f_{MAX} com uma taxa baixa de bits, em que essa taxa baixa de bits normalmente é ainda menor do que 1/10 ou inferior à taxa de bits exigida para a banda
10 baixa entre uma frequência de 0 e a frequência cruzada 1200. Além disso, a Fig. 12a ilustra a largura de banda BW_{AAC} de um codificador de áudio AAC direto, que é muito maior do que a frequência cruzada. Assim, as linhas não podem apenas ser descartadas, mas também podem ser adicionadas. Também, a variação
15 da largura de banda para um número constante de linhas dependendo da taxa de sampleagem local f_{SR} é ilustrada. Preferencialmente, o número de linhas a serem adicionadas ou excluídas com relação ao número de linhas para o caso normal é definido de forma que cada período dos dados codificados de AAC tenham uma frequência máxima
20 o mais próximo possível da frequência cruzada 1200. Portanto, quaisquer orifícios espectrais devido a uma redução da largura de banda por um lado e uma elevação ao transmitir informações em uma frequência acima da frequência cruzada no período codificado da banda baixa são evitados. Isso, por um lado, aumenta a qualidade
25 do sinal de áudio decodificado e, por outro, diminui a taxa de bits.

A adição real de linhas com relação a um número definido de linhas ou a uma exclusão de linhas com relação ao

número definido de linhas pode ser desempenhada antes da quantização das linhas, ou seja, na entrada do bloco 512 ou pode ser desempenhada após a quantização ou, dependendo do código de entropia específico, também pode ser desempenhada após a
5 codificação da entropia.

Além disso, é preferível trazer as variações da largura de banda para um nível mínimo e até eliminar as variações da largura de banda, mas, em outras implementações, até uma redução das variações da largura de banda ao determinar o número
10 de linhas dependendo da característica de curva aumenta a qualidade do áudio e diminui a taxa de bits exigida comparada a uma situação em que um número constante de linhas é aplicado independentemente de uma determinada característica de curva temporal.

15 Embora alguns aspectos tenham sido descritos no contexto de um aparato, é óbvio que esses aspectos também representam uma descrição do método correspondente, em que um bloco ou dispositivo corresponde a uma etapa de um método ou a um recurso da etapa de um método. De forma análoga, os aspectos
20 descritos no contexto da etapa de um método também representam uma descrição de um bloco correspondente ou item ou recurso de um aparato correspondente.

Dependendo de determinados requisitos de implementação, as materializações da invenção podem ser
25 implementadas em hardwares ou em softwares. A implementação pode ser desempenhada usando uma mídia de armazenamento digital, por exemplo, um disquete, um DVD, um CD, um ROM, um PROM, um EPROM, um EEPROM ou uma memória FLASH, tendo sinais de controle

eletronicamente legíveis armazenados nela, que colaboram (ou são capazes de colaborar) com um sistema de COMPUTADOR programável, de tal forma que o respectivo método seja desempenhado. Algumas materializações de acordo com a invenção compreendem um portador de dados tendo sinais de controle eletronicamente legíveis, que são capazes de colaborar com um sistema de COMPUTADOR programável, de tal forma que um dos métodos descritos neste documento seja desempenhado. Em geral, as materializações da presente invenção podem ser implementadas como um produto do programa de COMPUTADOR com um código de programa, o código de programa estando operante para desempenhar um dos métodos quando o produto do programa de COMPUTADOR for executado em um computador. O código do programa poderá, por exemplo, ser armazenado em um portador legível da máquina. Outras materializações compreender o programa de COMPUTADOR para desempenhar um dos métodos descritos neste documento, armazenados em um portador legível da máquina. Ou seja, uma materialização do método inventivo é, portanto, um programa de COMPUTADOR tendo um código de programa para desempenhar um dos métodos descritos neste documento quando o programa de COMPUTADOR for executado em um computador. Outra materialização dos métodos inventivos é, portanto, um portador de dados (ou uma mídia de armazenamento digital ou uma mídia legível pelo computador) compreendendo o programa de COMPUTADOR nele gravado para desempenhar um dos métodos descritos neste documento. Outra materialização do método inventivo é, portanto, um fluxo de dados ou uma sequência de sinais representando o programa de COMPUTADOR para desempenhar um dos métodos descritos neste documento. O fluxo de dados ou a sequência de sinais poderá, por exemplo, ser

configurada para ser transferida por meio de uma conexão de comunicação de dados, por exemplo, via Internet. Outra materialização compreende um processamento, o que significa, por exemplo, que um computador ou um dispositivo lógico programável, seja configurado ou adaptado para desempenhar um dos métodos descritos neste documento. Outra materialização compreende um computador tendo instalado em si o programa de COMPUTADOR para desempenhar um dos métodos descritos neste documento. Em algumas materializações, um dispositivo de lógica programável (por exemplo, um conjunto de portas programáveis em campo) poderá ser usado para desempenhar algumas ou todas as funcionalidades dos métodos descritos neste documento. Em algumas materializações, um conjunto de portas programáveis em campo poderá colaborar com um microprocessador para desempenhar um dos métodos descritos neste documento.

REIVINDICAÇÕES

1. **"PROVEDOR DO SINAL DE ATIVAÇÃO DE CURVA TEMPORAL, CODIFICADOR DO SINAL DE ÁUDIO, MÉTODO PARA FORNECER UM SINAL DE ATIVAÇÃO DE CURVA TEMPORAL E MÉTODO PARA CODIFICAR UM SINAL DE ÁUDIO"**, caracterizado pelo codificador de áudio para codificar um sinal de áudio, compreender:

um mecanismo de curva temporal (506);

um conversor de frequência de tempo (508) para desempenhar uma conversão de tempo/frequência de um sinal de áudio ou sem um sinal de áudio de tempo/frequência com curva temporal em uma representação espectral;

um quantizador (512) para quantizar valores de áudio, onde o quantizador é configurado para quantizar para valores de áudio zero abaixo de um limite de quantização;

um calculador de preenchimento de ruído (524) para estimar uma medição de energia de valores de áudio quantizados para zero em uma estrutura do sinal de áudio para obter uma medição de preenchimento de áudio;

um analisador de sinais de áudio (516,520) para analisar se a estrutura do sinal de áudio tem uma característica harmônica ou de fala;

um manipulador (602) para manipular a medição de preenchimento de ruído dependendo de uma característica harmônica ou de fala do sinal de áudio para obter uma medição manipulada do preenchimento de ruído; e

uma interface de saída (522) para gerar um sinal codificado para transmissão ou armazenamento, o sinal

codificado compreendendo a medição manipulada do preenchimento de ruído (530);

onde o manipulador (602) é configurado para aplicar um nível de ruído normal quando nenhuma curva temporal for aplicada e para manipular o nível de preenchimento de ruído a ser inferior ao normal quando a curva temporal estiver ativa.

2. Codificador de áudio de acordo com a reivindicação 1, **caracterizado** por compreender o analisador do sinal de áudio (516, 520) um acionador de tom para gerar uma indicação de um tom, quando um tom for localizado na estrutura do sinal de áudio; e

um o manipulador (602) configurado para reduzir a medição do preenchimento de ruído quando um tom for encontrado.

3. Codificador de áudio de acordo com qualquer uma das reivindicações 1 ou 2, **caracterizado** pelo analisador do sinal de áudio compreender um detector sonoro/não sonoro (520) para detectar se pelo menos uma parte da estrutura tem voz, e

pelo manipulador (602) estar configurado para reduzir a medição do preenchimento de ruído ou para zerar a medição do preenchimento de ruído quando a parte for detectada como sendo sonora; e

na qual o manipulador (602) está configurado para não manipular ou para manipular a medição de preenchimento de ruído para um grau menor quando a parte for detectada como sendo não sonoro.

4. Decodificador para decodificar um sinal de áudio codificado, **caracterizado** por compreender:

uma interface de entrada (539) para processar o sinal de áudio codificado para obter uma medição de preenchimento (543) e dados de áudio codificados (546);

um decodificador/requantizador (547, 550) para gerar dados requantizados;

um analisador de sinais (600) para recuperar informações, se uma estrutura dos dados de áudio tiver característica harmônica ou de fala; e

um mecanismo de preenchimento de ruído (552) para gerar dados de áudio de preenchimento de ruído,

onde o mecanismo de preenchimento de ruído (552) está configurado para gerar dados de preenchimento de ruído em resposta à medição de preenchimento de ruído e à característica harmônica ou de fala dos dados de áudio; e

um processador (556, 558, 560) para processar os dados requantizados e os dados de áudio de preenchimento de ruído para obter um sinal de áudio decodificado (564);

onde o sinal de áudio codificado compreende dados (542, 541) indicando se a estrutura dos dados de áudio tem uma característica harmônica ou de fala, e

onde o analisador de sinais (600) está configurado para analisar o sinal de áudio codificado para recuperar uma indicação de dados, se a estrutura dos dados de áudio tiver uma característica harmônica ou de fala;

onde os dados são uma indicação de que a estrutura foi submetido a um processamento de curva temporal

e

onde o processador compreende um mecanismo de eliminação de curva temporal (558) para eliminar a curva temporal de um sinal de áudio derivado de dados de preenchimento de ruído e dados requantizados.

5. Método para codificar um sinal de áudio, **caracterizado** por compreender:

curva temporal (506) de um sinal de áudio;

desempenho de (508) uma conversão de tempo/frequência de um sinal de áudio ou sem um sinal de áudio de tempo/frequência com curva temporal em uma representação espectral;

quantização (512) de valores de áudio, onde o quantizador está configurado para quantizar para valores de áudio zero abaixo de um limite de quantização;

estimativa (524) de uma medição de uma energia dos valores de áudio quantizados para zero para uma estrutura do sinal de áudio;

análise (516,520) se a estrutura do sinal de áudio ter uma característica harmônica ou de fala;

manipulação (602) da medição de preenchimento de ruído dependendo de uma característica harmônica ou de fala do sinal de áudio para obter uma medição de preenchimento de ruído manipulado de forma que um nível de ruído normal seja aplicado quando nenhuma curva temporal for aplicada e de forma que o nível de preenchimento de ruído seja manipulado para um nível menor do que o normal quando a curva temporal estiver ativada; e

geração (522) de um sinal codificado para transmissão ou armazenamento, o sinal codificado compreendendo a medição de preenchimento de áudio manipulada (530).

6. Método para decodificar um sinal de áudio codificado, **caracterizado** pelo sinal de áudio codificado compreender dados (542, 541) indicando se a estrutura dos dados de áudio tem uma característica harmônica ou de fala, compreendendo:

processamento (539) do sinal de áudio codificado para obter uma medição de preenchimento de ruído (543) e de dados de áudio codificados (546);

análise do sinal de áudio codificado para recuperar uma indicação de dados, se a estrutura dos dados de áudio ter uma característica harmônica ou de fala, onde os dados são uma indicação de que a estrutura foi submetido a um processamento de curva temporal;

geração (547, 550) de dados requantizados;

recuperação (600) de informações, se uma estrutura dos dados de áudio ter característica harmônica ou de fala; e

geração (552) de dados de áudio de preenchimento de ruído em resposta à medição de preenchimento de ruído e à característica harmônica ou de curva dos dados de áudio; e

processamento (556, 558, 560) dos dados requantizados e dos dados de áudio de preenchimento de ruído para obter um sinal de áudio decodificado (564) onde o

processamento compreende uma eliminação de curva temporal de um sinal de áudio derivado de dados de preenchimento de ruído e de dados requantizados.

7. Meio de armazenamento não transitório legível por máquina, caracterizado por conter conjunto de instruções que, quando executadas realizam as etapas do método de acordo com a reivindicação 5 ou as etapas do método da reivindicação 6.

8. Codificador de áudio para gerar um sinal de áudio codificado, **caracterizado** por compreender:

um analisador de sinais de áudio (516, 520) para analisar, se uma estrutura do sinal de áudio ter uma característica harmônica ou de fala;

um controlador de função de janela (504) para selecionar uma função de janela dependendo de uma característica harmônica ou de fala do sinal de áudio;

um janelador (502) para refletor o sinal de áudio usando a função de janela selecionada para obter uma estrutura janelada; e

um processador (508, 512) para processar ainda mais a estrutura janelada para obter o sinal de áudio codificado;

onde o controlador da função de janela (504) compreende um detector transiente (700) para detectar um transiente, onde o controlador da função de janela é configurado para alternar uma função de janela para um bloco extenso para uma função de janela para um bloco curto, quando um transiente for detectado e uma característica harmônica ou

de fala não for encontrada pelo analisador do sinal de áudio (516, 520) e para não alternar a função de janela para o bloco curto quando um transiente for detectado e uma característica harmônica ou de fala for localizada pelo analisador do sinal de áudio (516, 520); e

onde o controlador da função de janela (504) é configurado para alternar para uma função de janela (707) sendo mais extensa do que a função de janela para um bloco curto e tendo uma sobreposição mais curta (712) do que a função de janela (714) para um bloco extenso quando um transiente for detectado e o sinal tiver uma característica harmônica ou de fala, de forma que a função de janela (707) tendo a sobreposição mais curta seja usada para refletir uma fala inicial ou um início de um sinal harmônico.

9. Codificador de áudio para gerar um sinal de áudio codificado, **caracterizado** por compreender:

um analisador de sinais de áudio (516, 520) para analisar, se uma estrutura do sinal de áudio ter uma característica harmônica ou de fala;

um controlador da função de janela (504) para selecionar uma função de janela dependendo de uma característica harmônica ou de fala do sinal de áudio;

um janelador (502) para refletir o sinal de áudio usando a função de janela selecionada para obter uma estrutura refletida; e

um processador (508, 512) para processar ainda mais a estrutura refletida para obter o sinal de áudio codificado;

onde o detector transiente (700) está configurado para detectar uma característica quantitativa do sinal de áudio e para comparar a característica quantitativa com um limite controlável, onde um transiente é detectado quando a característica quantitativa tiver uma relação predeterminada quanto ao limite controlável e

onde o analisador do sinal de áudio está configurado para controlar o limite variável de forma que uma probabilidade para uma comutação para uma função de janela para um bloco curto seja reduzida quando o analisador de sinais de áudio (516, 520) tiver encontrado uma característica harmônica ou de fala.

