



República Federativa do Brasil

Ministério do Desenvolvimento, Indústria,
Comércio e Serviços

Instituto Nacional da Propriedade Industrial



(11) BR 122022008596-2 B1

(22) Data do Depósito: 27/10/2014

(45) Data de Concessão: 31/01/2023

(54) Título: DECODIFICADOR DE ÁUDIO E MÉTODO PARA FORNECER UMA INFORMAÇÃO DE ÁUDIO DECODIFICADA UTILIZANDO UMA DISSIMULAÇÃO DE ERRO QUE MODIFICA UM SINAL DE EXCITAÇÃO NO DOMÍNIO DE TEMPO

(51) Int.Cl.: G10L 19/005; G10L 25/90; G10L 19/08; G10L 19/02.

(30) Prioridade Unionista: 28/07/2014 EP 14178825; 31/10/2013 EP 13191133.

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

(72) Inventor(es): JÉRÉMIE LECOMTE.

(86) Pedido PCT: PCT EP2014073036 de 27/10/2014

(87) Publicação PCT: WO 2015/063045 de 07/05/2015

(85) Data do Início da Fase Nacional: 04/05/2022

(62) Pedido Original do Dividido: BR112016009805-6 - 27/10/2014

(57) Resumo: Decodificador áudio (200;400) para fornecer uma informação de áudio decodificada (220; 412) baseada em uma informação de áudio codificada (210;410). O decodificador áudio compreende uma dissimulação de erro (240; 480; 600) configurada para fornecer uma informação de áudio de dissimulação de erro (242;482;612) para dissimular uma perda de um quadro de áudio, em que a dissimulação de erro está configurada para modificar o sinal de excitação de domínio de tempo (452,456;610) obtido para um ou mais quadros de áudio que antecedem um quadro de áudio perdido, de modo a obter a informação de áudio da dissimulação de erro.

RELATÓRIO DESCRIPTIVO

"DECODIFICADOR DE ÁUDIO E MÉTODO PARA FORNECER UMA INFORMAÇÃO DE ÁUDIO DECODIFICADA UTILIZANDO UMA DISSIMULAÇÃO DE ERRO QUE MODIFICA UM SINAL DE EXCITAÇÃO NO DOMÍNIO DE TEMPO"

Pedido Dividido do pedido de patente No. BR 11 2016 009805 6 depositado em 27/10/2014.

[001] Especificação

[002] Área Técnica

[003] Os modelos de acordo com a invenção criam decodificadores áudio para fornecerem uma informação de áudio decodificada baseada em uma informação de áudio codificada.

[004] Alguns modelos de acordo com a invenção criam métodos para fornecerem uma informação de áudio decodificada baseada em uma informação de áudio codificada.

[005] Alguns modelos de acordo com a invenção criam programas de computador para executarem um dos ditos métodos.

[006] Alguns modelos de acordo com a invenção dizem respeito a uma dissimulação de domínio do tempo para um codec no domínio da transformação.

[007] Antecedentes da Invenção

[008] Em anos recentes existe um aumento da procura de uma transmissão digital e armazenamento de conteúdos de áudio. Contudo, os conteúdos de áudio são frequentemente transmitidos sobre canais inseguros, trazendo consigo o risco de as unidades de dados (por exemplo, pacotes) que compreendem um ou mais quadros de áudio (por exemplo, sob a forma de uma representação codificada, como, por exemplo, uma representação de domínio de tempo codificada ou uma representação de domínio de tempo codificada) são perdidas. Em algumas situações, será possível solicitar uma repetição (reenvio) de quadros de áudio (ou de unidades de dados, como pacotes,

compreendendo um ou mais quadros de áudio perdidos). Contudo, isto iria habitualmente originar um atraso significativo, e iria assim exigir um *buffering* extensivo de quadros de áudio. Noutros casos, é dificilmente possível solicitar uma repetição de quadros de áudio perdidos.

[009] De modo a obter uma boa qualidade de áudio, ou pelo menos aceitável, visto que os quadros de áudio se perdem sem fornecimento de *buffering* extensivo (que iria consumir uma grande quantidade de memória e que iria também significativamente degradar as capacidades em tempo real da codificação de áudio) é desejável existirem conceitos que tragam uma boa qualidade de áudio, ou pelo menos uma qualidade de áudio aceitável, mesmo no caso em que os quadros de áudio se percam.

[010] No passado, alguns conceitos de dissimulação de erros foram desenvolvidos, podendo ser aplicados em diferentes conceitos de codificação de áudio.

[011] A seguir será descrito um conceito de codificação de áudio convencional.

[012] No padrão 3gpp TS 26.290, uma decodificação de excitação codificada por transformação (decodificação TCX) com dissimulação de erros é explicada. A seguir, algumas explicações serão fornecidas, baseadas na seção “decodificação do modo TCX e síntese do sinal” na referência [1].

[013] Um decodificador TCX de acordo com o Padrão Internacional 3gpp TS 26.290 é ilustrado nas Figs. 7 e 8, em que as Figs. 7 e 8 ilustram diagramas de blocos do decodificador TCX. Contudo, a Fig. 7 ilustra aqueles blocos relevantes para a decodificação TCX em uma operação normal ou em um caso de uma perda parcial de pacotes. Pelo contrário, a Fig. 8 ilustra o processamento relevante da decodificação TCX no caso de dissimulação por eliminação de pacotes TCX-256.

[014] Por outras palavras, as Figs. 7 e 8 ilustram um diagrama de blocos do decodificador TCX incluindo os seguintes casos:

[015] Caso 1 (Fig. 8): Dissimulação por eliminação de pacotes em TCX-

256, quando o comprimento do quadro TCX é de 256 amostras e o respetivo pacote foi perdido, ou seja, $BFI_TCX = (1)$; e

[016] Caso 2 (Fig. 7): Decodificação TCX normal, possivelmente com perdas parciais de pacotes.

[017] A seguir algumas explicações serão fornecidas relativamente às Figs. 7 e 8.

[018] Tal como referido, a Fig. 7 ilustra um diagrama de blocos de um decodificador TCX que executa uma decodificação TCX em operação normal ou no caso de perda parcial de pacotes. O decodificador TCX 700 de acordo com a Fig. 7 recebe parâmetros TCX específicos 710 e fornece, baseado nisso, informação de áudio decodificada 712, 714.

[019] O decodificador áudio 700 comprehende um desmultiplexador “DEMUX TCX 720”, configurado para receber os parâmetros TCX específicos 710 e a informação “ BFI_TCX ”. O desmultiplexador 720 separa os parâmetros TCX específicos 710 e fornece uma informação de excitação codificada 722, uma informação de preenchimento de ruído codificada 724 e uma informação de ganho global codificada 726. O decodificador áudio 700 comprehende um decodificador de excitação 730, configurado para receber a informação de excitação codificada 722, a informação de preenchimento de ruído codificada 724 e a informação de ganho global codificada 726, assim como alguma informação adicional (como, por exemplo, uma indicação de baixo débito “ bit_rate_flag ”, uma informação “ BFI_TCX ” e uma informação de comprimento do quadro TCX. O decodificador de excitação 730 fornece, baseado nisto, um sinal de excitação de domínio de tempo 728 (também designado com “ X ”). O decodificador de excitação 730 comprehende um processador de informação de excitação 732, que desmultiplexa a informação de excitação codificada 722 e decodifica parâmetros de quantização de vetores algébricos. O processador de informação de excitação 732 fornece um sinal de excitação intermédio 734, que habitualmente se encontra em uma representação de domínio de frequência, e que é indicado com Y . O codificador de excitação 730

compreende também um injetor de ruído 736, configurado para injetar ruído em sub-bandas não quantizadas, para derivar um sinal de excitação preenchido de ruído 738 a partir do sinal de excitação intermédio 734. O sinal de excitação preenchido de ruído 738 está habitualmente de domínio de frequência, e é indicado com Z. O injetor de ruído 736 recebe uma informação de intensidade de ruído 742 a partir de um decodificador de nível de preenchimento de ruído 740. O decodificador de excitação comprehende também uma desacentuação de baixa frequência adaptável 744, configurada para executar uma operação de desacentuação de baixa frequência baseada no sinal de excitação preenchido de ruído 738, para desse modo obter um sinal de excitação processado 746, ainda de domínio de frequência, e indicado com X'. O decodificador de excitação 730 comprehende também um transformador de domínio de frequência para o domínio de tempo 748, configurado para receber o sinal de excitação processado 746 e para fornecer, baseado nisso, um sinal de excitação de domínio de tempo 750, associado a uma certa porção de tempo representada por um conjunto de parâmetros de domínio de frequência (por exemplo, do sinal de excitação processado 746). O decodificador de excitação 730 comprehende também um escalador 752, configurado para escalonamento do sinal de domínio de tempo 750 para desse modo obter um sinal de excitação de domínio de tempo de escala 754. O escalador 752 recebe uma informação de ganho global 756 do decodificador de ganho global 758, em que, por sua vez, o decodificador de ganho global 758 recebe a informação de ganho global codificada 726. O decodificador de excitação 730 comprehende também uma síntese de sobreposição e adição 760, que recebe sinais de excitação de domínio de tempo de escala 754 associados a uma pluralidade de porções de tempo. A síntese de sobreposição e adição 760 executa uma operação de sobreposição e adição (que pode incluir uma operação de janelas) baseada nos sinais de excitação de domínio de tempo de escala 754 para obter um sinal de excitação de domínio de tempo temporariamente combinado 728 para um período mais longo no tempo (mais longo do que os períodos no

tempo para os quais os sinais de excitação de domínio de tempo individuais 750, 754 são fornecidos).

[020] O decodificador áudio 700 comprehende também uma síntese LPC 770, que recebe o sinal de excitação de domínio de tempo 728 fornecido pela síntese de sobreposição e adição 760 e um ou mais coeficientes LPC que define uma função do filtro de síntese LPC 772. A síntese LPC 770 pode, por exemplo, compreender um primeiro filtro 774, que pode, por exemplo, filtrar por síntese o sinal de excitação de domínio de tempo 728, para desse modo obter um sinal de áudio descodificado 712. Como opção, a síntese LPC 770 pode compreender também um segundo filtro de síntese 774 utilizando outra função do filtro de síntese, para desse modo obter um sinal de áudio descodificado 714.

[021] A seguir, a decodificação TCX irá ser descrita no caso de uma dissimulação por eliminação de pacotes TCX-256. A Fig. 8 ilustra um diagrama de blocos do decodificador TCX neste caso.

[022] A dissimulação por eliminação de pacotes 800 recebe uma informação de *pitch* 810, também indicada com “pitch_tcx”, e que é obtida a partir do quadro TCX descodificado anterior. Por exemplo, a informação de *pitch* 810 pode ser obtida utilizando um estimador de *pitch* dominante 747 a partir do sinal de excitação processado 746 no decodificador de excitação 730 (durante a decodificação “normal”). Além disso, a dissimulação por eliminação de pacotes 800 recebe parâmetros LPC 812, que podem representar uma função do filtro de síntese LPC. Os parâmetros LPC 812 podem, por exemplo, ser idênticos aos parâmetros LPC 772. Por conseguinte, a dissimulação por eliminação de pacotes 800 pode ser configurada para fornecer, baseada na informação de distância 810 e nos parâmetros LPC 812, um sinal de dissimulação de erros 814, que pode ser considerada como uma informação de áudio de dissimulação de erro. A dissimulação por eliminação de pacotes 800 comprehende um tampão de excitação 820, que pode, por exemplo, servir de

tampão de uma excitação anterior. O tampão de excitação 820 pode, por exemplo, utilizar o livro de códigos adaptável de ACELP, e pode fornecer um sinal de excitação 822. A dissimulação por eliminação de pacotes 800 pode compreender ainda um primeiro filtro 824, uma função do filtro a qual pode ser definida tal como ilustrado na Fig. 8. Desse modo, o primeiro filtro 824 pode filtrar o sinal de excitação 822 baseado nos parâmetros LPC 812, para obter uma versão filtrada 826 do sinal de excitação 822. A dissimulação por eliminação de pacotes compreende também um limitador de amplitude 828, que pode limitar uma amplitude do sinal de excitação filtrado 826 baseado na informação alvo ou informação de nível rms_{wsyn} . Além disso, a dissimulação por eliminação de pacotes 800 pode compreender um segundo filtro 832, que pode ser configurado para receber o sinal de excitação filtrado de amplitude limitada 830 do limitador de amplitude 822 e para fornecer, baseado nisso, o sinal de dissimulação de erros 814. Uma função do filtro do segundo filtro 832 pode, por exemplo, ser definida tal como ilustrada na Fig. 8.