10. Método para gerar um sinal de áudio codificado, **caracterizado** por compreender:

análise (516, 520), se uma estrutura do sinal de áudio ter uma característica harmônica ou de fala;

seleção (504) de uma função de janela dependendo de uma característica harmônica ou de fala do sinal de áudio;

o janelamento (502) do sinal de áudio usando a função de janela selecionada para obter uma estrutura refletida; e

processamento (508, 512) da estrutura janelada para obter o sinal de áudio codificado;

onde uma comutação é desempenhada a partir de uma função de janela para um bloco extenso a uma função de janela para um bloco curto, quando um transiente for detectado e uma característica harmônica ou de fala não for

encontrada pela análise; e

onde uma comutação é desempenhada para uma função de janela (707) sendo maior do que a função de janela para um bloco curto e tendo uma porção de sobreposição mais curta do que a função de janela (714) para um bloco extenso, quando um transiente for detectado e o sinal tiver uma característica harmônica ou de fala, de forma que a função de janela (707) tendo uma sobreposição mais curta seja usada para refletir uma fala inicial ou um início de um sinal harmônico.

11. Método para gerar um sinal de áudio codificado, **caracterizado** por compreender:

análise (516, 520), se uma estrutura do sinal de áudio ter uma característica harmônica ou de fala;

seleção (504) de uma função de janela dependendo de uma característica harmônica ou de fala do sinal de áudio;

o janelamento (502) do sinal de áudio usando a função de janela selecionada para obter uma estrutura refletida; e

processamento (508, 512) da estrutura janelada para obter o sinal de áudio codificado;

onde uma característica quantitativa do sinal de áudio é detectada e comparada a um limite controlável, onde um transiente é detectado, quando a característica quantitativa tiver uma relação predeterminada com o limite controlável; e

onde o limite variável é controlado de forma

que uma probabilidade para uma comutação para uma função de janela para um bloco curto seja reduzida quando uma característica harmônica ou de fala tiver sido encontrada.

12. Meio de armazenamento não transitório legível por máquina, caracterizado por conter conjunto de instruções que, quando executadas realizam as etapas do método de acordo com a reivindicação 10 ou as etapas do método da reivindicação 11.

13. Codificador de áudio para gerar um sinal de áudio, **caracterizado** por compreender:

um mecanismo de curva temporal controlável (506) para introduzir uma curva temporal no sinal de áudio para obter um sinal de áudio com curva temporal;

um conversor de tempo/frequência (508) para converter pelo menos uma parte do sinal de áudio ou sem um sinal de áudio de tempo/frequência com curva temporal em uma representação espectral;

um estágio de formação de ruído temporal para desempenhar uma filtragem preditiva sobre a frequência da representação espectral de acordo com uma instrução de controle de formação de ruído temporal (803), onde a filtragem preditiva não é desempenhada quando não há uma instrução de controle de formação de ruído temporal;

um controlador de formação de ruído temporal (800, 802, 804) para gerar a instrução de controle de formação de ruído temporal com base na representação espectral,

onde o controlador de formação de ruído

temporal está configurado para aumentar uma probabilidade de desempenhar a filtragem preditiva sobre a frequência, quando a representação espectral estiver baseada em um sinal de áudio com curva temporal ou para diminuir a probabilidade de desempenhar a filtragem preditiva sobre a frequência, quando a representação espectral não estiver baseada em um sinal de áudio com curva temporal; e

um processador (512) para processar ainda mais uma saída do estágio de formação de ruído temporal para obter o sinal de áudio codificado (532);

onde o controlador de formação de ruído temporal (800, 802, 804) é configurado para estimar um ganho em uma taxa de bits ou uma qualidade, quando o sinal de áudio estiver sujeito à filtragem preditiva pelo estágio de formação de ruído temporal (510) para comparar (802) o ganho estimado com um limite de decisão, e

para decidir (802), em favor da filtragem preditiva, quando o ganho estimado está em uma relação predeterminada com o limite de decisão,

onde o controlador de formação de ruído temporal está ainda mais configurado para variar (804) o limite de decisão de forma que, para o mesmo ganho estimado, a filtragem preditiva seja ativada quando a representação espectral estiver baseada em um sinal com curva temporal e não estiver ativada quando a representação espectral não estiver baseada em um sinal de áudio com curva temporal.

14. Codificador de áudio de acordo com a reivindicação 13, **caracterizado** pelo mecanismo de curva

temporal compreender um classificador de sinais (520) para detectar vozeamento da fala e não vozeamento da fala, e

na qual o controlador de formação de ruído temporal (800, 802, 804) está configurado para aumentar a probabilidade quando um vozeamento da fala for detectado ou quando um não vozeamento da fala for detectada e a representação espectral está baseada no sinal de áudio com curva temporal.

15. Método para gerar um sinal de áudio, **caracterizado** por compreender:

curva temporal, (506) o sinal de áudio para obter um sinal de áudio com curva temporal;

conversão (508) de pelo menos uma parte do sinal de áudio ou sem um sinal de áudio de tempo/frequência com curva temporal em uma representação espectral;

desempenho de uma filtragem preditiva sobre a frequência da representação espectral de acordo com uma instrução de controle de formação de ruído temporal (803), onde a filtragem preditiva não é desempenhada quando não há uma instrução de controle de formação de ruído temporal;

geração (800, 802, 804) da instrução de controle de formação de ruído temporal baseada na representação espectral,

onde uma probabilidade para desempenhar a filtragem preditiva sobre a frequência é maior quando a representação espectral se baseia em um sinal de áudio com curva temporal ou onde a probabilidade para desempenhar a filtragem preditiva sobre a frequência é menor quando a

representação espectral não estiver baseada em um sinal de áudio sem curva temporal; e

processamento (512) de uma saída do estágio de formação de ruído temporal para obter o sinal de áudio codificado (532);

onde um ganho em uma taxa de bits ou uma qualidade, quando o sinal de áudio é submetido à filtragem preditiva pelo estágio de formação de ruído temporal (510), é estimado e

onde o ganho estimado é comparado a um limite de decisão, para decidir (802), em favor da filtragem preditiva, quando o ganho estimado está em uma relação predeterminada com o limite de decisão,

onde o limite de decisão é variado de forma que, para o mesmo ganho estimado, a filtragem preditiva é ativada quando a representação espectral se baseia em um sinal com curva temporal e não é ativada quando a representação espectral não se baseia em um sinal de áudio com curva temporal.

16. Meio de armazenamento não transitório legível por máquina, caracterizado por conter conjunto de instruções que, quando executadas realizam as etapas do método de acordo com a reivindicação 15.

17. Codificador de áudio para codificar um sinal de áudio, **caracterizado** por compreender:

um mecanismo de curva temporal (506) para introduzir a curva em um sinal de áudio usando uma característica variável de curva temporal;

um conversor de tempo/frequência (508) para converter um sinal de áudio com curva temporal em uma representação espectral tendo uma série de coeficientes espectrais; e

um processador (512) para processar um número variável de coeficientes espectrais para gerar um sinal de áudio codificado,

onde o processador (512, 1000) está configurado para definir de forma variável o número de coeficientes espectrais para uma estrutura do sinal de áudio com base na característica de curva temporal para a estrutura de forma que uma variação da largura de banda representada pelo número de coeficientes de frequência processados de acordo com cada estrutura seja reduzida ou eliminada.

18. Codificador de áudio de acordo com a reivindicação 17, **caracterizado** por compreender:

uma característica de curva temporal variável que compreende uma frequência de amostragem local (f_{SR}) para uma estrutura e na qual o processador (512, 1000) está configurado para aumentar um número de coeficientes espectrais, quando a frequência de amostragem local for aumentada ou na qual o processador (512, 1000) está configurado para diminuir o número de coeficientes espectrais, quando a frequência de amostragem for diminuída.

19. Codificador de áudio de acordo com qualquer uma das reivindicações 17 ou 18, **caracterizado** por compreender ainda mais um codificador de extensão de largura de banda para codificar uma banda espectral acima de uma

frequência cruzada (1200) usando parâmetros derivados de uma banda do sinal de áudio acima da frequência cruzada (1200), onde a frequência cruzada é uma frequência máxima de uma largura de banda alvo para cada estrutura.

20. Codificador de áudio de acordo com a reivindicação 19, **caracterizado** pelo sinal de áudio, antes de a curva temporal ser introduzida, ser retirado uma amostra usando uma frequência de amostragem normal (f_N) e na qual o processador (512, 1000) é configurado para usar um número predeterminado de coeficientes espectrais (N_N) derivados da frequência cruzada e da frequência de amostragem normal, quando a frequência de amostragem local for igual à frequência de amostragem normal ou para usar um número mais elevado de coeficientes espectrais comparado ao número predeterminado de coeficientes espectrais (N_N), quando a frequência de amostragem local for maior do que a frequência de amostragem normal (f_N) ou para usar um número menor comparado ao número predeterminado de coeficientes espectrais, quando a frequência de amostragem local for menor do que a frequência de amostragem normal (f_N).

21. Codificador de áudio de acordo com qualquer uma das reivindicações 17 a 20, **caracterizado** pelo processador compreender um quantizador para quantizar os coeficientes espectrais para obter coeficientes espectrais quantizados e um codificador de entropia para codificar a entropia dos coeficientes espectrais quantizados,

onde o processador (512, 1000) inclui um seletor para descartar coeficientes espectrais não incluídos

no número definido de coeficientes espectrais antes ou depois da quantização, de forma que o sinal de áudio codificado compreenda somente os coeficientes espectrais que não foram descartados, ou

onde o processador inclui um seletor para adicionar coeficientes espectrais exigidos pelo número definido de coeficientes espectrais antes ou depois da quantização, de forma que o sinal de áudio codificado compreenda também os coeficientes espectrais adicionados.

22. Método para codificar um sinal de áudio, **caracterizado** por compreender:

curva temporal (506) de um sinal de áudio usando uma característica variável de curva temporal;

conversão (508) de um sinal de áudio com curva temporal em uma representação espectral tendo uma série de coeficientes espectrais; e

processamento (512) de um número variável de coeficientes espectrais para gerar um sinal de áudio codificado,

onde um número variável de coeficientes espectrais para uma estrutura do sinal de áudio se baseia na característica de curva temporal para a estrutura, de forma que uma variação da largura de banda representada pelo número processado de coeficientes de frequência de acordo com cada estrutura seja reduzida ou eliminada.

23. Meio de armazenamento não transitório legível por máquina, caracterizado por conter conjunto de instruções que, quando executadas realizam as etapas do

método de acordo com a reivindicação 22.

24. Provedor do sinal de ativação de curva temporal (100; 230; 234), **caracterizado** por ser para fornecer um sinal de ativação de curva temporal (112; 232; 234p) de acordo com uma representação (110; 234e; 234k) de um sinal de áudio, que compreende:

um provedor de informações de compactação de energia (120; 234f; 234l; 325; 370) configurado para fornecer informações de compactação de energia (122; 234m; 234n; 326; 374) descrevendo uma compactação de energia em uma representação do espectro transformada com curva temporal (222) do sinal de áudio; e

um comparador (130; 234o) configurado para comparar as informações de compactação de energia (122; 234m; 234n; 326; 374) com um valor de referência e para fornecer o sinal de ativação de curva temporal (112; 232; 234p) dependendo de um resultado da comparação.

25. Provedor do sinal de ativação de curva temporal (100; 230; 234) de acordo com a reivindicação 24, **caracterizado** pelo provedor de informações de compactação de energia (120; 234f; 234l) estar configurado para fornecer uma medição do nivelamento espectral descrevendo a representação do espectro transformada com curva temporal (234e; 234k) do sinal de áudio como as informações de compactação de energia (122; 234m; 234n).

26. Provedor do sinal de ativação com curva temporal (100; 230; 234) de acordo com a reivindicação 25, **caracterizado** pelo provedor de informações de compactação de

energia (120; 234f; 234l) estar configurado para computador um quociente de um significado geométrico do espectro de potência transformada com curva temporal (234e; 234k) do sinal de áudio e um significado aritmético do espectro de potência transformada com curva temporal (234e; 234k) do sinal de áudio para obter a medição do nivelamento espectral.

27. Provedor do sinal de ativação de curva temporal (100; 230; 234) de acordo com uma das afirmações 24 a 26, **caracterizado** pelo provedor de informações de compactação de energia (120; 234f; 234l) está configurado para enfatizar uma parte de frequência elevada da representação do espectro transformada com curva temporal (234e; 234k) em comparação a uma parte de frequência mais baixa da representação do espectro transformada com curva temporal (234e; 234k) para obter as informações de compactação de energia (122; 234m; 234n).

28. Provedor do sinal de ativação com curva temporal (100;230; 234) de acordo com qualquer uma das reivindicações 24 a 27, **caracterizado** pelo provedor de informações de compactação de energia (120; 234m; 234n) estar configurado para obter uma pluralidade das medições baseadas em banda do nivelamento espectral e para computar uma média da pluralidade das medidas baseadas em banda do nivelamento espectral para obter as informações de compactação de energia (122,234m;234n).

29. Provedor do sinal de ativação com curva temporal (100;230;234) de acordo com a reivindicação 24, **caracterizado** pelo provedor de informações de compactação de

energia (120;234f;234l;325) está configurado para fornecer uma medida da entropia perceptual descrevendo a representação do espectro transformada com curva temporal (234e;234k) do sinal de áudio como as informações de compactação de energia (122;234m;234n).

30. Provedor do sinal de ativação com curva temporal (100; 230; 234; 325) de acordo com a reivindicação 29, **caracterizado** pelo provedor de informações de compactação de energia (120;234f;234l;325) estar configurado para computar um número estimado (nl) de linhas diferentes de zero para um ou mais bandas com fator de escala da representação espectral transformada com curva temporal (234e; 234k) do sinal de áudio a partir de informações sobre um fator de formação (ffac(n)) da banda do fator de escala e para computar a medida de entropia perceptual (326) para uma banda com fator de escala em consideração, usando uma multiplicação do número estimado (nl) de linhas diferentes de zero e uma medição de energia da banda com fator de escala em consideração.