[023] A seguir, alguns detalhes relativos à decodificação e dissimulação de erros serão descritos.

[024] No Caso 1 (dissimulação por eliminação de pacotes em TCX-256), não existe informação disponível para decodificar a amostra 256 do quadro TCX. A síntese TCX é encontrada através do processamento da excitação anterior atrasada por T , em que $T=pitch_tcx$ é um desfasamento de *pitch* estimado no quadro TCX previamente decodificado, através de um filtro não linear aproximadamente equivalente a $1/\hat{A}(z)$. Um filtro não linear é utilizado em vez de $1/\hat{A}(z)$ para evitar cliques na síntese. Este filtro é decomposto em três passos:

[025] **Passo 1:** filtragem através de

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

[026] para mapear a excitação atrasada por T no domínio alvo TCX;

[027] **Passo 2:** aplicação de um limitador (a magnitude é limitada para \pm

$rms_{wsyn})$

[028] **Passo 3:** filtragem através de

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

[029] para procurar a síntese. De salientar que o *buffer OVLP_TCX* é definido como zero neste caso.

[030] **Decodificação dos parâmetros algébricos VQ**

[031] No Caso 2, a decodificação TCX envolve decodificação dos parâmetros algébricos VQ descrevendo cada bloco quantizado \hat{B}'_k do espetro de escala X' , em que X' é tal como descrito no Passo 2 da Seção 5.3.5.7 do 3gpp TS 26.290. De relembrar que X' possui uma dimensão N , em que $N = 288,576$ e 1152 para TCX-256, 512 e 1024 respetivamente, e que cada bloco B'_k possui a dimensão 8 . O número K dos blocos B'_k é assim 36 , 72 e 144 para TCX-256, 512 e 1024 respetivamente. Os parâmetros algébricos VQ para cada bloco B'_k são descritos no Passo 5 da Seção 5.3.5.7. Para cada bloco B'_k , três conjuntos de índices binários são enviados pelo codificador:

a) o índice do livro de códigos n_k , transmitido em código unário tal como descrito no Passo 5 da Seção 5.3.5.7;

b) a classificação l_k de um ponto de rede c em um chamado livro de códigos base, que indica qual a permutação tem de ser aplicada a um *leader* específico (ver Passo 5 da Seção 5.3.5.7) para obter um ponto de rede c ;

c) e, se o bloco quantizado \hat{B}'_k (um ponto de estrutura) não estava no livro de códigos base, os 8 índices do índice de extensão de Voronoi \mathbf{k} são calculados no sub-passo V1 do Passo 5 na Seção; a partir dos índices da extensão de Voronoi, um vetor de extensão \mathbf{z} pode ser calculado tal como na referência [1] do 3gpp TS 26.290. O número de bits em cada componente do vetor de índice \mathbf{k} é dado pela ordem de extensão r , que pode ser obtida a partir do valor do código unário do índice n_k . O fator de escala M da extensão de Voronoi é dado por $M = 2^r$.

[032] Depois, a partir do fator de escala M , o vetor da extensão de Voronoi \mathbf{z} (um ponto de rede em RE_8) e o ponto de rede \mathbf{c} no livro de códigos base (também um ponto de rede em RE_8), cada bloco à escala quantizado \hat{B}'_k pode ser calculado como

$$\hat{B}'_k = M \mathbf{c} + \mathbf{z}$$

[033] Quando não existe extensão de Voronoi (ou seja, $n_k < 5$, $M=1$ e $\mathbf{z}=0$), o livro de códigos base é o livro de códigos Q_0 , Q_2 , Q_3 ou Q_4 da referência [1] de 3gpp TS 26.290. Nenhuns bits são então necessários para transmitir o vetor \mathbf{k} . De outro modo, quando a extensão de Voronoi é utilizada, porque \hat{B}'_k é suficientemente grande, então apenas Q_3 ou Q_4 da referência [1] é utilizado como um livro de códigos base. A seleção de Q_3 ou Q_4 é implícita no valor do índice do livro de códigos n_k , tal como descrito no Passo 5 da Seção 5.3.5.7.

[034] Estimativa do valor de *pitch* dominante

[035] A estimativa de *pitch* dominante é efetuada de modo que o próximo quadro a ser descodificado possa ser adequadamente extrapolado se corresponder ao TCX-256 e se o pacote associado tiver sido perdido. Esta estimativa é baseada na suposição de que o pico da magnitude máxima no espetro do TCX alvo corresponde ao *pitch* dominante. A pesquisa para o M máximo é restrinida a uma frequência abaixo de $F_s/64$ kHz

$$M = \max_{i=1..N/32} (X'_{2i})^2 + (X'_{2i+1})^2$$

[036] e o índice mínimo $1 \leq i_{max} \leq N/32$ tal como $(X'_{2i})^2 + (X'_{2i+1})^2 = M$ é também considerado. Depois o *pitch* dominante é estimado no número de amostras como $T_{est} = N / i_{max}$ (este valor pode não ser inteiro). De relembrar que o *pitch* dominante é calculado para dissimulação por eliminação de pacotes em TCX-256. Para evitar problemas de *buffering* (o buffer de excitação é limitado a 256 amostras), se $T_{est} > 256$ amostras, *pitch_tcx* é definido como 256; caso contrário, se $T_{est} \leq 256$, um período múltiplo de *pitch* em 256 amostras são evitadas através da definição de *pitch_tcx* para

$$pitch_{tcx} = \max \{ \lfloor n T_{est} \rfloor \mid n \text{ integer} > 0 \text{ and } n T_{est} \leq 256 \}$$

[037] em que [...] indica o arredondamento para o número inteiro mais próximo na direção de $-\infty$.

[038] A seguir, alguns conceitos convencionais adicionais serão brevemente discutidos.

[039] Em ISO_IEC_DIS_23003_3 (referência [3]), uma decodificação TCX que aplique MDCT é explicada no contexto da Fala Unificada e Codec de Áudio.

[040] No estado da arte AAC (ver, por exemplo, referência [4]), apenas um modo de interpolação é descrito. De acordo com a referência [4], o núcleo decodificador AAC inclui uma função de dissimulação que aumenta o atraso do decodificador em um quadro.

[041] Na patente europeia EP 1207519 B1 (referência [5]), é descrito fornecer um método decodificador da fala e de compensação de erros capaz de obter melhoria adicional para fala decodificada em um quadro no qual um erro é detectado. De acordo com a patente, um parâmetro de codificação da fala inclui informação de modo que expressa características de cada segmento curto (quadro) da fala. O codificador da fala calcula de modo adaptável parâmetros de desfasamento e parâmetros de ganho utilizados para decodificação da fala de acordo com a informação de modo. Além disso, o decodificador da fala controla de modo adaptável a taxa de ganho de excitação adaptável e ganho fixo do ganho de excitação de acordo com a informação de modo. Além disso, o conceito de acordo com a patente compreende o controlo de modo adaptável dos parâmetros de ganho de excitação adaptável e parâmetros de ganho de excitação fixa utilizados para decodificação da fala de acordo com valores dos parâmetros de ganho descodificados em uma unidade de decodificação normal na qual nenhum erro é detectado imediatamente após uma unidade de decodificação cujos dados codificados são detectados para conter um erro.

[042] Do ponto de vista da técnica anterior, existe uma necessidade de uma melhoria adicional da dissimulação de erros, que fornece uma melhor

impressão de audição.

[043] 3. Resumo da Invenção

[044] Um modelo de acordo com a invenção cria um decodificador áudio para fornecer uma informação de áudio decodificada baseada em uma informação de áudio codificada. O decodificador áudio compreende uma dissimulação de erros configurada para fornecer uma informação de áudio de dissimulação de erro para dissimular uma perda de um quadro de áudio (ou mais do que uma perda de quadros) após um quadro de áudio codificado em uma representação de domínio de frequência, utilizando um sinal de excitação de domínio de tempo.

[045] Este modelo de acordo com a invenção é baseado na verificação de que uma dissimulação de erros melhorada pode ser obtida fornecendo a informação de áudio de dissimulação de erro baseada em um sinal de excitação de domínio de tempo mesmo se o quadro de áudio que antecede um quadro de áudio perdido for codificado em uma representação de domínio de frequência. Por outras palavras, foi reconhecido que uma qualidade de uma dissimulação de erros é habitualmente melhor se a dissimulação de erros for executada baseada em um sinal de excitação de domínio de tempo, quando comparado a uma dissimulação de erros executada em um domínio de frequências, de modo que vale a pena mudar para a dissimulação de domínio de tempo, utilizando um sinal de excitação de domínio de tempo, mesmo se o conteúdo áudio que antecede o quadro de áudio perdido for codificado de domínio de frequência (isto é, em uma representação de domínio de frequência). Isto é, por exemplo, verdadeiro para um sinal monofônico e maioritariamente para fala.

[046] Por conseguinte, esta invenção permite obter uma boa dissimulação de erros mesmo se o quadro de áudio que antecede o quadro de áudio perdido for codificado de domínio de frequência (isto é, em uma representação de domínio de frequência).

[047] Em um modelo preferido, a representação de domínio de frequência

compreende uma representação codificada de uma pluralidade de valores espectrais e uma representação codificada de uma pluralidade de fatores de escala para o escalonamento dos valores espectrais, ou o decodificador de áudio é configurado para derivar uma pluralidade de fatores de escala para o escalonamento dos valores espectrais da representação codificada dos parâmetros LPC. Isso poderia ser efetuado através da utilização de FDNS (Modulação do Ruído de domínio de Frequências). Contudo, foi verificado que vale a pena derivar um sinal de excitação de domínio de tempo (que pode servir como uma excitação para uma síntese LPC) mesmo se o quadro de áudio que antecede o quadro de áudio perdido for originalmente codificado na representação de domínio de frequência compreendendo informação significativamente diferente (em especial, uma representação codificada de uma pluralidade de valores espectrais em uma representação codificada de uma pluralidade de fatores de escala para o escalonamento dos valores espectrais). Por exemplo, no caso de TCX não enviamos fatores de escala (de um codificador para um decodificador) mas LPC e depois no decodificador transformamos o LPC em uma representação do fator de escala para os bins MDCT. Por outras palavras, no caso de TCX enviamos o coeficiente LPC e depois no decodificador transformamos esses coeficientes LPC em uma representação do fator de escala para TCX em USAC ou no AMR-WB+ não existe de todo fator de escala.

[048] Em um modelo preferido, o decodificador de áudio compreende um núcleo decodificador de domínio de frequência configurado para aplicar um escalonamento baseado no fator de escala a uma pluralidade de valores espectrais derivados da representação de domínio de frequência. Neste caso, a dissimulação de erros é configurada para fornecer a informação de áudio de dissimulação de erro para dissimulação de uma perda de um quadro de áudio após um quadro de áudio codificado na representação de domínio de frequência compreendendo uma pluralidade de fatores de escala codificados utilizando um sinal de excitação de domínio de tempo derivado da

representação de domínio de frequência. Este modelo de acordo com a invenção é baseado na verificação de que a derivação do sinal de excitação de domínio de tempo proveniente da representação de domínio de frequência mencionada em cima fornece habitualmente um melhor resultado de dissimulação de erros quando comparado com uma dissimulação de erros executada diretamente de domínio de frequência. Por exemplo, o sinal de excitação é criado baseado na síntese do quadro anterior, então não importa se o quadro anterior é um domínio de frequência (MDCT, FFT...) ou um quadro de domínio de tempo. Contudo, vantagens especiais podem ser observadas se o quadro anterior foi um domínio de frequência. Além disso, dever-se-á ter em atenção que resultados especialmente bons são obtidos, por exemplo, para sinal monofônico como a fala. Como outro exemplo, os fatores de escala podem ser transmitidos como coeficientes LPC, por exemplo utilizando uma representação polinomial que é então convertida para fatores de escala no lado decodificador.