31. Provedor do sinal de ativação com curva temporal (100;230;234) de acordo com a reivindicação 24, **caracterizado** pelo provedor de informações de compactação de energia (120;234f;234l;370) estar configurado para fornecer uma medida de autocorrelação (374) descrevendo uma autocorrelação de uma representação de domínio de tempo com curva temporal do sinal de áudio (234e; 234k) como as informações sobre compactação de energia.

32. Provedor do sinal de ativação com curva

temporal (100;230;234) de acordo com a reivindicação 31, **caracterizado** pelo provedor de informações de compactação de energia (120;234f;234l;370) está configurado para determinar uma soma dos valores absolutos de uma função de autocorrelação normalizada da representação com curva temporal (234e;234k) do sinal de áudio para obter as informações sobre compactação de energia.

33. Provedor do sinal de ativação com curva temporal (100;230) de acordo com qualquer uma das reivindicações de 24 a 32, **caracterizado** pelo provedor do sinal de ativação com curva temporal compreender um calculador do valor de referência configurado para computar o valor de referência a partir de uma representação do espectro sem curva temporal do sinal de áudio (210) ou a partir de uma representação de domínio de tempo sem curva temporal do sinal de áudio (210); e

onde o computador é configurado para formar um valor proporcional usando as informações sobre compactação de energia (122) descrevendo uma compactação de energia em uma representação do espectro transformada com curva temporal do sinal de áudio e do valor de referência, e para comparar o valor proporcional com um ou mais valores limite para obter o sinal de ativação com curva temporal como resultado da comparação.

34. Provedor do sinal de ativação com curva temporal (230;234) de acordo com qualquer uma das reivindicações 24 a 32, **caracterizado** pelo provedor do sinal de ativação com curva temporal compreende um calculador do

valor de referência configurado para computar o valor de referência a partir de uma representação com curva temporal do sinal de entrada (210), a curva temporal usando informações padrão de contorno de curva temporal (288); e

onde o comparador é configurado para formar um valor proporcional usando as informações de compactação de energia (234e) descrevendo uma compactação de energia em uma representação com curva temporal do sinal e áudio e o valor de referência, e para comparar o valor proporcional com um ou mais valores limite para obter o sinal de ativação com curva temporal como resultado da comparação.

35. Codificador de sinais de áudio (200) para codificar um sinal de áudio de entrada (210) para obter uma representação codificada (212) do sinal de áudio de entrada, **caracterizado** por compreender:

um transformador com curva temporal (220) configurado para fornecer uma representação espectral transformada com curva temporal (222) a partir do sinal de áudio de entrada (210) usando um contorno de curva temporal;

um provedor do sinal de ativação com curva temporal (100; 230; 234) de acordo com uma das afirmações 24 a 34, onde o provedor do sinal de ativação com curva temporal é configurado para receber o sinal do áudio de entrada (210) e para fornecer o sinal de ativação com curva temporal (112; 232; 234p); e

um controlador (240) configurado para fornecer seletivamente, dependendo do sinal de ativação com curva temporal (112; 232; 234p), informações de contorno de

curva temporal descobertas recentemente (286), descrevendo uma parte não constante do contorno de curva temporal ou informações padrão do contorno de curva temporal (288), descrevendo uma parte constante do contorno de curva temporal para o transformador de curva temporal (220) para descrever o contorno de curva temporal usado pelo transformador de curva temporal (220).

36. Codificador do sinal de áudio de acordo com a reivindicação 35, **caracterizado** pelo codificador do sinal de áudio compreender uma interface de saída (280) configurada para incluir a representação espectral transformada da curva temporal (222) na representação codificada (212) do sinal de áudio e

para incluir seletivamente, dependendo do sinal de ativação de curva temporal (232), informações de contorno de curva temporal na representação codificada (212) do sinal de áudio.

37. Método (400) para fornecer um sinal de ativação de curva temporal a partir de um sinal de áudio, **caracterizado** por compreender:

fornecimento (410) de informações de compactação de energia, descrevendo uma compactação de energia em uma representação espectral transformada com curva temporal do sinal de áudio;

comparação (420) das informações de compactação de energia com um valor de referência; e

fornecimento (430) do sinal de ativação de curva temporal dependendo do resultado da comparação.

38. Método (450) para codificar um sinal de áudio de entrada para obter uma representação codificada do sinal de áudio de entrada, **caracterizado** por compreender:

forneimento (470) de um sinal de ativação de curva temporal de acordo com a reivindicação 37, onde as informações de compactação de energia descrevem uma compactação de energia em uma representação do espectro transformada com curva temporal do sinal de áudio de entrada; e

forneimento seletivo (480), dependendo do sinal de ativação de curva temporal, de uma descrição da representação espectral transformada com curva temporal do sinal de áudio de entrada ou uma descrição de uma representação espectral sem transformação com curva temporal do sinal de áudio de entrada para inclusão na representação codificada do sinal de áudio de entrada.

39. Meio de armazenamento não transitório legível por máquina, **caracterizado por** conter conjunto de instruções que, quando executadas realizam as etapas do método de acordo com a reivindicação 37 ou as etapas do método de acordo com a reivindicação 38.

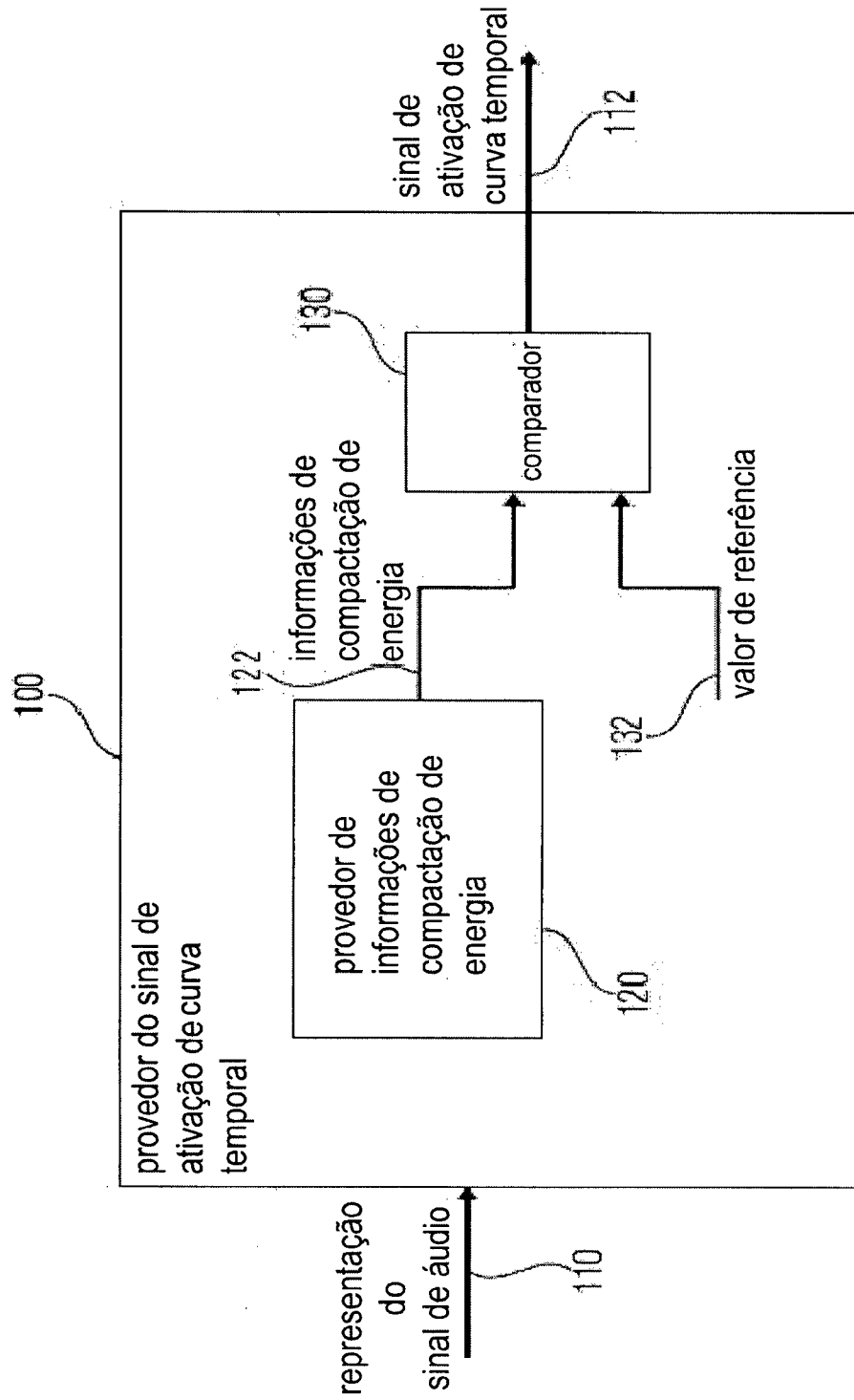


FIG 1

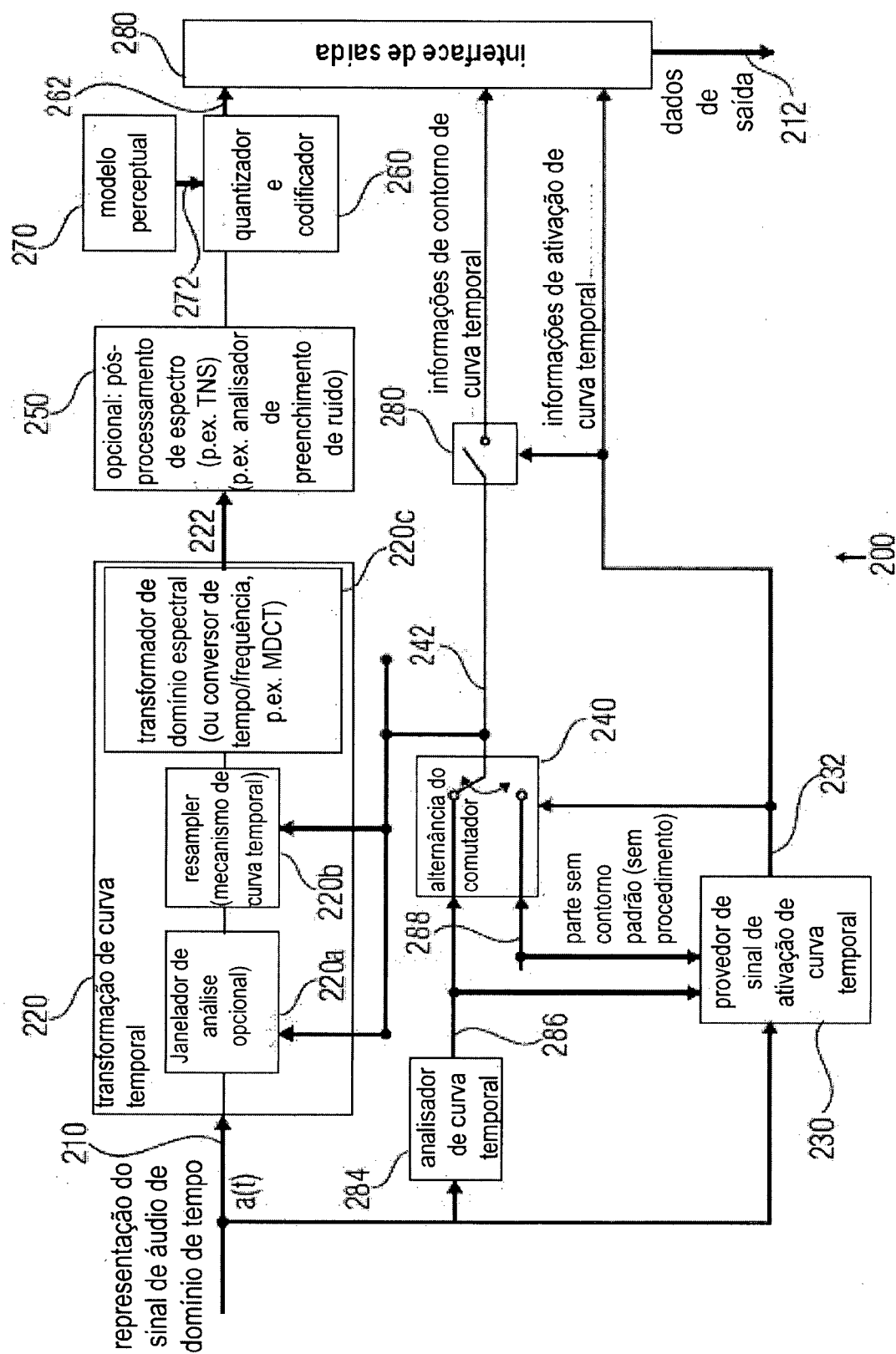


FIG 2A

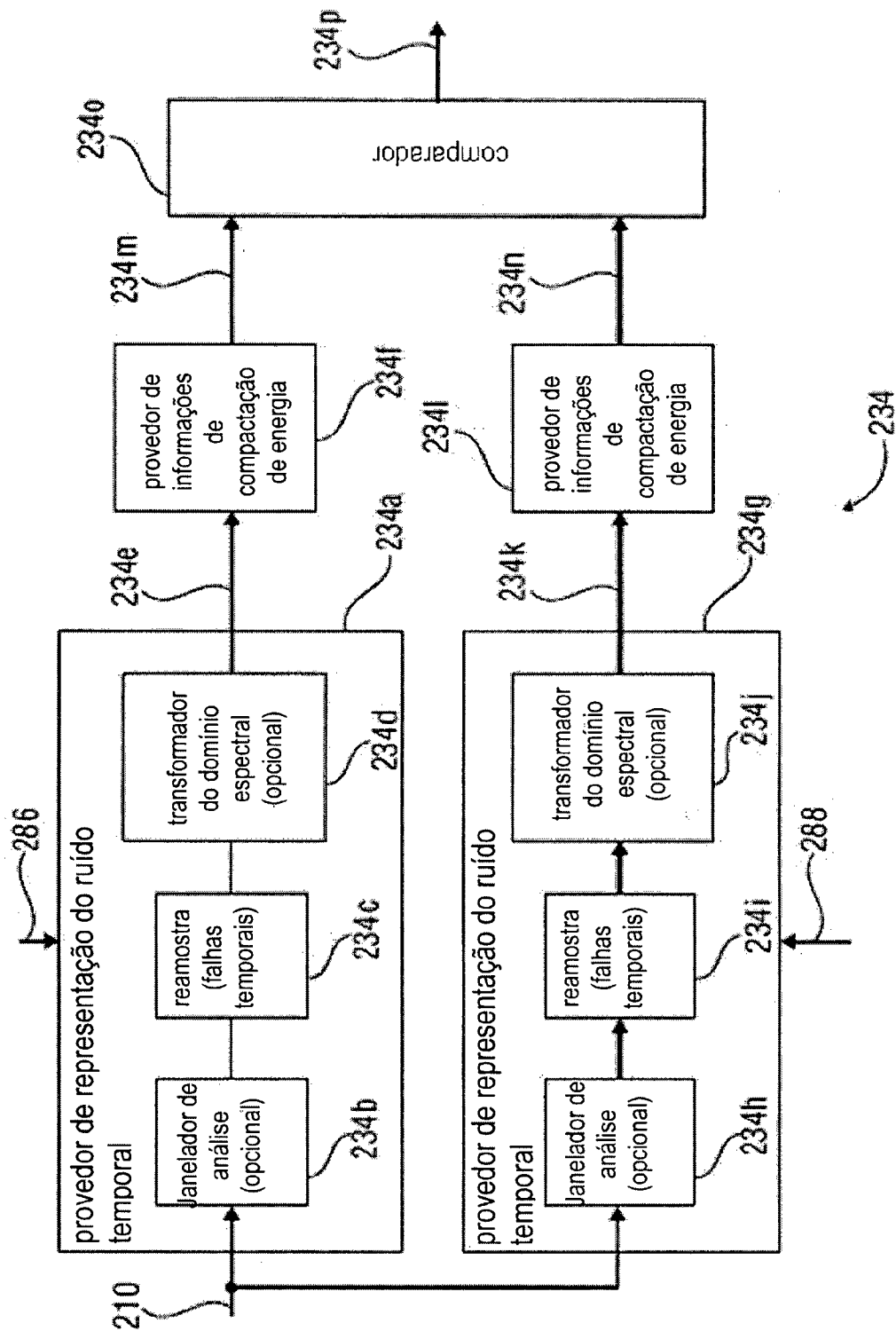
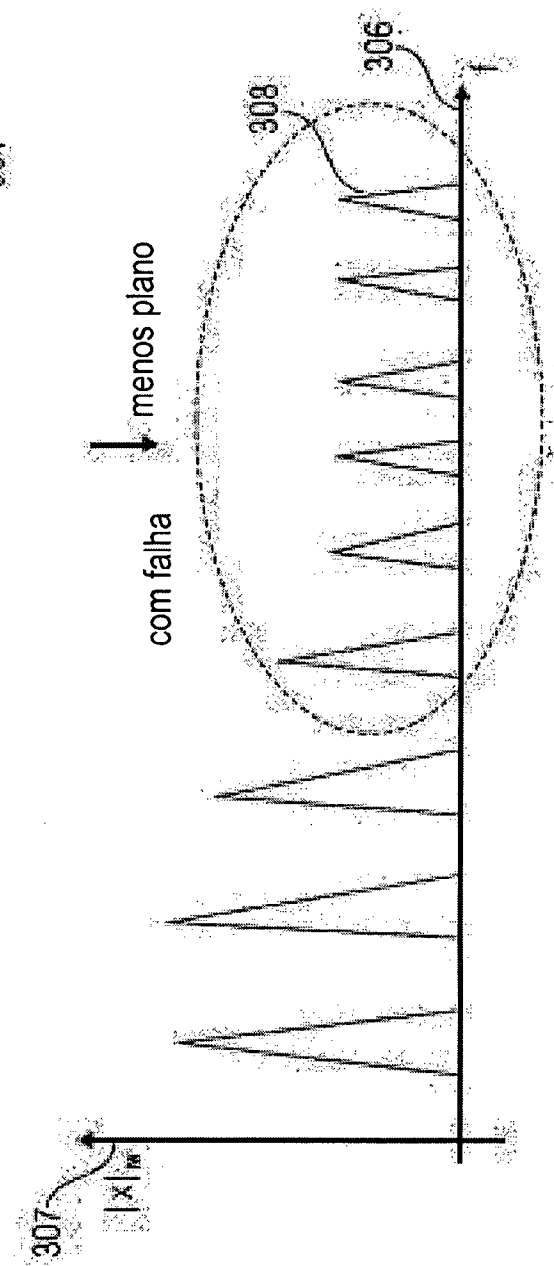
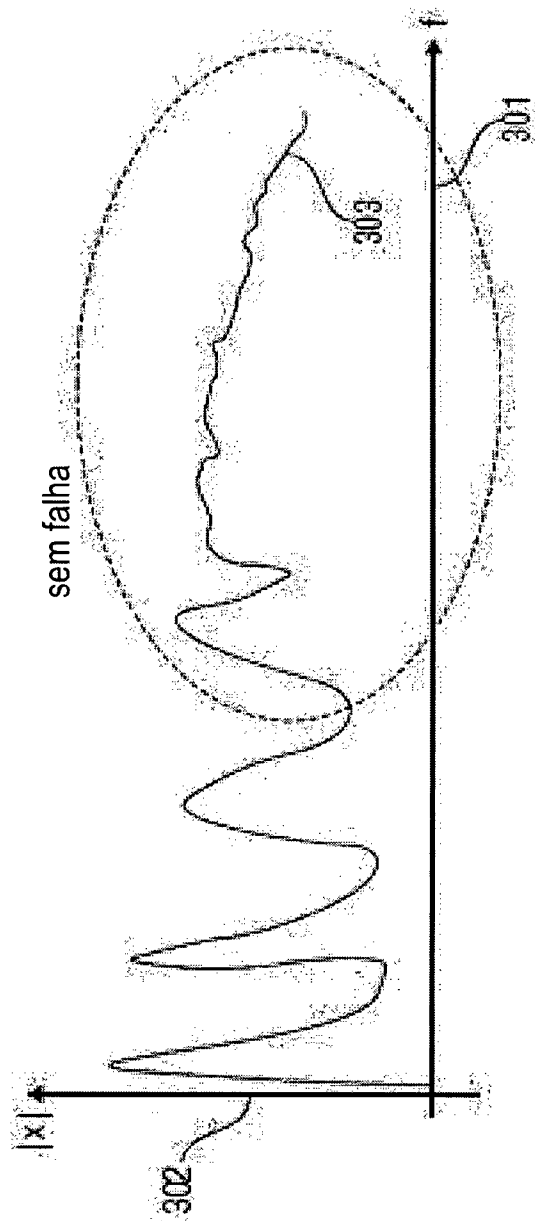