[049] Em um modelo preferido, o decodificador áudio comprehende um núcleo decodificador de domínio de frequência configurado para derivar uma representação do sinal de áudio de domínio de tempo da representação de domínio de frequência sem utilizar um sinal de excitação de domínio de tempo como uma quantidade intermédia para o quadro de áudio codificado na representação de domínio de frequência. Por outras palavras, foi verificado que o uso de um sinal de excitação de domínio de tempo para uma dissimulação de erros é vantajoso mesmo se o quadro de áudio que antecede o quadro de áudio perdido for codificado em um modo de frequência “verdadeira” que não usa qualquer sinal de excitação de domínio de tempo como uma quantidade intermédia (e que é consequentemente não baseado em uma síntese LCP).

[050] Em um modelo preferido, a dissimulação de erros é configurada para obter o sinal de excitação de domínio de tempo baseado no quadro de áudio codificado na representação de domínio de frequência para fornecer a informação de áudio de dissimulação de erro para dissimular o quadro de áudio

perdido utilizando o dito sinal de excitação de domínio de tempo. Por outras palavras, foi reconhecido que o sinal de excitação de domínio de tempo, utilizado para a dissimulação de erros, deverá ser derivado do quadro de áudio codificado na representação de domínio de frequência anterior ao quadro de áudio perdido, pois este sinal de excitação de domínio de tempo derivado do quadro de áudio codificado na representação de domínio de frequência anterior ao quadro de áudio perdido fornece uma boa representação de um conteúdo áudio do quadro de áudio que antecede o quadro de áudio perdido, de modo que a dissimulação de erros possa ser executada com esforço moderado e boa precisão.

[051] Em um modelo preferido, a dissimulação de erros é configurada para executar uma análise LPC baseada no quadro de áudio codificado na representação de domínio de frequência que antecede o quadro de áudio perdido, para obter um conjunto de parâmetros de codificação preditiva linear e o sinal de excitação de domínio de tempo que representa um conteúdo áudio do quadro de áudio codificado na representação de domínio de frequência anterior ao quadro de áudio perdido. Foi verificado que vale o esforço executar uma análise LPC, para derivar os parâmetros de codificação preditiva linear e o sinal de excitação de domínio de tempo, mesmo se o quadro de áudio que antecede o quadro de áudio perdido for codificado em uma representação de domínio de frequência (que não contém quaisquer parâmetros de codificação preditiva linear e nenhuma representação do sinal de excitação de domínio de tempo), visto que uma informação de áudio de dissimulação de erro de boa qualidade pode ser obtida para muitos sinais de áudio de entrada, baseados no dito sinal de excitação de domínio de tempo. Em alternativa, a dissimulação de erros pode ser configurada para executar uma análise LPC baseada no quadro de áudio codificado na representação de domínio de frequência anterior ao quadro de áudio perdido, para obter o sinal de excitação de domínio de tempo que representa um conteúdo áudio do quadro de áudio codificado na representação de domínio de frequência anterior ao quadro de áudio perdido.

Ainda em alternativa, o decodificador áudio pode ser configurado para obter um conjunto de parâmetros de codificação preditiva linear utilizando uma estimativa do parâmetro de codificação preditiva linear, ou o decodificador áudio pode ser configurado para obter um conjunto de parâmetros de codificação preditiva linear baseado em um conjunto de fatores de escala utilizando uma transformação. Por outras palavras, os parâmetros LPC podem ser obtidos utilizando a estimativa do parâmetro LPC. Isso pode ser efetuado através de *windowing/autocorr/levinson de durbin* baseado no quadro de áudio codificado na representação de domínio de frequência ou pela transformação do fator de escala anterior diretamente para a representação LPC.

[052] Em um modelo preferido, a dissimulação de erros é configurada para obter uma informação de *pitch* (ou desfasamento) descrevendo um *pitch* do quadro de áudio codificado de domínio de frequência anterior ao quadro de áudio perdido, e para fornecer a informação de áudio de dissimulação de erro dependente da informação de *pitch*. Tendo em consideração a informação de *pitch*, pode ser determinado que a informação de áudio de dissimulação de erro (habitualmente uma informação de áudio de dissimulação de erro que abrange a duração temporal de pelo menos um quadro de áudio perdido) esteja bem adaptada ao conteúdo áudio atual.

[053] Em um modelo preferido, a dissimulação de erros é configurada para obter a informação de *pitch* baseada no sinal de excitação de domínio de tempo derivado do quadro de áudio codificado na representação de domínio de frequência anterior ao quadro de áudio perdido. Foi verificado que uma derivação da informação de *pitch* do sinal de excitação de domínio de tempo traz consigo uma elevada precisão. Além disso, foi verificado ser vantajoso se a informação de *pitch* for bem adaptada ao sinal de excitação de domínio de tempo, visto que a informação de *pitch* é utilizada para uma modificação do sinal de excitação de domínio de tempo. Ao derivar a informação de *pitch* do sinal de excitação de domínio de tempo, uma relação próxima pode ser obtida.

[054] Em um modelo preferido, a dissimulação de erros é configurada para

avaliar uma correlação cruzada do sinal de excitação de domínio de tempo, para determinar uma informação de *pitch* grosseira. Além disso, a dissimulação de erros pode ser configurada para refinar a informação de *pitch* grosseira utilizando uma pesquisa de ciclo fechado em redor de um *pitch* determinado pela informação de *pitch* grosseira. Por conseguinte, uma informação de *pitch* altamente precisa pode ser obtida com moderado esforço computacional.

[055] Em um modelo preferido, a dissimulação de erros do decodificador áudio pode ser configurada para obter uma informação de *pitch* baseada em uma informação paralela da informação de áudio codificada.

[056] Em um modelo preferido, a dissimulação de erros pode ser configurada para obter uma informação de *pitch* baseada em uma informação de *pitch* disponível para um quadro de áudio previamente descodificado.

[057] Em um modelo preferido, a dissimulação de erros é configurada para obter uma informação de *pitch* baseada em uma pesquisa de *pitch* executada em um sinal de domínio de tempo ou em um sinal residual.

[058] Por outras palavras, o *pitch* pode ser transmitido como informação paralela ou pode também ser proveniente do quadro anterior se existir LPC, por exemplo. A informação de *pitch* pode também ser transmitida no fluxo de bits se estiver disponível no codificador. Podemos opcionalmente fazer pesquisa de *pitch* no sinal de domínio de tempo diretamente ou no residual, que dê habitualmente melhores resultados no residual (sinal de excitação de domínio de tempo).

[059] Em um modelo preferido, a dissimulação de erros é configurada para copiar um ciclo de *pitch* do sinal de excitação de domínio de tempo derivado do quadro de áudio codificado na representação de domínio de frequência anterior ao quadro de áudio perdido uma vez ou várias vezes, de modo a obter um sinal de excitação para uma síntese da dissimulação de erros do sinal de áudio. Ao copiar o sinal de excitação de domínio de tempo uma vez ou várias vezes, pode ser determinado que o componente determinístico (isto é, significativamente periódico) da informação de áudio de dissimulação de erro é

obtido com boa exatidão e é uma boa continuação do componente determinístico (por ex.: significativamente periódico) do conteúdo áudio do quadro de áudio que antecede o quadro de áudio perdido.

[060] Em um modelo preferido, a dissimulação de erros é configurada para o filtro de baixa frequência do ciclo de *pitch* do sinal de excitação de domínio de tempo a partir da representação de domínio de frequência do quadro de áudio codificado na representação de domínio de frequência anterior ao quadro de áudio perdido utilizando um filtro dependente da taxa de amostragem, uma largura de banda a qual é dependente de uma taxa de amostragem do quadro de áudio codificado em uma representação de domínio de frequência. Por conseguinte, o sinal de excitação de domínio de tempo pode ser adaptado a uma largura de banda de áudio disponível, que resulta em uma boa impressão de audição da informação de áudio de dissimulação de erro. Por exemplo, é preferível passar por baixo apenas no primeiro quadro perdido, e de preferência, passamos também por baixo apenas se o sinal não for 100% estável. Contudo, dever-se-á ter em atenção que a filtragem passa-baixo é opcional, e pode ser executada apenas no primeiro ciclo de *pitch*. Por exemplo, o filtro pode ser dependente da taxa de amostragem, de modo que a frequência de corte seja dependente da largura de banda.

[061] Em um modelo preferido, a dissimulação de erros é configurada para predizer um *pitch* no final de um quadro perdido para adaptar o sinal de excitação de domínio de tempo, ou uma ou mais cópias deste, para o *pitch* previsto. Por conseguinte, alterações de *pitch* esperadas durante o quadro de áudio perdido podem ser consideradas. Consequentemente, perturbações em uma transição entre a informação de áudio de dissimulação de erro e uma informação de áudio de um quadro descodificado de modo adequado a seguir a um ou mais quadros de áudio perdidos são evitadas (ou pelo menos reduzidas, visto que se trata apenas de um *pitch* de predição e não o verdadeiro). Por exemplo, a adaptação vai desde o último *pitch* bom para o de predição. Isso é feito pela resincronização de impulsos [7].

[062] Em um modelo preferido, a dissimulação de erros é configurada para combinar um sinal de excitação de domínio de tempo extrapolado e um sinal de ruído, de modo a obter um sinal de entrada para uma síntese LPC. Neste caso, a dissimulação de erros é configurada para executar a síntese LPC, em que a síntese LPC é configurada para filtrar o sinal de entrada da síntese LPC dependendo dos parâmetros de codificação de predição linear, de modo a obter a informação de áudio de dissimulação de erro. Por conseguinte, ambos um componente determinístico (por exemplo, aproximadamente periódico) do conteúdo áudio e um componente do tipo ruído do conteúdo áudio podem ser considerados. Por conseguinte, é determinado que a informação de áudio de dissimulação de erro compreende uma impressão de audição “natural”.

[063] Em um modelo preferido, a dissimulação de erros é configurada para calcular um ganho do sinal de excitação de domínio de tempo extrapolado, utilizado para obter o sinal de entrada para a síntese LPC, utilizando uma correlação de domínio de tempo executada baseada em uma representação de domínio de tempo do quadro de áudio codificado de domínio de frequência anterior ao quadro de áudio perdido, em que um desfasamento da correlação é definido dependendo de uma informação de *pitch* obtida baseada no sinal de excitação de domínio de tempo. Por outras palavras, uma intensidade de um componente periódico é determinada no quadro de áudio que antecede o quadro de áudio perdido, e esta intensidade determinada do componente periódico é utilizada para obter a dissimulação da informação de áudio. Contudo, foi verificado que o cálculo anteriormente mencionado da intensidade do componente periódico fornece resultados especialmente bons, visto que o sinal de áudio atual de domínio de tempo do quadro de áudio que antecede o quadro de áudio perdido é considerado. Em alternativa, uma correlação de domínio da excitação ou diretamente de domínio de tempo pode ser utilizada para obter a informação de *pitch*. Contudo, existem também diferentes possibilidades, dependendo de qual modelo é utilizado. Em um modelo, a informação de *pitch* poderia ser apenas o *pitch* obtido do ltp do último quadro

ou o *pitch* transmitido como informação paralela ou o calculado.