FIG 2B



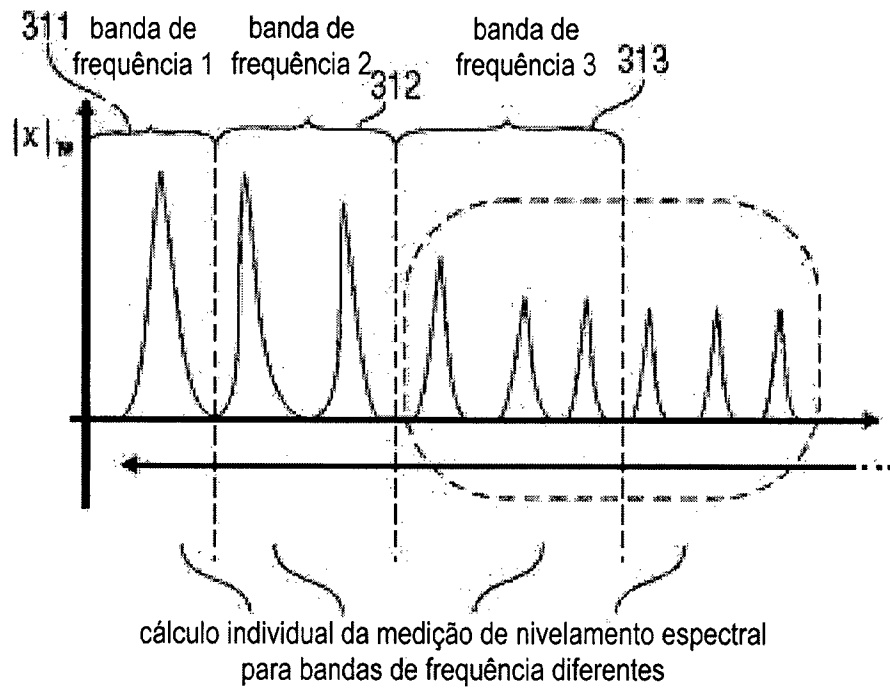


FIG 3C

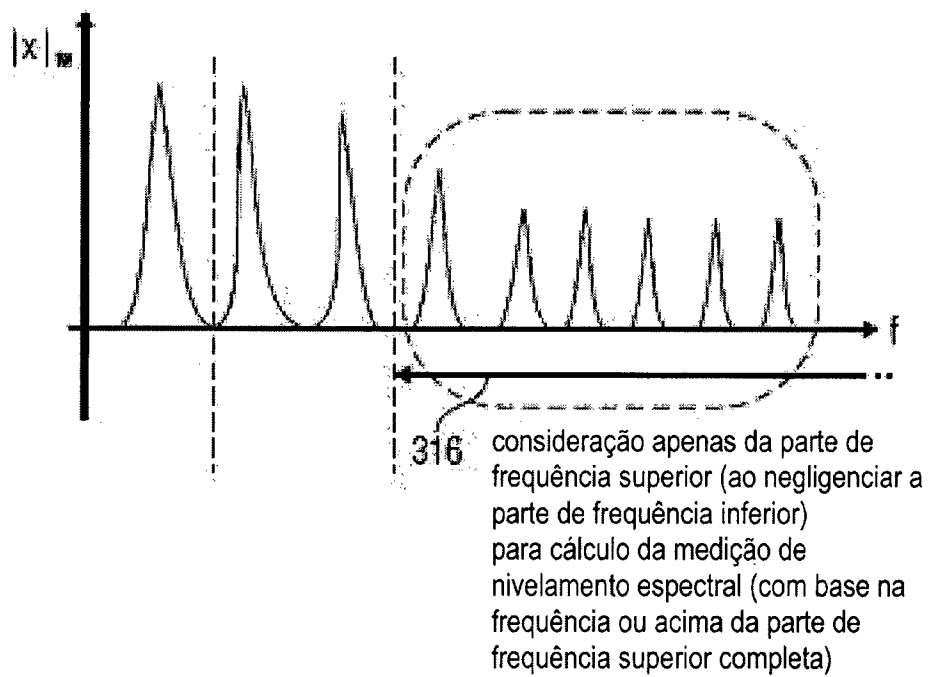


FIG 3D

6/25

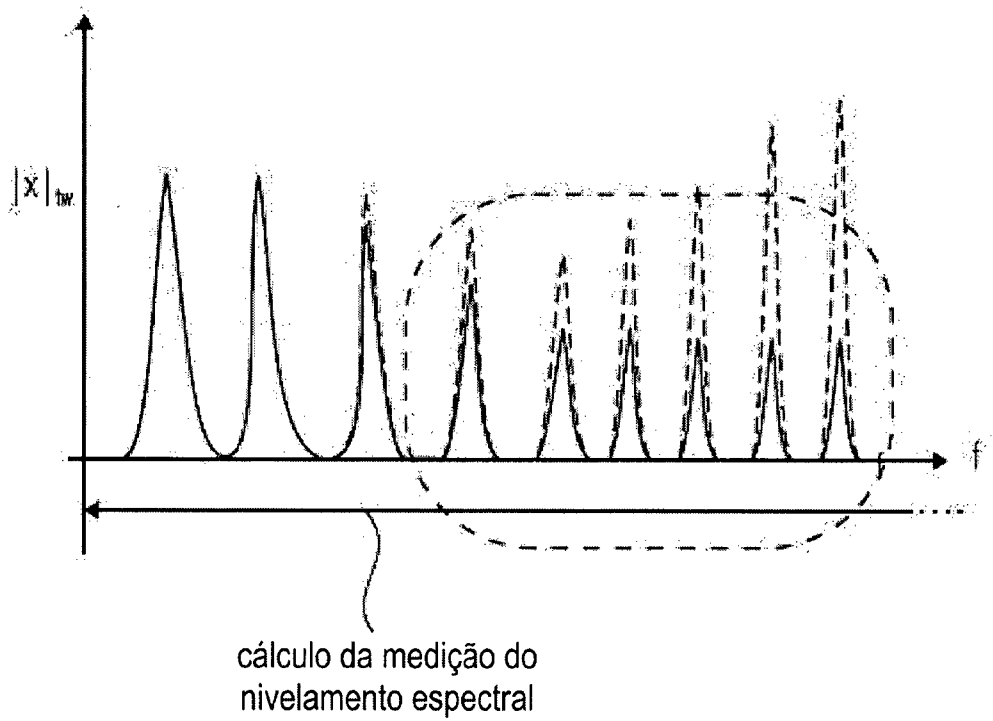


FIG 3E

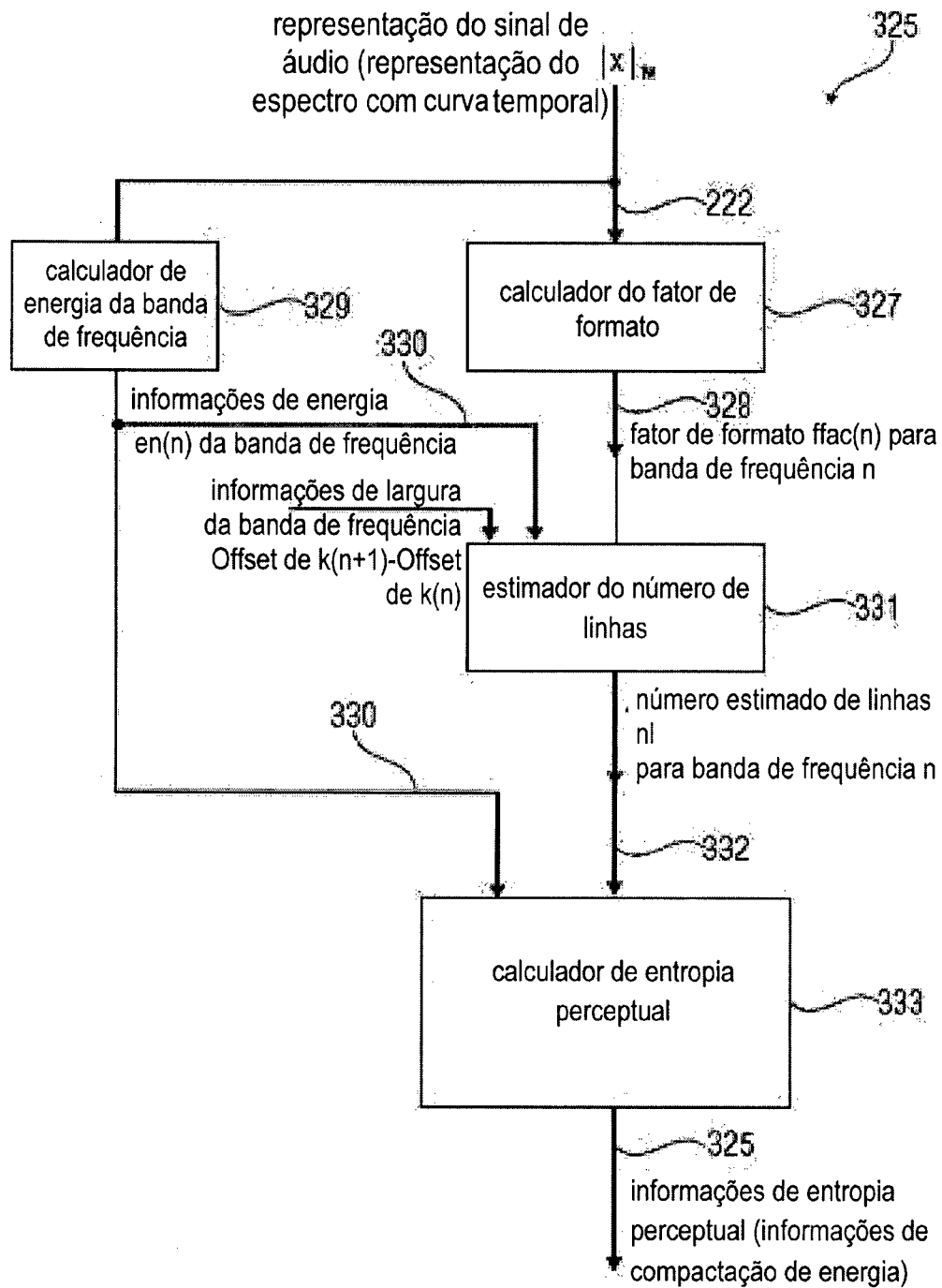


FIG 3F

8/25

sinal de tempo com
curva temporal

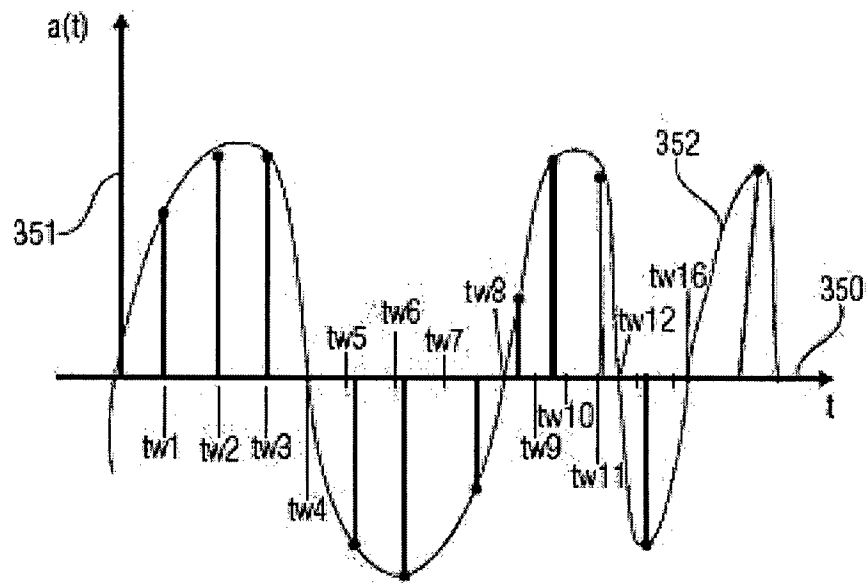


FIG 3G

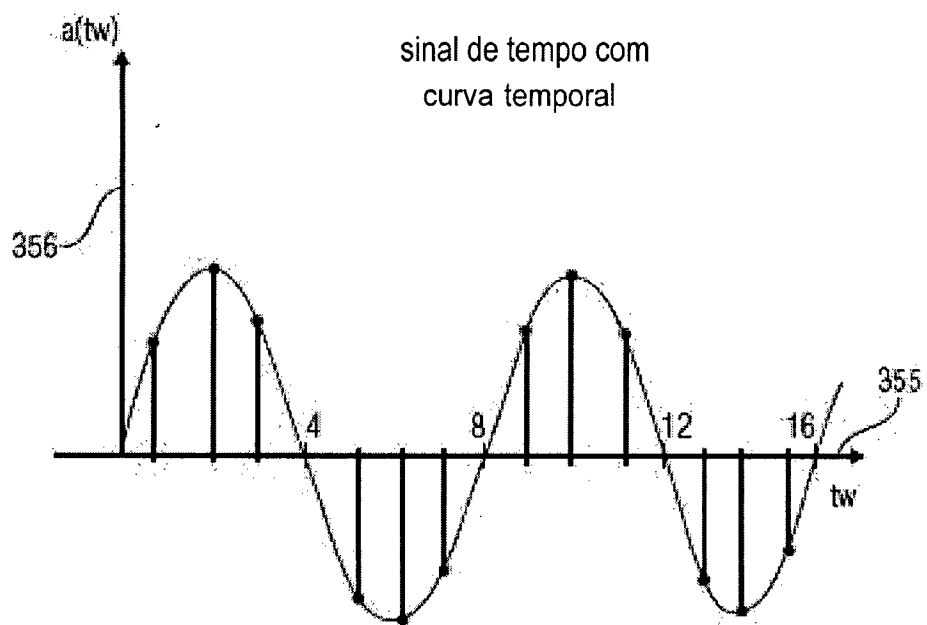


FIG 3H

9/25

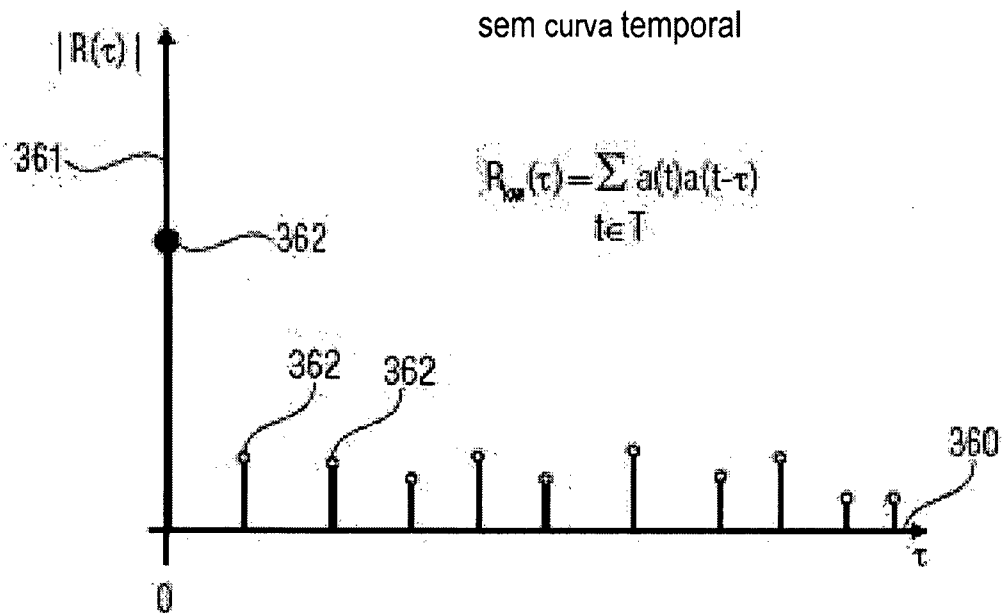


FIG 3I

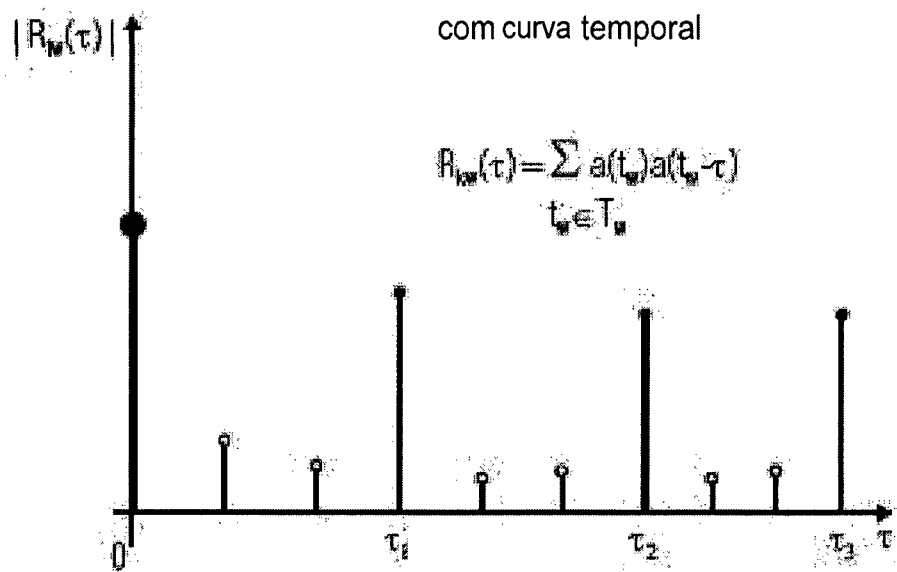