[064] Em um modelo preferido, a dissimulação de erros é configurada para filtrar passa por baixo o sinal de ruído combinado com o sinal de excitação de domínio de tempo extrapolado. Foi verificado que a filtragem de baixa frequência do sinal de ruído (tipicamente alimentado na síntese LPC) resulta em uma impressão de audição natural. Por exemplo, a característica de alta frequência pode alterar com a quantidade de quadros perdidos, após uma certa quantidade de perdas de quadros poderá já não existir alta frequência. A característica de alta frequência pode ser também dependente da taxa de amostragem com a qual o decodificador está a executar. Por exemplo, a alta frequência é dependente da taxa de amostragem, e a característica do filtro pode mudar ao longo do tempo (com consecutivas perdas de quadro). A característica de alta frequência pode também ser opcionalmente alterada sobre consecutivas perdas de quadro de modo que depois de uma certa quantidade de perdas de quadro já não existe filtragem para apenas obter a totalidade do ruído em forma de banda para obter um bom ruído aceitável fechado para o ruído de fundo.

[065] Em um modelo preferido, a dissimulação de erros é configurada para seletivamente alterar a forma espectral do sinal de ruído (562) utilizando o filtro de pré-ênfase no qual o sinal de ruído é combinado com o sinal de excitação de domínio de tempo extrapolado se o quadro de áudio codificado em uma representação de domínio de frequência anterior ao quadro de áudio for um quadro de áudio vozeado ou compreende um início. Tem sido verificado que a impressão de audição da informação de áudio de dissimulação de erro pode ser melhorada por esse conceito. Por exemplo, nalguns casos é melhor diminuir os ganhos e moldá-los e nalguns locais é melhor aumentá-los.

[066] Em um modelo preferido, a dissimulação de erros é configurada para calcular um ganho do sinal de ruído dependendo de uma correlação de domínio de tempo, executada baseada em uma representação do domínio de tempo do quadro de áudio codificado na representação de domínio de

frequência anterior ao quadro de áudio perdido. Tem sido verificado que essa determinação do ganho do sinal de ruído fornece resultados especialmente exatos, visto que o sinal atual de domínio de tempo associado ao quadro de áudio que antecede o quadro de áudio perdido pode ser considerado. Utilizando este conceito, é possível ser capaz de obter uma energia do quadro dissimulado próxima da energia do quadro de áudio anterior. Por exemplo, o ganho para o sinal de ruído pode ser gerado medindo a energia do resultado: excitação do sinal de entrada – excitação baseada no *pitch* gerado.

[067] Em um modelo preferido, a dissimulação de erros é configurada para modificar um sinal de excitação de domínio de tempo obtido baseado em um ou mais quadros de áudio que antecedem um quadro de áudio perdido, de modo a obter a informação de áudio de dissimulação de erro. Foi verificado que a modificação do sinal de excitação de domínio de tempo permite adaptar o sinal de excitação de domínio de tempo a uma evolução temporal desejada. Por exemplo, a modificação do sinal de excitação de domínio de tempo permite “desvanecer” o componente determinístico (por exemplo, significativamente periódico) do conteúdo áudio na informação de áudio de dissimulação de erro. Além disso, a modificação do sinal de excitação de domínio de tempo permite também adaptar o sinal de excitação de domínio de tempo a uma variação do (estimada ou esperada) de *pitch*. Isto permite ajustar as características da informação de áudio de dissimulação de erro ao longo do tempo.

[068] Em um modelo preferido, a dissimulação de erros é configurada para utilizar uma ou mais cópias modificadas do sinal de excitação de domínio de tempo obtido baseado em um ou mais quadros de áudio que antecedem um quadro de áudio perdido, de modo a obter a informação de dissimulação de erros. As cópias modificadas do sinal de excitação de domínio de tempo podem ser obtidas com um esforço moderado, e a modificação pode ser executada utilizando um simples algoritmo. Desse modo, as características desejadas da informação de áudio de dissimulação de erro podem ser obtidas com esforço moderado.

[069] Em um modelo preferido, a dissimulação de erros é configurada para modificar o sinal de excitação de domínio de tempo obtido baseado em um ou mais quadros de áudio que antecedem um quadro de áudio perdido, ou uma ou mais cópias deste, para assim reduzir um componente periódico da informação de áudio de dissimulação de erro. Por conseguinte, pode ser considerado que a correlação entre o conteúdo áudio do quadro de áudio que antecede o quadro de áudio perdido e o conteúdo de áudio de um ou mais quadros de áudio perdidos diminua ao longo do tempo. Também, pode ser evitado que uma impressão auditiva não natural seja originada por uma longa conservação de um componente periódico da informação de áudio de dissimulação de erro.

[070] Em um modelo preferido, a dissimulação de erros é configurada à escala ao sinal de excitação de domínio de tempo obtido baseado em um ou mais quadros de áudio que antecedem o quadro de áudio perdido, ou uma ou mais cópias deste, para desse modo modificar o sinal de excitação de domínio de tempo. Foi verificado que a operação de escalonamento pode ser executada com pouco esforço, em que o sinal de excitação de domínio de tempo escalonado tipicamente fornece uma boa informação de áudio de dissimulação de erro.

[071] Em um modelo preferido, a dissimulação de erros é configurada para reduzir gradualmente um ganho aplicado à escala ao sinal de excitação de domínio de tempo obtido baseado em um ou mais quadros de áudio que antecedem um quadro de áudio perdido, ou a uma ou mais cópias deste. Por conseguinte, um desvanecimento do componente periódico pode ser obtido na informação de áudio de dissimulação de erro.

[072] Em um modelo preferido, a dissimulação de erros é configurada para ajustar uma velocidade utilizada para gradualmente reduzir um ganho aplicado à escala ao sinal de excitação de domínio de tempo obtido baseado em um ou mais quadros de áudio que antecedem um quadro de áudio perdido, ou a uma ou mais cópias deste, dependendo de um ou mais parâmetros de um ou mais quadros de áudio que antecedem o quadro de áudio perdido, e/ou dependendo

de um número de quadros de áudio consecutivamente perdidos. Por conseguinte, é possível regular a velocidade na qual o componente determinístico (por exemplo, pelo menos aproximadamente periódico) é desvanecido na informação de áudio de dissimulação de erro. A velocidade do desvanecimento pode ser adaptada a características específicas do conteúdo áudio, que pode ser tipicamente observado de um ou mais parâmetros de um ou mais quadros de áudio que antecedem o quadro de áudio perdido. Em alternativa, ou além disso, o número de quadros de áudio perdidos consecutivos pode ser considerado aquando da determinação da velocidade utilizada para desvanecer o componente determinístico (por exemplo, pelo menos aproximadamente periódico) da informação de áudio de dissimulação de erro, o que ajuda a adaptar a dissimulação de erros à situação específica. Por exemplo, o ganho da parte tonal e o ganho da parte ruidosa pode ser desvanecido em separado. O ganho para a parte tonal pode convergir para zero após uma certa quantidade de perdas de quadros enquanto o ganho do ruído pode convergir para o ganho determinado para atingir um certo ruído aceitável.

[073] Em um modelo preferido, a dissimulação de erros é configurada para regular a velocidade utilizada para gradualmente reduzir um ganho aplicado à escala ao sinal de excitação de domínio de tempo obtido baseado em um ou mais quadros de áudio que antecedem um quadro de áudio perdido, ou a uma ou mais cópias deste, dependendo de um comprimento de um período de *pitch* do sinal de excitação de domínio de tempo, de modo que uma entrada do sinal de excitação de domínio de tempo em uma síntese LPC seja desvanecida mais rapidamente para sinais com um comprimento mais curto do período de *pitch* quando comparado com os sinais dotados de um comprimento maior do período de *pitch*. Por conseguinte, pode ser evitado que os sinais com um comprimento mais curto do período de *pitch* sejam repetidos com muita frequência com elevada intensidade, pois isto iria tipicamente resultar em uma impressão de audição não natural. Desse modo, uma qualidade global da

informação de áudio de dissimulação de erro pode ser melhorada.

[074] Em um modelo preferido, a dissimulação de erros é configurada para regular a velocidade utilizada para gradualmente reduzir um ganho aplicado à escala ao sinal de excitação de domínio de tempo obtido baseado em um ou mais quadros de áudio que antecedem um quadro de áudio perdido, ou a uma ou mais cópias deste, dependendo de um resultado de uma análise de *pitch* ou uma predição de *pitch*, de modo que um componente determinístico da entrada do sinal de excitação de domínio de tempo em uma síntese LPC é desvanecida mais rapidamente para sinais com uma alteração de *pitch* maior por unidade de tempo quando comparado aos sinais com uma alteração de *pitch* mais pequena por unidade de tempo, e/ou de modo que um componente determinístico da entrada do sinal de excitação de domínio de tempo em uma síntese LPC é desvanecida mais rapidamente para sinais para os quais uma predição de *pitch* falha quando comparado com os sinais para os quais a predição de *pitch* é bem-sucedida. Por conseguinte, o desvanecimento pode ser feito mais rapidamente para os sinais nos quais existe uma grande incerteza de *pitch* quando comparado com sinais para os quais existe uma incerteza mais pequena de *pitch*. Contudo, ao desvanecer um componente determinístico mais rapidamente para os sinais que compreendem uma incerteza comparativamente grande de *pitch*, perturbações audíveis podem ser evitadas ou pelo menos significativamente reduzidas.

[075] Em um modelo preferido, a dissimulação de erros é configurada para escalonamento de tempo o sinal de excitação de domínio de tempo obtido baseado em um ou mais quadros de áudio que antecedem um quadro de áudio perdido, ou a uma ou mais cópias deste, dependendo de uma predição de um *pitch* para o tempo de um ou mais quadros de áudio perdidos. Por conseguinte, o sinal de excitação de domínio de tempo pode ser adaptado a um *pitch* variável, de modo que a informação de áudio de dissimulação de erro compreenda uma impressão de audição mais natural.

[076] Em um modelo preferido, a dissimulação de erros é configurada para

fornecer a informação de áudio de dissimulação de erro durante um tempo mais longo do que uma duração temporal de um ou mais quadros de áudio perdidos. Por conseguinte, é possível executar uma operação de sobreposição e adição baseada na informação de áudio de dissimulação de erro, que ajuda a reduzir perturbações de bloqueio.

[077] Em um modelo preferido, a dissimulação de erros é configurada para executar uma sobreposição e adição da informação de áudio de dissimulação de erro e uma representação de domínio de tempo de um ou mais quadros de áudio recebidos de modo adequado a seguir a um ou mais quadros de áudio perdidos. Desse modo, é possível evitar (ou pelo menos reduzir) perturbações de bloqueio.

[078] Em um modelo preferido, a dissimulação de erros é configurada para derivar a informação de áudio de dissimulação de erro baseada em pelo menos três quadros parcialmente sobrepostos ou janelas que antecedem um quadro de áudio perdido ou uma janela perdida. Por conseguinte, a informação de áudio de dissimulação de erro pode ser obtida com boa exatidão mesmo para modos de codificação nos quais mais de dois quadros (ou janelas) se encontram sobrepostos (sendo que essa sobreposição pode ajudar a reduzir o atraso).

[079] Outro modelo de acordo com a invenção cria um método para fornecer uma informação de áudio decodificada baseada em uma informação de áudio codificada. O método compreende o fornecimento de uma informação de áudio de dissimulação de erro para a dissimulação de uma perda de um quadro de áudio a seguir a um quadro de áudio codificado em uma representação de domínio de frequência utilizando um sinal de excitação de domínio de tempo. Este método está baseado nas mesmas considerações que o decodificador áudio anteriormente mencionado.

[080] Ainda outro modelo de acordo com a invenção cria um programa de computador para executar o dito método quando o programa de computador é executado em um computador.

[081] Outro modelo de acordo com a invenção cria um decodificador áudio para fornecer uma informação de áudio decodificada baseada em uma informação de áudio codificada. O decodificador áudio compreende uma dissimulação de erros configurada para fornecer uma informação de áudio de dissimulação de erro para dissimular uma perda de um quadro de áudio. A dissimulação de erros é configurada para modificar um sinal de excitação de domínio de tempo baseado em um ou mais quadros de áudio que antecedem um quadro de áudio perdido, de modo a obter a informação de áudio de dissimulação de erro.

[082] Este modelo de acordo com a invenção está baseado na ideia de que uma dissimulação de erros com uma boa qualidade de áudio pode ser obtida baseada em um sinal de excitação de domínio de tempo, em que uma modificação do sinal de excitação de domínio de tempo obtido baseado em um ou mais quadros de áudio que antecedem um quadro de áudio perdido permite uma adaptação da informação de áudio de dissimulação de erro às alterações esperadas (ou previstas) do conteúdo áudio durante o quadro perdido. Por conseguinte, perturbações e, em especial, uma impressão de audição não natural, podem ser evitadas. Consequentemente, um fornecimento melhorado de uma informação de áudio de dissimulação de erro é obtido, de modo que quadros de áudio perdidos possam ser dissimulados com resultados melhorados.

[083] Em um modelo preferido, a dissimulação de erros é configurada para utilizar uma ou mais cópias modificadas do sinal de excitação de domínio de tempo obtido para um ou mais quadros de áudio que antecedem um quadro de áudio perdido, de modo a obter a informação da dissimulação de erros. Utilizando uma ou mais cópias modificadas do sinal de excitação de domínio de tempo obtido para um ou mais quadros de áudio que antecedem um quadro de áudio perdido, uma boa qualidade da informação de áudio de dissimulação de erro pode ser obtida com pouco esforço computacional.

[084] Em um modelo preferido, a dissimulação de erros é configurada para

modificar o sinal de excitação de domínio de tempo obtido para um ou mais quadros de áudio que antecedem um quadro de áudio perdido, ou uma ou mais cópias deste, para desse modo reduzir um componente periódico da informação de áudio de dissimulação de erro ao longo do tempo. Ao reduzir o componente periódico da informação de áudio de dissimulação de erro ao longo do tempo, uma conservação longa não natural de um som determinístico (por exemplo, aproximadamente periódico) pode ser evitada, ajudando a fazer com que a informação de áudio de dissimulação de erro pareça natural.

[085] Em um modelo preferido, a dissimulação de erros é configurada para o escalonamento do sinal de excitação de domínio de tempo obtido baseado em um ou mais quadros de áudio que antecedem o quadro de áudio perdido, ou uma ou mais cópias deste, para desse modo modificar o sinal de excitação de domínio de tempo. O escalonamento do sinal de excitação de domínio de tempo constitui uma maneira especialmente eficiente de variar a informação de áudio de dissimulação de erro ao longo do tempo.

[086] Em um modelo preferido, a dissimulação de erros é configurada para reduzir gradualmente um ganho aplicado à escala ao sinal de excitação de domínio de tempo obtido para um ou mais quadros de áudio que antecedem um quadro de áudio perdido, ou a uma ou mais cópias deste. Foi verificado que reduzindo gradualmente o ganho aplicado à escala ao sinal de excitação de domínio de tempo obtido para um ou mais quadros de áudio que antecedem um quadro de áudio perdido, ou a uma ou mais cópias deste, permite obter um sinal de excitação de domínio de tempo para o fornecimento da informação de áudio de dissimulação de erro, de modo que os componentes determinísticos (por exemplo, pelo menos componentes aproximadamente periódicos) são desvanecidos. Por exemplo, poderá não existir apenas um ganho. Por exemplo, poderemos ter um ganho para a parte tonal (também referida como parte aproximadamente periódica), e um ganho para a parte do ruído. Ambas excitações (ou componentes de excitação) podem ser desvanecidas em separado com fator de velocidade diferente e depois as duas excitações

resultantes (ou componentes de excitação) podem ser combinadas antes de serem alimentadas ao LPC para síntese. No caso em que não temos qualquer estimativa de ruído de fundo, o fator de desvanecimento para o ruído e para a parte tonal pode ser idêntico, e depois podemos ter apenas um desvanecimento a aplicar aos resultados das duas excitações multiplicadas por o seu próprio ganho e combinados.

[087] Desse modo, pode ser evitado que a informação de áudio de dissimulação de erro compreenda um componente de áudio determinístico temporariamente estendido (por exemplo, pelo menos aproximadamente periódico), que poderia tipicamente fornecer uma impressão de audição não natural.

[088] Em um modelo preferido, a dissimulação de erros é configurada para regular uma velocidade utilizada para reduzir gradualmente um ganho aplicado à escala ao sinal de excitação de domínio de tempo para um ou mais quadros de áudio que antecedem um quadro de áudio perdido, ou a uma ou mais cópias deste, dependendo de um ou mais parâmetros de um ou mais quadros de áudio que antecedem o quadro de áudio perdido, e/ou dependendo de um número de quadros de áudio perdidos consecutivos. Desse modo, a velocidade de desvanecimento do componente determinístico (por exemplo, pelo menos aproximadamente periódico) na informação de áudio de dissimulação de erro pode ser adaptada à situação específica com esforço computacional moderado. Visto que o sinal de excitação de domínio de tempo utilizado para o fornecimento da informação de áudio de dissimulação de erro é tipicamente uma versão à escala (à escala utilizando o ganho mencionado em cima) do sinal de excitação de domínio de tempo obtido para o um ou mais quadros de áudio que antecedem o quadro de áudio perdido, uma variação do dito ganho (utilizada para derivar o sinal de excitação de domínio de tempo para o fornecimento da informação de áudio de dissimulação de erro) constitui um método simples e no entanto eficaz para adaptar a informação de áudio de dissimulação de erro às necessidades específicas. Contudo, a velocidade de

desvanecimento é também controlável com muito pouco esforço.

[089] Em um modelo preferido, a dissimulação de erros é configurada para regular a velocidade utilizada para reduzir gradualmente um ganho aplicado à escala ao sinal de excitação de domínio de tempo obtido baseado em um ou mais quadros de áudio que antecedem um quadro de áudio perdido, ou a uma ou mais cópias deste, dependendo de um comprimento de um período de *pitch* do sinal de excitação de domínio de tempo, de modo que uma entrada do sinal de excitação de domínio de tempo em uma síntese LPC seja desvanecida mais rapidamente para os sinais com um comprimento mais curto co que o período de *pitch* quando comparado com os sinais com um comprimento maior do período de *pitch*. Por conseguinte, o desvanecimento é executado mais rapidamente para os sinais com um comprimento mais curto do período de *pitch*, o que evita que o período de *pitch* seja copiado demasiadas vezes (o que iria tipicamente resultar em uma impressão de audição não natural).

[090] Em um modelo preferido, a dissimulação de erros é configurada para regular a velocidade utilizada para reduzir gradualmente um ganho aplicado à escala ao sinal de excitação de domínio de tempo obtido para um ou mais quadros de áudio que antecedem um quadro de áudio perdido, ou a uma ou mais cópias deste, dependendo de um resultado de uma análise de *pitch* ou uma predição de *pitch*, de modo que um componente determinístico de uma entrada do sinal de excitação de domínio de tempo em uma síntese LPC seja mais rapidamente desvanecido para os sinais com uma alteração de *pitch* maior por unidade de tempo quando comparado aos sinais com uma alteração de *pitch* mais pequena por unidade de tempo, e/ou de modo que um componente determinístico da entrada do sinal de excitação de domínio de tempo em uma síntese LPC é desvanecida mais rapidamente para sinais para os quais uma predição de *pitch* falha quando comparado com os sinais para os quais a predição de *pitch* é bem-sucedida. Por conseguinte, um componente determinístico (por exemplo, pelo menos aproximadamente periódico) é mais rapidamente desvanecido para os sinais para os quais existe uma maior

incerteza de *pitch* (sendo que uma maior alteração de *pitch* por unidade de tempo, ou mesmo uma falha da predição de *pitch*, indica uma incerteza comparativamente grande do *pitch*). Desse modo, perturbações, que poderiam surgir de um fornecimento de uma informação de áudio de dissimulação de erro altamente determinístico em uma situação na qual o *pitch* atual é incerto, podem ser evitadas.

[091] Em um modelo preferido, a dissimulação de erros é configurada para escalonamento de tempo o sinal de excitação de domínio de tempo obtido para (ou baseado em) um ou mais quadros de áudio que antecedem um quadro de áudio perdido, ou a uma ou mais cópias deste, dependendo de uma predição de um *pitch* para o tempo do um ou mais quadros de áudio perdidos. Por conseguinte, o sinal de excitação de domínio de tempo, utilizado para o fornecimento da informação de áudio de dissimulação de erro, é modificado (quando comparado ao sinal de excitação de domínio de tempo obtido para ou baseado em) um ou mais quadros de áudio que antecedem um quadro de áudio perdido, de modo que o *pitch* do sinal de excitação de domínio de tempo siga os requisitos de um período de tempo do quadro de áudio perdido. Por consequência, uma impressão de audição, que pode ser alcançada pela informação de áudio de dissimulação de erro, pode ser melhorada.

[092] Em um modelo preferido, a dissimulação de erros é configurada para obter um sinal de excitação de domínio de tempo, utilizado para descodificar um ou mais quadros de áudio que antecedem o quadro de áudio perdido, e para modificar o dito sinal de excitação de domínio de tempo, utilizado para descodificar um ou mais quadros de áudio que antecedem o quadro de áudio perdido, para obter um sinal de excitação de domínio de tempo modificado. Neste caso, a dissimulação de domínio de tempo é configurada para fornecer a informação de áudio de dissimulação de erro baseada no sinal de excitação de domínio de tempo modificado. Por conseguinte, é possível reutilizar um sinal de excitação de domínio de tempo, que já foi utilizado para descodificar um ou mais quadros de áudio que antecedem o quadro de áudio perdido. Desse

modo, um esforço computacional pode ser mantido muito pequeno, se o sinal de excitação de domínio de tempo já tiver sido adquirido para a decodificação de um ou mais quadros de áudio que antecedem o quadro de áudio perdido.

[093] Em um modelo preferido, a dissimulação de erros é configurada para obter uma informação de *pitch*, utilizada para descodificar um ou mais quadros de áudio que antecedem o quadro de áudio perdido. Neste caso, a dissimulação de erros é também configurada para fornecer a informação de áudio de dissimulação de erro dependendo da dita informação de *pitch*. Por conseguinte, a informação de *pitch* anteriormente utilizada pode ser reutilizada, evitando um esforço computacional para uma nova computação da informação de áudio. Desse modo, a dissimulação de erros é especialmente computacionalmente eficiente. Por exemplo, no caso da ACELP temos 4 desfasamentos de *pitch* e ganhos por quadro. Podemos utilizar os últimos dois quadros para podermos predizer o *pitch* no final do quadro que temos que dissimular.

[094] Depois, comparando com o codec de domínio de frequência anteriormente descrito onde apenas um ou dois *pitch* por quadro são derivados (podemos ter mais do que dois mas isso iria adicionar muita complexidade para não muito ganho em qualidade), no caso de um codec de comutação que é por exemplo, perda – ACELP – FD então, temos uma precisão muito melhor visto que os *pitch* são transmitidos no fluxo de bits e são baseados no sinal de entrada original (não no descodificado tal como feito no decodificador). No caso de um débito binário elevado, por exemplo, podemos também enviar um desfasamento de *pitch* e informação de ganho, ou informação LTP, por quadro codificado de domínio de frequência.

[095] Em um modelo preferido, o decodificador áudio a dissimulação de erros pode ser configurado para obter uma informação de *pitch* baseada em uma informação paralela da informação de áudio codificada.

[096] Em um modelo preferido, a dissimulação de erros pode ser configurada para obter uma informação de *pitch* baseada em uma informação

de *pitch* disponível para um quadro de áudio previamente codificado.

[097] Em um modelo preferido, a dissimulação de erros é configurada para obter uma informação de *pitch* baseada em uma pesquisa de *pitch* executada em um sinal de domínio de tempo ou em um sinal residual.

[098] Por outras palavras, o *pitch* pode ser transmitido como informação paralela ou poderia também provir do quadro anterior se existir um LTP por exemplo. A informação de *pitch* poderia ser também transmitida no fluxo de bits se estiver disponível no codificador. Podemos como opção fazer a pesquisa de *pitch* no sinal de excitação de domínio de tempo diretamente ou no residual, que tenha habitualmente melhores resultados no residual (sinal de excitação de domínio de tempo).