FIG 3J

10/25

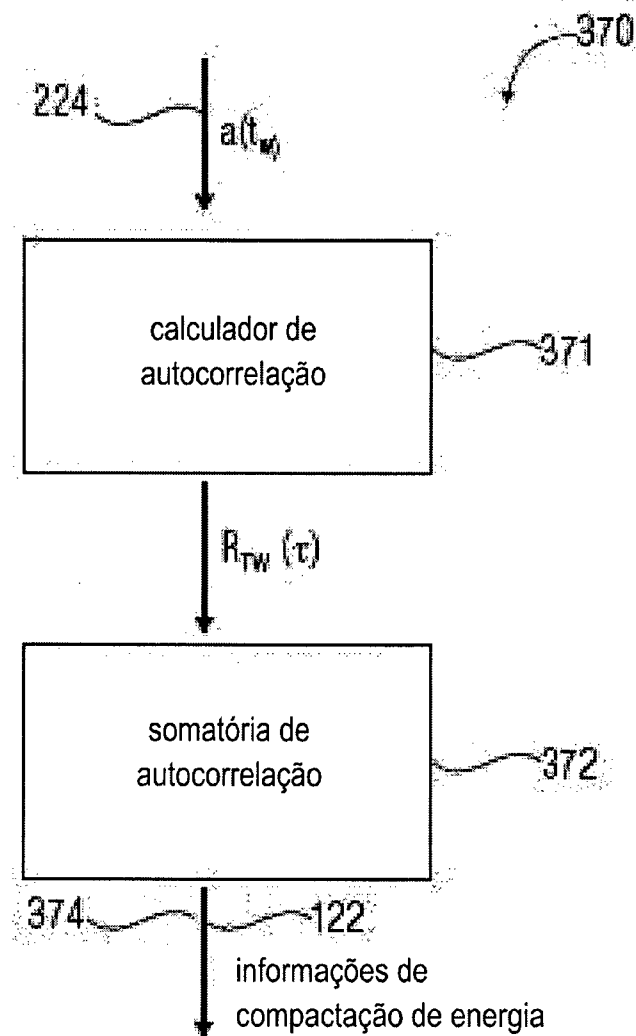


FIG 3K

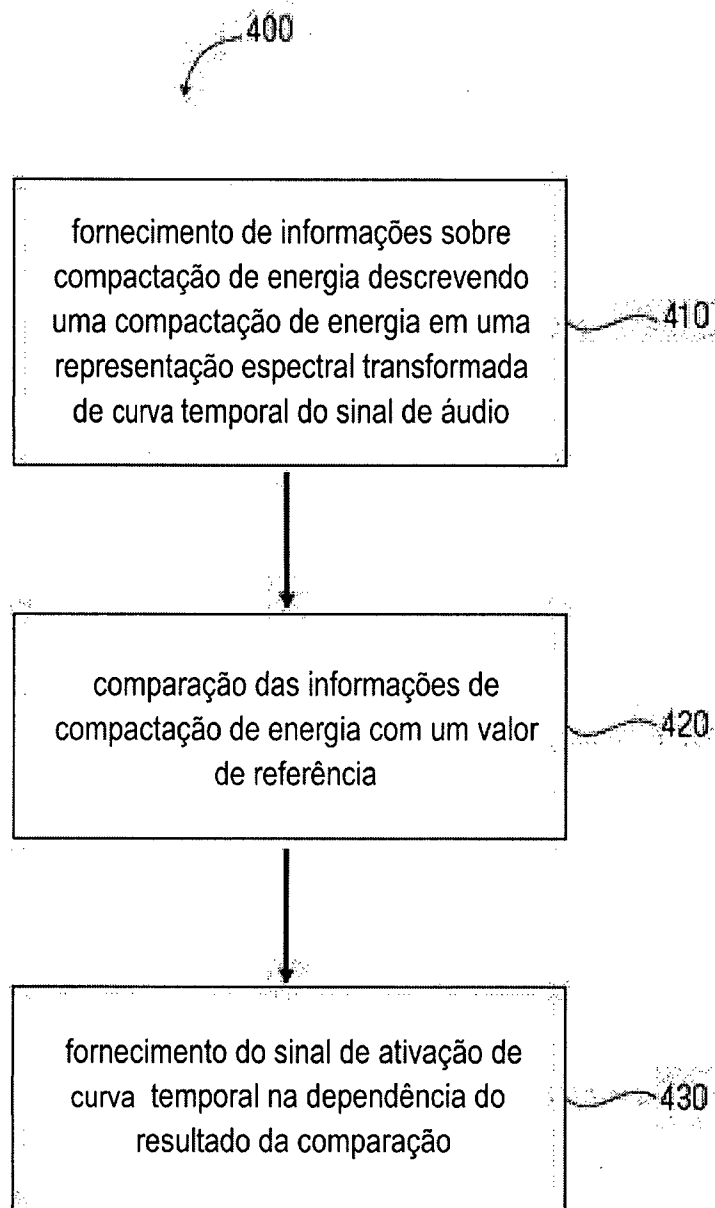


FIG 4A

12/25

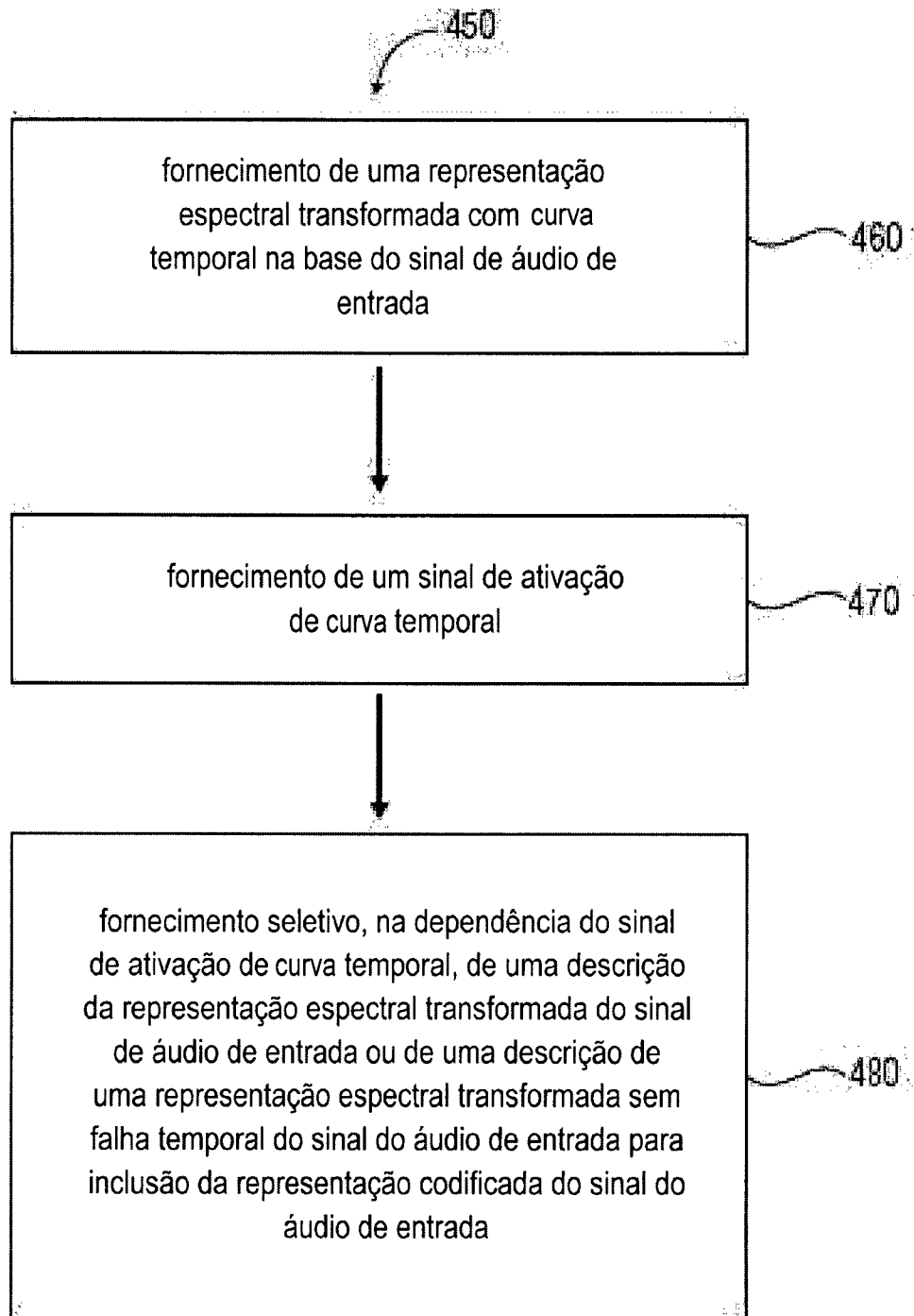


FIG 4B

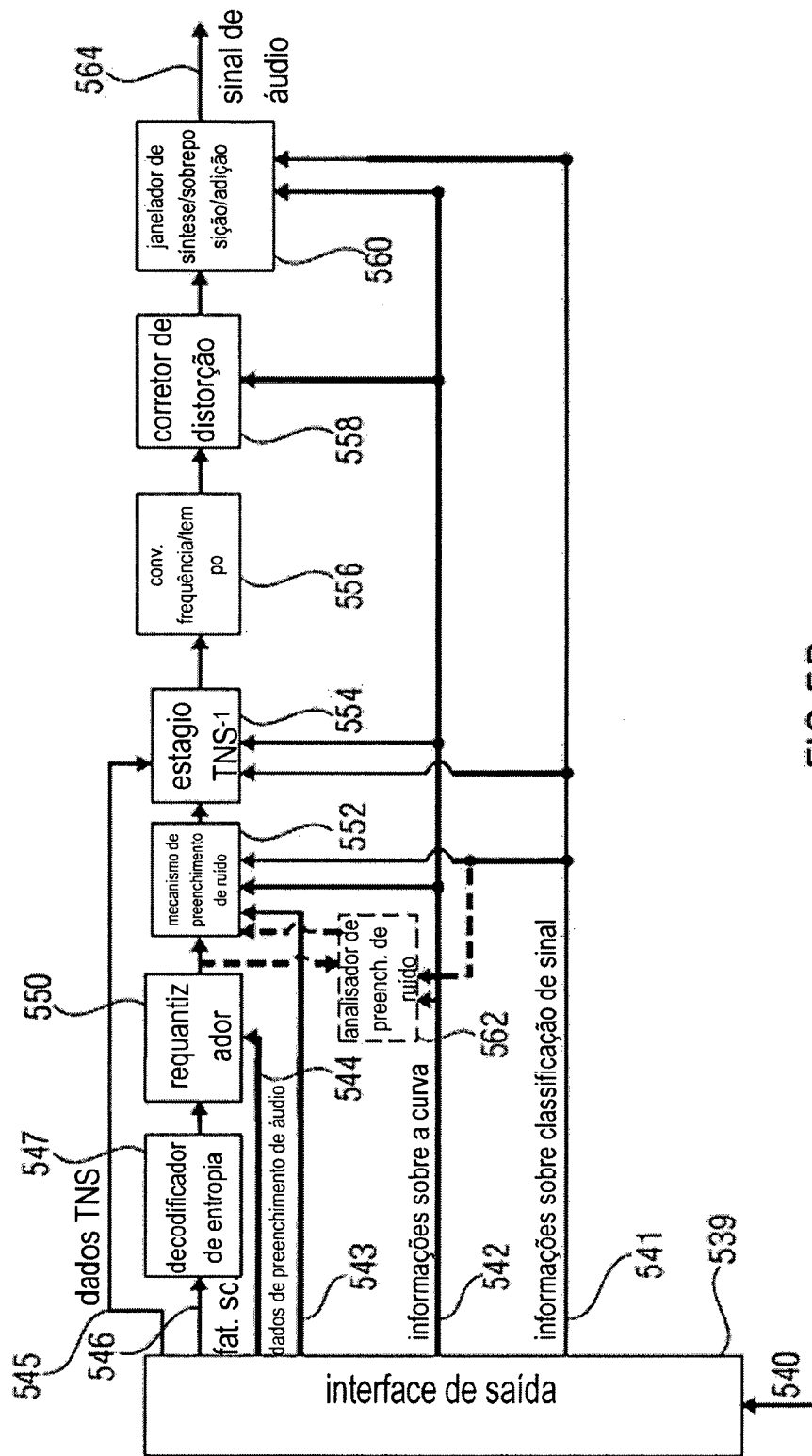


FIG 5B
(DECODIFICADOR)

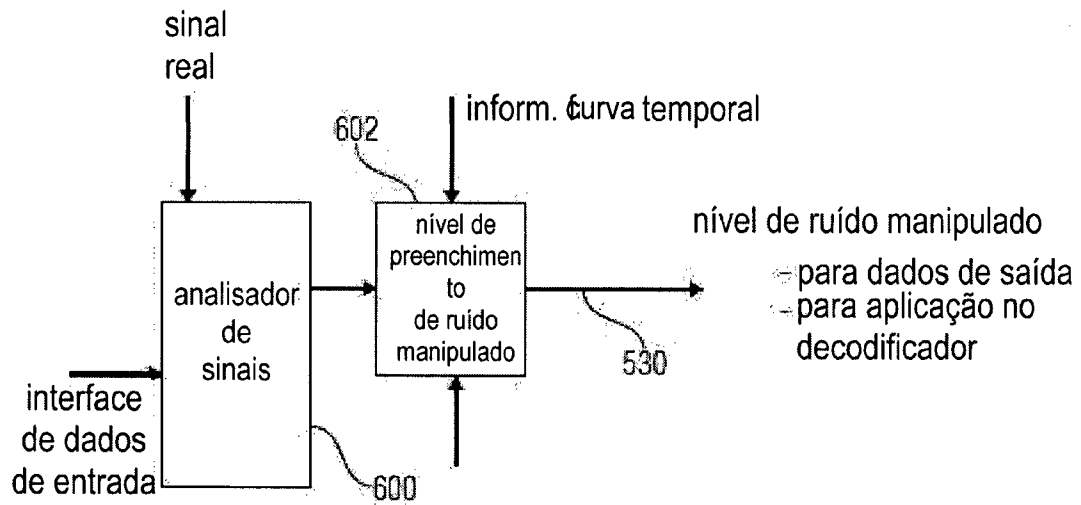


FIG 6A

classif. fala		curva temporal		classif. sinais		nível preench. Ruído
V	UV	SIM	NÃO	Fala	sem fala	
			X		X	normal
		X			X	baixo
X		X		X		muito baixo/zero
	X		X	X		normal