[099] Em um modelo preferido, a dissimulação de erros é configurada para obter um conjunto de coeficientes de predição linear, que tem sido utilizado para descodificar um ou mais quadros de áudio que antecedem o quadro de áudio perdido. Neste caso, a dissimulação de erros é configurada para fornecer a informação de áudio de dissimulação de erro dependendo do dito conjunto de coeficientes de predição linear. Desse modo, a eficiência da dissimulação de erros é aumentada pela reutilização da informação anteriormente gerada (ou previamente decodificada), como por exemplo o conjunto de coeficientes de predição linear anteriormente utilizado.

[0100] Em um modelo preferido, a dissimulação de erros é configurada para extrapolar um novo conjunto de coeficientes de predição linear baseados no conjunto de coeficientes de predição linear, utilizados para descodificar um ou mais quadros de áudio que antecedem o quadro de áudio perdido. Neste caso, a dissimulação de erros é configurada para utilizar um novo conjunto de coeficientes de predição linear para fornecer a informação da dissimulação de erros. Ao derivar o novo conjunto de coeficientes de predição linear, utilizados para fornecer a informação de áudio de dissimulação de erro, a partir de um conjunto de coeficientes de predição linear anteriormente utilizados utilizando uma extração, um recálculo total dos coeficientes de predição linear pode

ser evitado, ajudando a manter o esforço computacional razoavelmente pequeno. Além disso, ao efetuar uma extrapolação baseada no conjunto de coeficientes de predição linear anteriormente utilizados, pode ser garantido que o novo conjunto de coeficientes de predição linear é pelo menos idêntico ao conjunto de coeficientes de predição linear anteriormente utilizados, ajudando a evitar descontinuidades aquando do fornecimento da informação de dissimulação de erros. Por exemplo, após uma certa quantidade de perda de quadros tendemos a uma estimativa de ruído de fundo em forma LPC. A velocidade desta convergência pode, por exemplo, depender da característica do sinal.

[0101] Em um modelo preferido, a dissimulação de erros é configurada para obter uma informação sobre uma intensidade de um componente do sinal determinístico em um ou mais quadros de áudio que antecedem um quadro de áudio perdido. Neste caso, a dissimulação de erros é configurada para comparar a informação sobre uma intensidade de um componente do sinal determinístico em um ou mais quadros de áudio que antecedem um quadro de áudio perdido com um valor limite, para decidir se deverá introduzir um componente determinístico de um sinal de excitação de domínio de tempo em uma síntese LPC (síntese baseada em coeficiente preditivo linear), ou se deverá introduzir apenas um componente de ruído de um sinal de excitação de domínio de tempo em uma síntese LPC. Por conseguinte, é possível omitir o fornecimento de um componente determinístico (por exemplo, pelo menos aproximadamente periódico) da dissimulação da informação de áudio no caso de existir apenas um pequeno contributo do sinal determinístico no interior de um ou mais quadros que antecedem o quadro de áudio perdido. Foi verificado que isto ajuda a obter uma boa impressão de audição.

[0102] Em um modelo preferido, a dissimulação de erros é configurada para obter uma informação de *pitch* descrevendo um *pitch* do quadro de áudio que antecede o quadro de áudio perdido, e para fornecer a informação de áudio de dissimulação de erro dependendo com a informação de *pitch*. Por conseguinte,

é possível adaptar o *pitch* da informação da dissimulação de erros ao *pitch* do quadro de áudio que antecede o quadro de áudio perdido. Por conseguinte, descontinuidades são evitadas e uma impressão de audição natural pode ser obtida.

[0103] Em um modelo preferido, a dissimulação de erros é configurada para obter uma informação de *pitch* baseada no sinal de excitação de domínio de tempo associado ao quadro de áudio que antecede o quadro de áudio perdido. Foi verificado que a informação de *pitch* obtida baseada no sinal de excitação de domínio de tempo é especialmente fiável, e é também muito bem adaptada ao processamento do sinal de excitação de domínio de tempo.

[0104] Em um modelo preferido, a dissimulação de erros é configurada para avaliar uma correlação cruzada do sinal de excitação de domínio de tempo (ou, em alternativa, de um sinal de áudio de domínio de tempo), para determinar uma informação de *pitch* grosseira, e para refinar a informação de *pitch* grosseira utilizando uma pesquisa de ciclo fechado em redor de um *pitch* determinado (ou descrito) pela informação de *pitch* grosseira. Foi verificado que este conceito permite obter uma informação de *pitch* muito precisa diretamente no sinal de domínio de tempo enquanto em algumas outras efetuamos a pesquisa de *pitch* no sinal de excitação de domínio de tempo.

[0105] Em um modelo preferido, a dissimulação de erros é configurada para obter a informação de *pitch* para o fornecimento da informação de áudio de dissimulação de erro baseada em uma informação de *pitch* anteriormente computada, utilizada para uma decodificação de um ou mais quadros de áudio que antecedem o quadro de áudio perdido, e baseada em uma avaliação de uma correlação cruzada do sinal de excitação de domínio de tempo, modificada para obter um sinal de excitação de domínio de tempo modificado para o fornecimento da informação de áudio de dissimulação de erro. Foi verificado que tendo em consideração ambas a informação de *pitch* anteriormente computada e a informação de *pitch* obtida baseada no sinal de excitação de domínio de tempo (utilizando uma correlação cruzada) melhora a credibilidade

da informação de *pitch* e consequentemente ajuda a evitar perturbações e/ou descontinuidades.

[0106] Em um modelo preferido, a dissimulação de erros é configurada para selecionar um pico da correlação cruzada, de uma pluralidade de picos da correlação cruzada, como um pico representando um *pitch* dependendo na informação de *pitch* anteriormente computada, de modo que um pico seja escolhido representando um *pitch* que se encontra mais próximo do *pitch* representado pela informação de *pitch* anteriormente computada. Por conseguinte, possíveis ambiguidades da correlação cruzada que pode, por exemplo, resultar em múltiplos picos, podem ser ultrapassadas. A informação de *pitch* anteriormente computada é assim utilizada para selecionar o pico “adequado” da correlação cruzada, ajudando a significativamente aumentar a fiabilidade. Por outro lado, o verdadeiro sinal de excitação de domínio de tempo é considerado principalmente para a determinação de *pitch*, fornecendo uma boa exatidão (significativamente melhor do que uma exatidão obtenível baseada em apenas a informação de áudio anteriormente computada).

[0107] Em um modelo preferido, o decodificador áudio a dissimulação de erros pode ser configurado para obter uma informação de áudio baseada em uma informação paralela da informação de áudio codificada.

[0108] Em um modelo preferido, a dissimulação de erros pode estar configurada para obter uma informação de *pitch* baseada em uma informação de *pitch* disponível para um quadro de áudio anteriormente descodificado.

[0109] Em um modelo preferido, a dissimulação de erros é configurada para obter uma informação de *pitch* baseada em uma pesquisa de *pitch* efetuada em um sinal de domínio de tempo ou em um sinal residual.

[0110] Por outras palavras, o *pitch* pode ser transmitido como informação paralela ou poderia também provir do quadro anterior se existir, por exemplo, LTP. A informação de áudio poderia também ser transmitida no fluxo de bits se estivesse disponível no codificador. Podemos opcionalmente pesquisar o *pitch* no sinal de domínio de tempo diretamente ou no residual, fornecendo

habitualmente melhores resultados no residual (sinal de excitação de domínio de tempo).

[0111] Em um modelo preferido, a dissimulação de erros é configurada para copiar um ciclo de *pitch* do sinal de excitação de domínio de tempo associado ao quadro de áudio que antecede o quadro de áudio perdido uma vez ou várias vezes, de modo a obter um sinal de excitação (ou pelo menos um componente determinístico deste) para uma síntese da informação de áudio de dissimulação de erro. Copiando o ciclo de *pitch* do sinal de excitação de domínio de tempo associado ao quadro de áudio que antecede o quadro de áudio perdido uma vez ou várias vezes, e modificando a dita uma ou mais cópias utilizando um algoritmo de modificação comparativamente simples, o sinal de excitação (ou pelo menos o componente determinístico deste) para a síntese da informação de áudio de dissimulação de erro pode ser obtido com pouco esforço computacional. Contudo, reutilizando o quadro de áudio perdido (copiando o dito sinal de excitação de domínio de tempo) evita descontinuidades sonoras.

[0112] Em um modelo preferido, a dissimulação de erros é configurada para filtrar a baixa frequência o ciclo de *pitch* do sinal de excitação de domínio de tempo associado ao quadro de áudio que antecede o quadro de áudio perdido utilizando um filtro dependente da taxa de amostragem, uma largura de banda a qual é dependente de uma taxa de amostragem do quadro de áudio codificado em uma representação de domínio de frequência. Por conseguinte, o sinal de excitação de domínio de tempo é adaptado a uma largura de banda do sinal do decodificador áudio, resultando em uma boa reprodução do conteúdo áudio.

[0113] Para detalhes e melhorias opcionais, é feita referência às explicações fornecidas em cima.

[0114] Por exemplo, é preferível a baixa frequência apenas no primeiro quadro perdido, e de preferência, também a baixa frequência se o sinal não for vozeado. Contudo, dever-se-á ter em atenção que a filtragem de baixa

frequência é opcional. Ainda, o filtro pode ser dependente da taxa de amostragem, de modo que a frequência de corte seja independente da largura de banda.

[0115] Em um modelo preferido, a dissimulação de erros é configurada para predizer um *pitch* em uma parte final de um quadro perdido. Neste caso, a dissimulação de erros é configurada para adaptar o sinal de excitação de domínio de tempo, ou uma ou várias cópias deste, ao *pitch* de predição. Modificando o sinal de excitação de domínio de tempo, de modo que o sinal de excitação de domínio de tempo atualmente utilizado para o fornecimento da informação de áudio de dissimulação de erro seja modificado relativamente ao sinal de excitação de domínio de tempo associado a um quadro de áudio que antecede o quadro de áudio perdido, mudanças de *pitch* esperadas (ou de predição) durante o quadro de áudio perdido podem ser consideradas, de modo que a informação de áudio de dissimulação de erro seja bem adaptada à evolução atual (ou pelo menos à evolução esperada ou de predição) do conteúdo áudio. Por exemplo, a adaptação passa desde o último bom *pitch* para o de predição. Isso é feito pela ressincronização de impulsos [7].

[0116] Em um modelo preferido, a dissimulação de erros é configurada para combinar um sinal de excitação de domínio de tempo extrapolado e um sinal de ruído, de modo a obter um sinal de entrada para uma síntese LPC. Neste caso, a dissimulação de erros é configurada para executar a síntese LPC, em que a síntese LPC é configurada para filtrar o sinal de entrada da síntese LPC dependendo dos parâmetros de codificação de predição linear, de modo a obter a informação de áudio de dissimulação de erro. Combinando o sinal de excitação de domínio de tempo (tipicamente uma versão modificada do sinal de excitação de domínio de tempo derivado de um ou mais quadros de áudio que antecedem o quadro de áudio perdido) e um sinal de ruído, ambos componentes determinísticos (por exemplo, aproximadamente periódicos) e componentes de ruído do conteúdo áudio podem ser considerados na dissimulação de erros. Desse modo, poder-se-á obter que a informação de

áudio de dissimulação de erro forneça uma impressão de audição idêntica à impressão de audição fornecida pelos quadros que antecedem o quadro perdido.

[0117] Também, combinando um sinal de excitação de domínio de tempo e um sinal de ruído, de modo a obter o sinal de entrada para a síntese LPC (que pode ser considerado como um sinal de excitação de domínio de tempo combinado), é possível variar uma percentagem do componente determinístico do sinal de entrada de áudio para a síntese LPC enquanto mantém uma energia (do sinal de entrada da síntese LPC, ou mesmo do sinal de saída da síntese LPC). Por conseguinte, é possível variar as características da informação de áudio de dissimulação de erro (por exemplo, características de tonalidade) sem significativamente alterar uma energia ou sonoridade do sinal de áudio de dissimulação de erro, de modo que seja possível modificar o sinal de excitação de domínio de tempo sem provocar distorções audíveis não aceitáveis.