FIG 6B

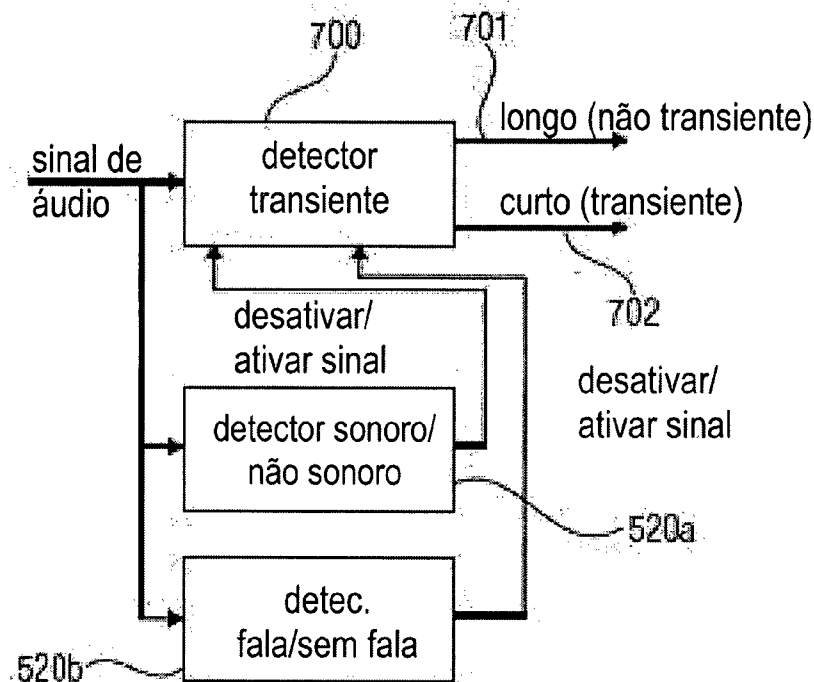


FIG 7A

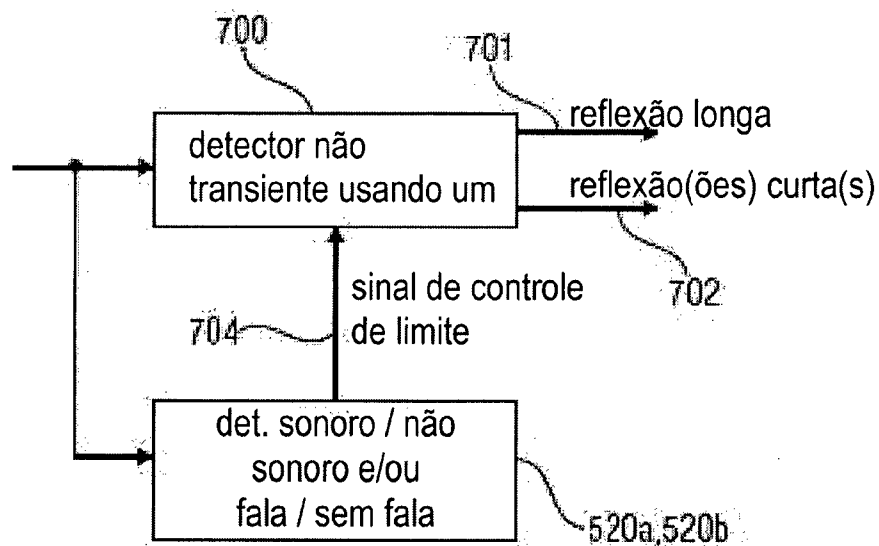


FIG 7B

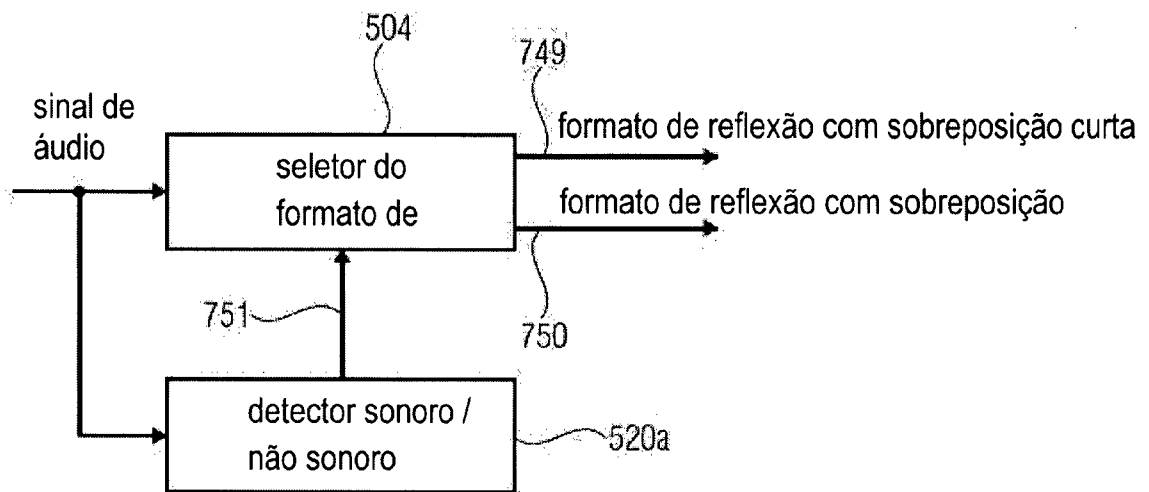


FIG 7C

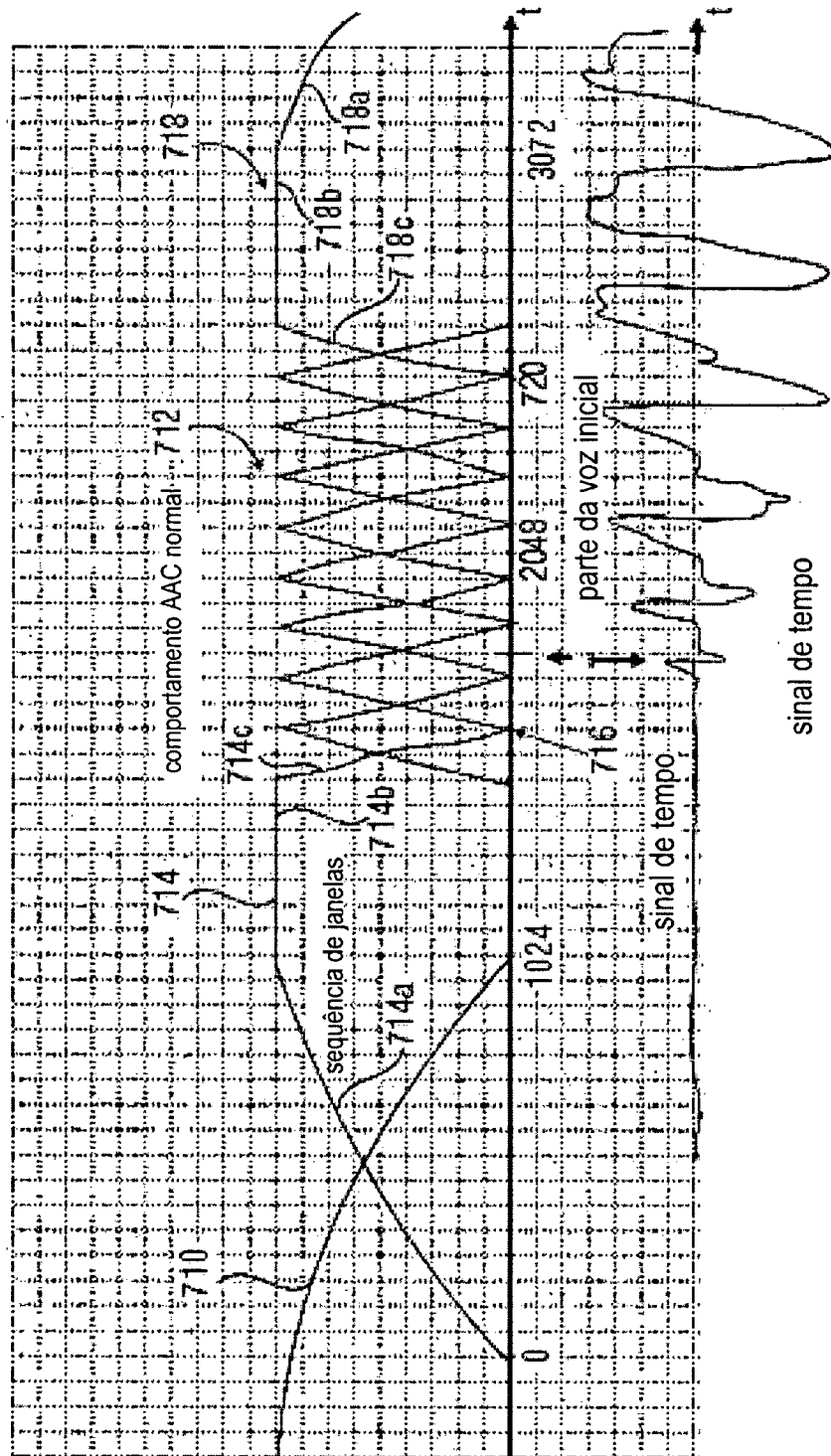


FIG 7D

com prevenção à alternância de bloco —
com detecção de ataques e sobreposição baixa - - -

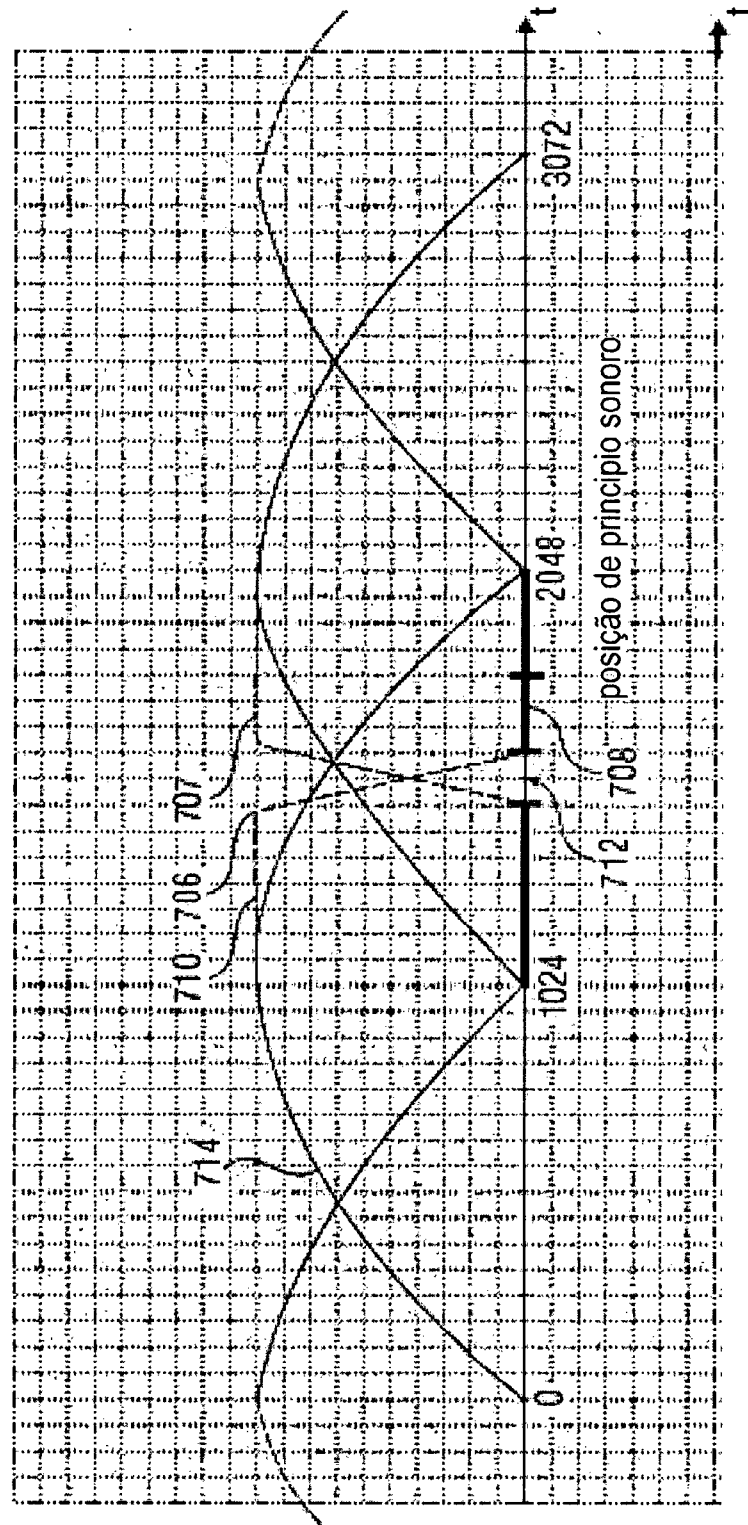


FIG 7E

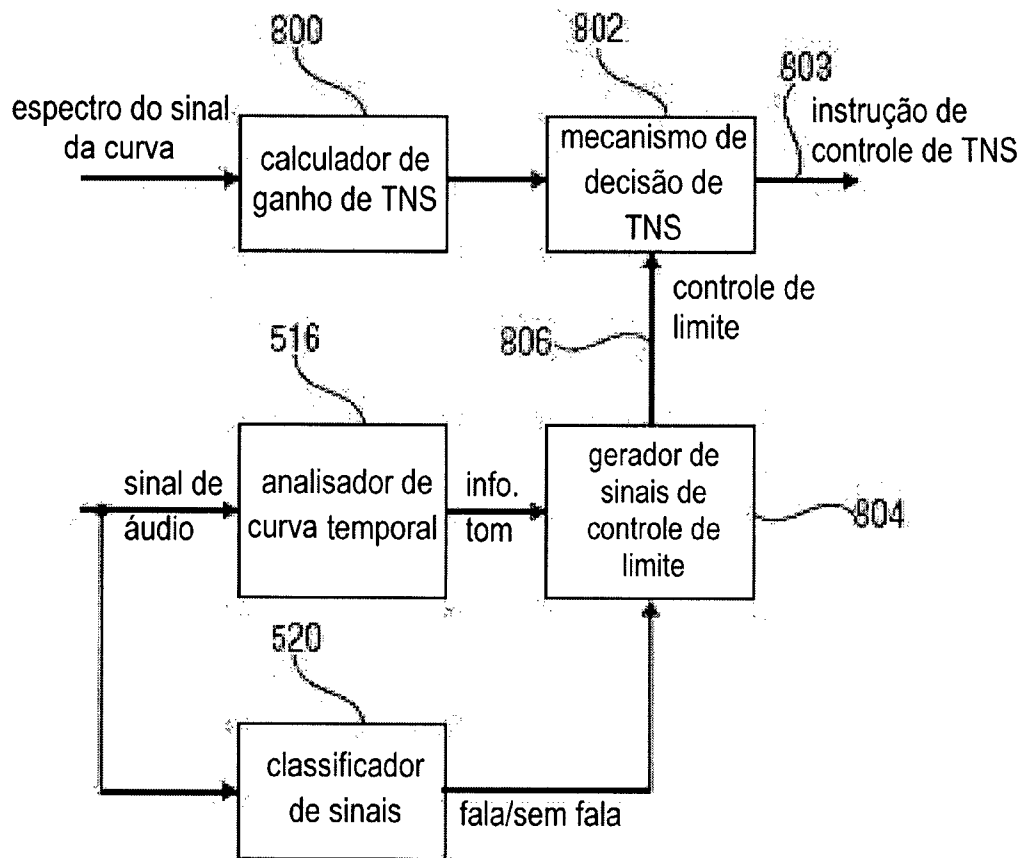


FIG 8A

		sonoro	não sonoro	limite de decis. TNS
controle de tom		classif. sinal		
SIM	NÃO	fala	sem fala	
	X		X	normal
X			X	inferior
X		X		(igual) inferior

FIG 8B

FIG 9A

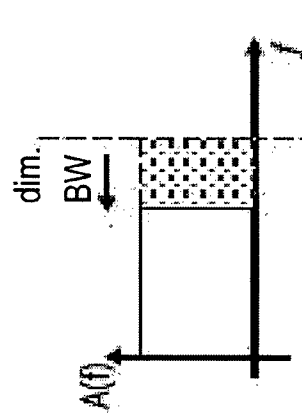
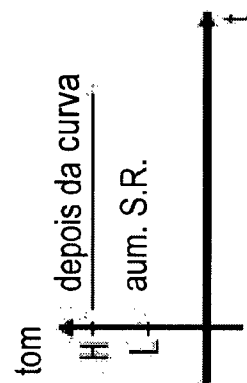
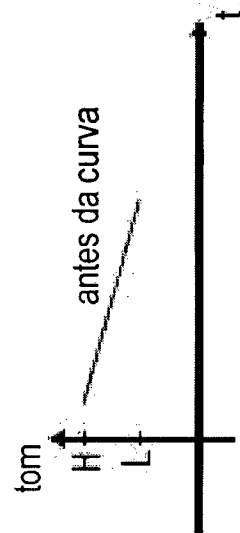