[0118] Um modelo de acordo com a invenção cria um método para fornecer uma informação de áudio decodificada baseada em uma informação de áudio codificada. O método compreende o fornecimento de uma informação de áudio de dissimulação de erro para dissimulação de uma perda de um quadro de áudio. O fornecimento de uma informação de áudio de dissimulação de erro compreende a modificação de um sinal de excitação de domínio de tempo obtido baseado em um ou mais quadros de áudio que antecedem um quadro de áudio perdido, de modo a obter uma informação de áudio de dissimulação de erro.

[0119] Este método é baseado nas mesmas considerações do decodificador áudio descrito em cima.

[0120] Um modelo adicional de acordo com a invenção cria um programa de computador para executar o dito método quando o programa de computador é executado em um computador.

[0121] Breve Descrição dos Desenhos

[0122] Modelos desta invenção irão subsequentemente ser descritos tendo como referência as figuras incluídas, nas quais:

[0123] A Fig. 1 ilustra um esquema de blocos de um decodificador áudio, de acordo com um modelo da invenção;

[0124] A Fig. 2 ilustra um esquema de blocos de um decodificador áudio, de acordo com outro modelo desta invenção;

[0125] A Fig. 3 ilustra um esquema de blocos de um decodificador áudio, de acordo com outro modelo desta invenção;

[0126] A Fig. 4 ilustra um esquema de blocos de um decodificador áudio, de acordo com outro modelo desta invenção;

[0127] A Fig. 5 ilustra um esquema de blocos de uma dissimulação de domínio de tempo para um codificador por transformação;

[0128] A Fig. 6 ilustra um esquema de blocos de uma dissimulação de domínio de tempo para um codec de comutação;

[0129] A Fig. 7 ilustra um diagrama de blocos de um decodificador TCX que executa uma decodificação TCX em uma operação normal ou no caso de perda parcial de pacotes;

[0130] A Fig. 8 ilustra um esquema de blocos de um decodificador TCX que executa uma decodificação TCX no caso de uma dissimulação de eliminação por pacotes TCX-356;

[0131] A Fig. 9 ilustra um fluxograma de um método para fornecer uma informação de áudio decodificada baseada em uma informação de áudio codificada, de acordo com um modelo desta invenção; e

[0132] A Fig. 10 ilustra um fluxograma de um método para fornecer uma informação de áudio decodificada baseada em uma informação de áudio codificada, de acordo com outro modelo desta invenção.

[0133] A Fig. 11 ilustra um esquema de blocos de um decodificador áudio, de acordo com outro modelo desta invenção.

[0134] Descrição Detalhada dos Modelos

[0135] 1. Decodificador de Áudio de Acordo com a Fig. 1

[0136] A Fig. 1 ilustra um esquema de blocos de um decodificador áudio 100, de acordo com um modelo da invenção. O decodificador áudio 100 recebe uma informação de áudio codificada 110, que pode, por exemplo, compreender um quadro de áudio codificado em uma representação de domínio de frequência. A informação de áudio codificada pode, por exemplo, ser recebida através de um canal inseguro, de modo que uma perda de quadro surja de vez em quando. O decodificador áudio 100 fornece ainda, baseado na informação de áudio codificada 110, a informação de áudio decodificada 112.

[0137] O decodificador áudio 100 pode compreender ainda uma decodificação/ processamento 120, que fornece a informação de áudio decodificada baseada em uma informação de áudio codificada na ausência de uma perda de quadros.

[0138] O decodificador áudio 100 compreende ainda uma dissimulação de erro 130, que fornece uma informação de áudio de dissimulação de erro. A dissimulação de erro 130 é configurada para fornecer uma informação de áudio de dissimulação de erro 132 para dissimular uma perda de um quadro de áudio a seguir a um quadro de áudio codificado na representação de domínio de frequência, utilizando um sinal de excitação de domínio de tempo.

[0139] Por outras palavras, a decodificação/processamento 120 pode fornecer uma informação de áudio decodificada 122 para quadros de áudio que são codificados sob a forma de uma representação de domínio de frequência, isto é, sob a forma de uma representação codificada, valores codificados os quais descrevem intensidades em diferentes binários de frequência. Colocando outras palavras, a decodificação/processamento 120 pode, por exemplo, compreender um decodificador áudio de domínio de frequência, que deriva um conjunto de valores espectrais da informação de áudio codificada 110 e executa um domínio de frequência para domínio de tempo por transformação para desse modo derivar uma representação de domínio de tempo que

constitui a informação de áudio decodificada 122 ou que forma a base para o fornecimento da informação de áudio decodificada 122 no caso de existir um pós-processamento adicional.

[0140] Contudo, a dissimulação de erro 130 não executa a dissimulação de erro de domínio de frequência mas sim utiliza um sinal de excitação de domínio de tempo, que pode, por exemplo, servir para excitar um filtro de síntese, como por exemplo um filtro de síntese LPC, que fornece uma representação de domínio de tempo de um sinal de áudio (por exemplo, a informação de áudio de dissimulação de erro) baseada no sinal de excitação de domínio de tempo e também baseada nos coeficientes do filtro LPC (coeficientes do filtro de codificação de predição linear).

[0141] Por conseguinte, a dissimulação de erro 130 fornece a informação de áudio de dissimulação de erro 132, que pode, por exemplo, ser um sinal de áudio de domínio de tempo, para quadros de áudio perdidos, em que o sinal de excitação de domínio de tempo utilizado pela dissimulação de erro 130 pode ser baseada em, ou derivada de, um ou mais quadros de áudio anteriores, recebidos de modo adequado (que antecedem o quadro de áudio perdido), codificados sob a forma de uma representação de domínio de frequência. Para concluir, o decodificador áudio 100 pode executar uma dissimulação de erro (isto é, fornecer uma informação de áudio de dissimulação de erro 132), que reduz uma degradação de uma qualidade de áudio devido à perda de um quadro de áudio baseado em uma informação de áudio codificada, na qual pelo menos alguns quadros de áudio são codificados em uma representação de domínio de frequência. Foi verificado que a execução da dissimulação de erro utilizando um sinal de excitação de domínio de tempo mesmo se um quadro a seguir a um quadro de áudio codificado recebido de modo adequado na representação de domínio de frequência for perdido, traz consigo uma qualidade de áudio melhorada quando comparado com uma dissimulação de erro que é executada no domínio de frequência (por exemplo, utilizando uma representação de domínio de frequência do quadro de áudio codificado na

representação de domínio de frequência que antecede o quadro de áudio perdido). Isto deve-se ao facto de que uma transição suave entre a informação de áudio decodificada associada ao quadro de áudio recebido de modo adequado que antecede o quadro de áudio perdido e a informação de áudio de dissimulação de erro associada ao quadro de áudio perdido pode ser obtida utilizando um sinal de excitação de domínio de tempo, visto que a síntese de sinal, tipicamente executada baseada no sinal de excitação de domínio de tempo, ajuda a evitar descontinuidades. Desse modo, uma boa (ou pelo menos aceitável) impressão de audição pode ser obtida utilizando o decodificador áudio 100, mesmo se um quadro de áudio for perdido a seguir a um quadro de áudio codificado recebido de modo adequado na representação de domínio de frequência. Por exemplo, a abordagem de domínio de tempo traz melhoria no sinal monofónico, como a fala, pois encontra-se mais próximo ao que é feito no caso de dissimulação de codec de fala. A utilização de LPC ajuda a evitar descontinuidades e dá uma melhor forma dos quadros.

[0142] Além disso, dever-se-á ter em atenção que o decodificador áudio 100 pode ser complementado por qualquer uma das características e funcionalidades descritas a seguir, seja individualmente ou combinadas.

[0143] 2. Decodificador Áudio de Acordo com a Fig. 2

[0144] A Fig. ilustra um esquema de blocos de um decodificador áudio 200, de acordo com um modelo desta invenção. O decodificador áudio 200 é configurado para receber uma informação de áudio codificada 210 e para fornecer, baseado nisto, uma informação de áudio decodificada 220. A informação de áudio codificada 210 pode, por exemplo, tomar a forma de uma sequência de quadros de áudio codificados em uma representação de domínio de tempo, codificada em uma representação de domínio de frequência, ou codificada em ambas uma representação de domínio de tempo e uma representação de domínio de frequência. Por outras palavras, todos os quadros da informação de áudio codificada 210 podem ser codificados em uma representação de domínio de frequência, ou todos os quadros da informação

de áudio codificada 210 podem ser codificados em uma representação de domínio de tempo (por exemplo, sob a forma de um sinal de excitação de domínio de tempo codificado e parâmetros de síntese do sinal codificado, como por exemplo, parâmetros LPC). Em alternativa, alguns quadros da informação de áudio codificada podem ser codificados em uma representação de domínio de frequência, e alguns outros quadros da informação de áudio codificada podem ser codificados em uma representação de domínio de tempo, por exemplo, se o decodificador áudio 200 dos um decodificador de áudio de comutação que pode alternar entre diferentes modos de decodificação. A informação de áudio decodificada 220 pode, por exemplo, ser uma representação de domínio de tempo de um ou mais canais de áudio.

[0145] O decodificador áudio 200 pode tipicamente compreender uma decodificação/processamento 220, que pode, por exemplo, fornecer uma informação de áudio decodificada 232 para quadros de áudio recebidos de modo adequado. Por outras palavras, a decodificação/processamento 230 pode executar uma decodificação de domínio de frequência (por exemplo, uma decodificação do tipo AAC, ou idêntico) baseada em um ou mais quadros de áudio codificados em uma representação de domínio de frequência. Em alternativa, ou além disso, a decodificação/processamento 230 pode ser configurada para executar uma decodificação de domínio de tempo (ou decodificação de domínio de predição linear) baseada em um ou mais quadros de áudio codificados em uma representação de domínio de tempo (ou, por outras palavras, em uma representação de domínio de predição linear), como, por exemplo, uma decodificação de predição linear TCX excitada (TCX = transformação de excitação codificada) ou uma decodificação ACELP (decodificação algébrica de livro de códigos excitada de predição linear). Como opção, a decodificação/processamento 230 pode ser configurada para mudar entre diferentes modos de decodificação.