FIG 9B

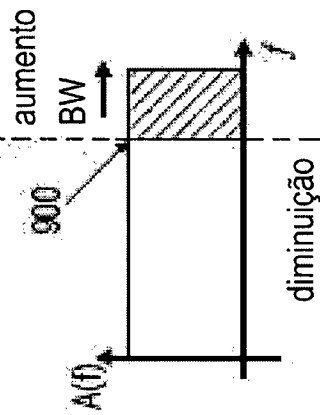
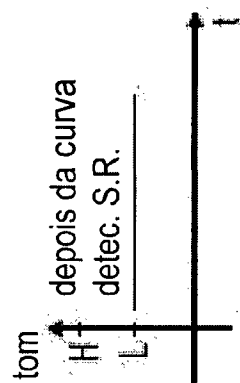
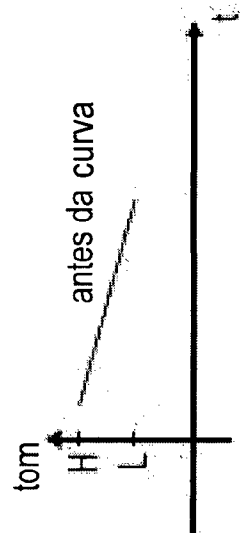


FIG 9C

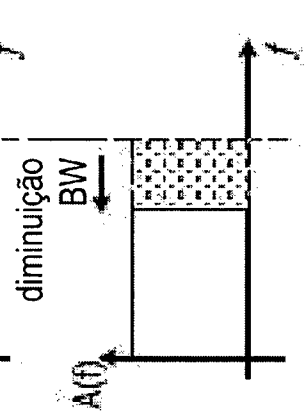
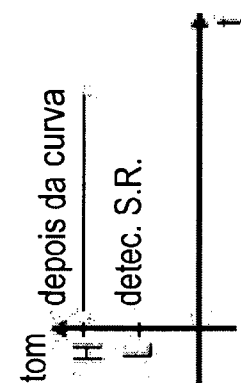
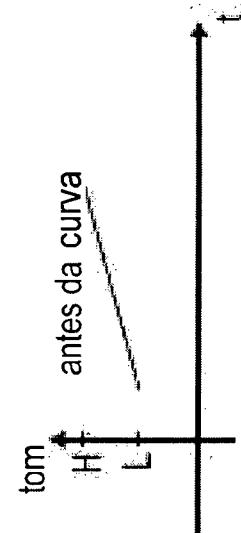


FIG 9D

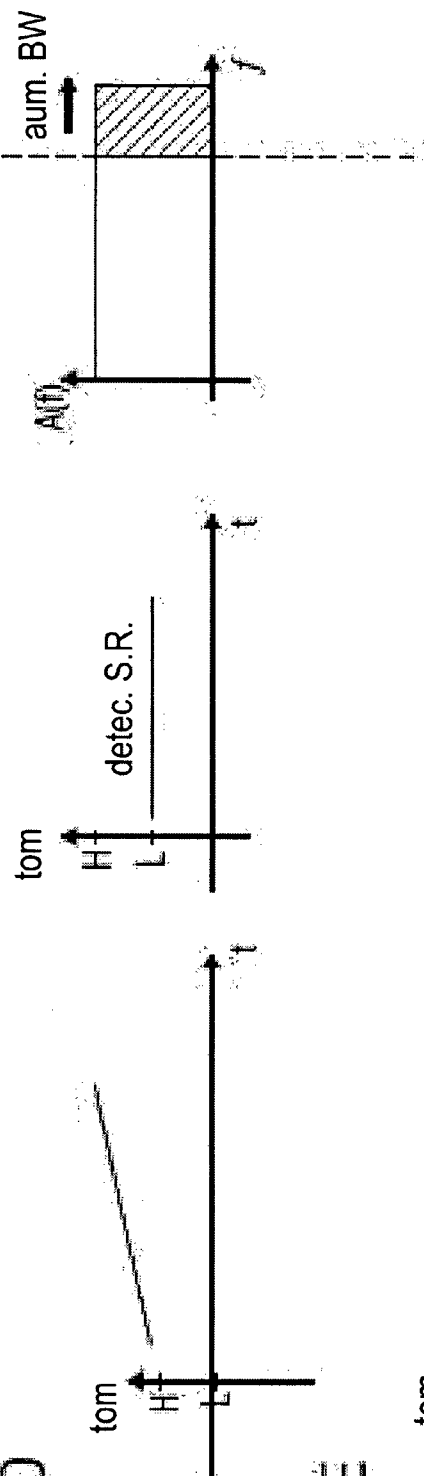
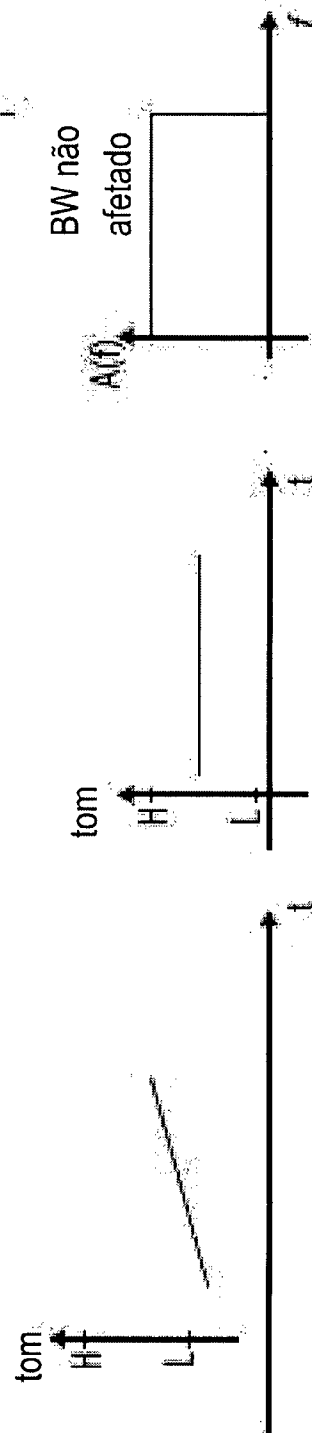


FIG 9E



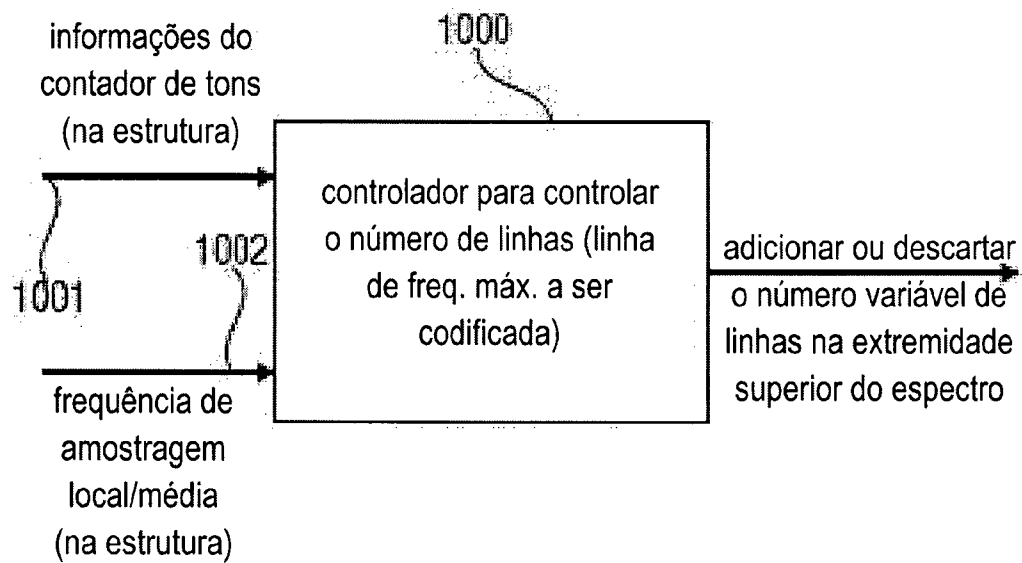


FIG 10A

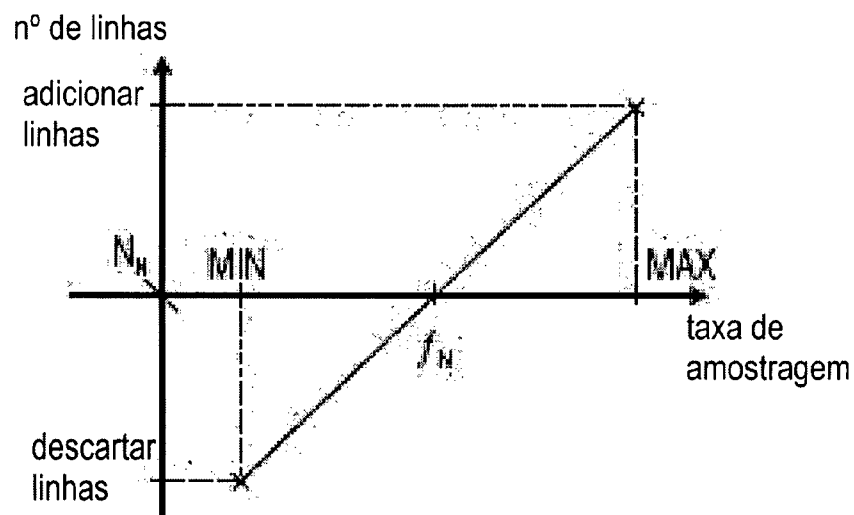


FIG 10B

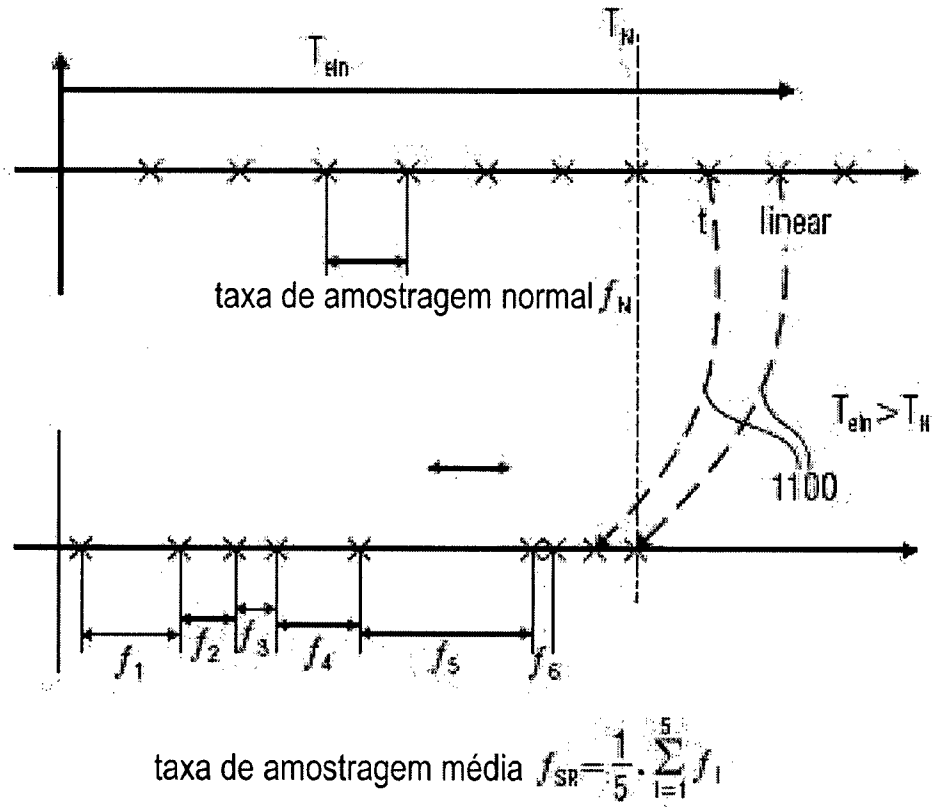


FIG 11

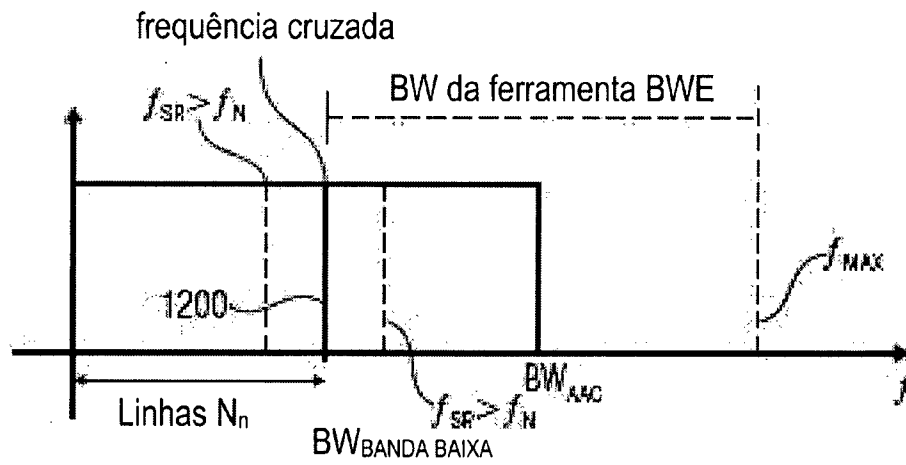


FIG 12A

f_{sr} : taxa de amostragem local
 Δf : largura de uma linha
espectral/distância entre linhas
espectrais
 N : número de linhas conforme
determinado pelo
processador
 T_{ah} : comprimento da estrutura
sem curva (depende da
característica de curva)

T_{ah}	f_{sr}	Δf	N
$=T_N$	$=f_N$	$=\Delta f_N$	$=N_N$
$>T_N$	$>f_N$	$>\Delta f_N$	$<N_N$ excluir linhas
$<T_N$	$<f_N$	$<\Delta f_N$	$>N_N$ adicionar linhas

FIG 12B