[0146] O decodificador áudio 200 compreende ainda uma dissimulação de erro 240, configurada para fornecer uma informação de áudio de dissimulação

de erro 242 para um ou mais quadros de áudio perdidos. A dissimulação de erro 240 é configurada para fornecer a informação de áudio de dissimulação de erro 242 para dissimular uma perda de um quadro de áudio (ou mesmo uma perda de quadros de áudio múltiplos). A dissimulação de erro 240 é configurada para modificar um sinal de excitação de domínio de tempo obtido baseado em um ou mais quadros de áudio que antecedem um quadro de áudio perdido, de modo a obter a informação de áudio de dissimulação de erro 242. Por outras palavras, a dissimulação de erro 240 pode obter (ou derivar) um sinal de excitação de domínio de tempo para (ou baseado nisso) um ou mais quadros de áudio codificados que antecedem um quadro de áudio perdido, e pode modificar o dito sinal de excitação de domínio de tempo, obtido para (ou baseado nisso) um ou mais quadros de áudio recebidos de modo adequado que antecedem um quadro de áudio perdido, para desse modo obter (através da modificação) um sinal de excitação de domínio de tempo utilizado para fornecer a informação de áudio de dissimulação de erro 242. Por outras palavras, o sinal de excitação de domínio de tempo modificado pode ser utilizado como uma entrada (ou como um componente de uma entrada) para uma síntese (por exemplo, síntese LPC) da informação de áudio de dissimulação de erro associada ao quadro de áudio perdido (ou mesmo com múltiplos quadros de áudio perdidos). Com o fornecimento da informação de áudio de dissimulação de erro 242 baseada no sinal de excitação de domínio de tempo obtido baseado em um ou mais quadros de áudio recebidos de modo adequado que antecedem o quadro de áudio perdido, as descontinuidades audíveis podem ser evitadas. Por outro lado, modificando o sinal de excitação de domínio de tempo derivado para (ou de) um ou mais quadros de áudio que antecedem o quadro de áudio perdido, e fornecendo a informação de áudio de dissimulação de erro baseada no sinal de excitação de domínio de tempo modificado, é possível considerar várias características do conteúdo de áudio (por exemplo, uma alteração de *pitch*), e é também possível evitar uma impressão de audição não natural (por exemplo, através do “desvanecimento”

de um componente de sinal determinístico (por exemplo, pelo menos aproximadamente periódico)). Desse modo, pode obter-se que a informação de áudio de dissimulação de erro 242 compreende alguma semelhança com a informação de áudio decodificada 232 obtida baseada em quadros de áudio descodificados de modo adequado que antecedem o quadro de áudio perdido, e pode ainda ser obtido que a informação de áudio de dissimulação de erro 242 compreenda um conteúdo áudio um tanto ou quanto diferente quando comparado à informação de áudio decodificada 232 associada ao quadro de áudio que antecede o quadro de áudio perdido modificando um pouco o sinal de excitação de domínio de tempo. A modificação do sinal de excitação de domínio de tempo utilizado para o fornecimento da informação de áudio de dissimulação de erro (associada ao quadro de áudio perdido) pode, por exemplo, compreender uma escala de amplitude ou uma escala de tempo. Contudo, outros tipos de modificação (ou mesmo uma combinação de uma escala de amplitude e uma escala de tempo) são possíveis, sendo que de preferência um certo grau de relacionamento entre o sinal de excitação de domínio de tempo obtido (como uma informação de entrada) pela dissimulação de erro e o sinal de excitação de domínio de tempo modificado deverá permanecer.

[0147] Para concluir, o decodificador áudio 200 permite fornecer a informação de áudio de dissimulação de erro 242, de modo que a informação de áudio de dissimulação de erro forneça uma boa impressão de audição mesmo no caso em que um ou mais quadros de áudio perdidos sejam perdidos. A dissimulação de erro é executada baseada em um sinal de excitação de domínio de tempo, em que uma variação das características do sinal do conteúdo de áudio durante o quadro de áudio perdido seja considerada pela modificação do sinal de excitação de domínio de tempo obtido baseado em um ou mais quadros de áudio que antecedem um quadro perdido.

[0148] Além disso, dever-se-á ter em atenção que o decodificador áudio 200 pode ser complementado por qualquer uma das características e

funcionalidades aqui descritas, individualmente ou combinadas.

[0149] 3. Decodificador Áudio de Acordo com a Fig. 3

[0150] A Fig. 3 ilustra um esquema de blocos de um decodificador áudio 300, de acordo com outro modelo desta invenção.

[0151] O decodificador áudio 300 é configurado para receber uma informação de áudio codificada 310 e para fornecer, baseado nisso, uma informação de áudio decodificada 312. O decodificador áudio 300 comprehende um analisador de fluxo de bits 320, que pode ser também designado como um “deformatador de fluxo de bits” ou “analisador sintático de fluxo de bits”. O analisador de fluxo de bits 320 recebe a informação de áudio codificada 310 e fornece, baseado nisso, uma representação de domínio de frequência 322 e possivelmente informação de controlo adicional 324. A representação de domínio de frequência 322 pode, por exemplo, compreender valores espectrais codificados 326, fatores de escala codificados 328 e, como opção, uma informação paralela adicional 330 que podem por exemplo controlar etapas de processamento de controlo específicas, como, por exemplo, um preenchimento de ruído, um processamento intermédio ou um pós-processamento. O decodificador áudio 300 comprehende também uma decodificação do valor espectral 340 configurada para receber os valores espectrais codificados 326, e para fornecer, baseado nisso, um conjunto de valores espectrais descodificados 342. O decodificador áudio 300 pode também compreender uma decodificação do fator de escala 350, que pode ser configurada para receber os fatores de escala codificados 328 e para fornecer, baseado nisso, um conjunto de fatores de escala descodificados 352.

[0152] Em alternativa à decodificação do fator de escala, uma conversão do fator LPC à escala 354 pode ser utilizada, por exemplo, no caso de que a informação de áudio codificada compreende uma informação LPC codificada, em vez de uma informação do fator de escala. Contudo, em alguns modos de codificação (por exemplo, no modo de decodificação TCX do decodificador áudio USAC ou no decodificador áudio EVS) um conjunto de coeficientes LPC

pode ser utilizado para derivar um conjunto de fatores de escala em paralelo ao decodificador áudio. Esta funcionalidade pode ser alcançada pela conversão do fator LPC à escala 354.

[0153] O decodificador áudio 300 pode compreender também um escalador 360, que pode ser configurado para aplicar um conjunto de fatores escalonados 352 ao conjunto de valores espectrais 342, para desse modo obter um conjunto de valores espectrais escalonados descodificados 362. Por exemplo, uma primeira banda de frequências compreendendo valores espectrais descodificados múltiplos 342 pode ser escalonada utilizando um primeiro fator de escala, e uma segunda banda de frequências compreendendo valores espectrais descodificados múltiplos 342 pode ser escalonada utilizando um segundo fator de escala. Por conseguinte, o conjunto de valores espectrais escalonados descodificados 362 é obtido. O decodificador áudio 300 pode compreender ainda um processamento opcional 366, que pode aplicar algum processamento aos valores espectrais escalonados descodificados 362. Por exemplo, o processamento opcional 366 pode compreender um preenchimento de ruído ou algumas outras operações.

[0154] O decodificador áudio 300 compreende também um domínio de frequência ao domínio do tempo por transformação 370, configurado para receber os valores espectrais escalonados descodificados 362, ou uma versão processada 368 destes, e para fornecer uma representação de domínio de tempo 372 associada a um conjunto de valores espectrais escalonados descodificados 362. Por exemplo, o domínio de frequência ao domínio do tempo por transformação 370 pode fornecer uma representação de domínio de tempo 372, associada a um quadro ou subquadro do conteúdo de áudio. Por exemplo, o domínio de frequência ao domínio do tempo por transformação pode receber um conjunto de coeficientes MDCT (que podem ser considerados como valores espectrais escalonados descodificados) e fornecer, baseado nisso, um bloco de amostras de domínio de tempo, que pode formar a representação de domínio de tempo 372.

[0155] O decodificador áudio 300 pode opcionalmente compreender um pós-processamento 376, que pode receber a representação de domínio de tempo 372 e de algum modo modificar a representação de domínio de tempo 372, para desse modo obter uma versão pós-processada 378 da representação de domínio de tempo 372.

[0156] O decodificador áudio 300 pode compreender também uma dissimulação de erro 380 que pode, por exemplo, receber uma representação de domínio de tempo 372 do domínio de frequência ao domínio do tempo por transformação 370 e que pode, por exemplo, fornecer uma informação de áudio da dissimulação de erro 382 para um ou mais quadros perdidos. Por outras palavras, se um quadro de áudio for perdido, de modo que, por exemplo, nenhum valor espectral codificado 326 estão disponíveis para o dito quadro de áudio (ou subquadro de áudio), a dissimulação de erro 380 pode fornecer a informação de áudio da dissimulação de erro baseada na representação de domínio de tempo 372 associada a um ou mais quadros de áudio que antecedem o quadro de áudio perdido. A informação de áudio da dissimulação de erro pode tipicamente ser uma representação de domínio de tempo de um conteúdo áudio.

[0157] Dever-se-á ter em atenção que a dissimulação de erro 380 pode, por exemplo, executar a funcionalidade da dissimulação de erro 130 descrita em cima. Também, a dissimulação de erro 380 pode, por exemplo, compreender a funcionalidade da dissimulação de erro 500 descrita tendo como referência a Fig. 5. Contudo, regra geral falando, a dissimulação de erro 380 pode compreender quaisquer das características e funcionalidades aqui descritas relativamente à dissimulação de erro.

[0158] Relativamente à dissimulação de erro, dever-se-á ter em atenção que a dissimulação de erro não acontece ao mesmo tempo da decodificação do quadro. Por exemplo se o quadro n é bom então efetuamos uma decodificação normal, e no final salvamos alguma variável que irá ajudar se tivermos de dissimular o próximo quadro, então se n+1 for perdido sinalizamos

a função de dissimulação que dá a variável proveniente do quadro bom anterior. Iremos também atualizar algumas variáveis para ajudar a próxima perda de quadro ou na recuperação para o próximo quadro bom.

[0159] O decodificador áudio 300 comprehende também uma combinação de sinais 390, configurada para receber a representação de domínio de tempo 372 (ou a representação de domínio de tempo pós-processada 378 no caso de existir um pós-processamento 376). Além disso, a combinação de sinal 390 pode receber a informação de áudio de dissimulação de erro 382, tipicamente também uma representação de domínio de tempo de um sinal de áudio de dissimulação de erro fornecida para um quadro de áudio perdido. A combinação de sinais 390 pode, por exemplo, combinar representações de domínio de tempo associadas aos quadros de áudio posteriores. No caso de existirem quadros de áudio posteriores descodificados de modo adequado, a combinação de sinais 390 pode combinar (por exemplo, sobrepor e adicionar) representações de domínio de tempo associadas a estes quadros de áudio posteriores descodificados de modo adequado. Contudo, se um quadro de áudio for perdido, a combinação de sinais 390 pode combinar (por exemplo, sobrepor e adicionar) a representação de domínio de tempo associada ao quadro de áudio descodificado de modo adequado que antecede o quadro de áudio perdido e a informação de áudio de dissimulação de áudio associada ao quadro de áudio perdido, para desse modo existir uma transição suave entre o quadro de áudio recebido de modo adequado e o quadro de áudio perdido. Do mesmo modo, a combinação de sinais 390 pode ser configurada para combinar (por exemplo, sobrepor e adicionar) a informação de áudio de dissimulação de erro associada ao quadro de áudio perdido e a representação de domínio de tempo associada a outro quadro de áudio descodificado de modo adequado a seguir ao quadro de áudio perdido (ou outra informação de áudio de dissimulação de erro associada a outro quadro de áudio perdido no caso de quadros de áudio consecutivos múltiplos se terem perdido).

[0160] Por conseguinte, a combinação de sinais 390 pode fornecer uma

informação de áudio decodificada 312, de modo que a representação de domínio de tempo 372, ou uma sua versão pós-processada 378, seja fornecido para quadros de áudio descodificados de modo adequado, e de modo que a informação de áudio de dissimulação de erro 382 seja fornecida para quadros de áudio perdidos, em que uma operação de sobreposição e adição seja tipicamente executada entre a informação de áudio (independentemente de se é fornecida pelo domínio de frequência ao domínio do tempo por transformação 370 ou pela dissimulação de erro 380) de quadros de áudio posteriores. Visto que alguns codecs são dotados de alguns efeitos de escada na parte de sobreposição e de adição que precisam ser cancelados, como opção podemos criar algum efeito de escada em metade de um quadro que criámos para executar a sobreposição adição.

[0161] Dever-se-á ter em atenção que a funcionalidade do decodificador áudio 300 é idêntica à funcionalidade do decodificador áudio 100 de acordo com a Fig. 1, em que detalhes adicionais são ilustrados na Fig. 3. Além disso, dever-se-á ter em atenção que o descodificados áudio 300 de acordo com a Fig. 3 pode ser complementado por quaisquer das características e funcionalidades aqui descritas. Em especial, a dissimulação de erro 380 pode ser complementada por quaisquer características e funcionalidades aqui descritas relativamente à dissimulação de erro.

[0162] 4. Decodificador Áudio 400 de Acordo com a Fig. 4

[0163] A Fig. 4 ilustra um decodificador áudio 400, de acordo com outro modelo desta invenção. O decodificador áudio 400 é configurado para receber uma informação de áudio codificada e para fornecer, baseado nisso, uma informação de áudio decodificada 412. O decodificador áudio 400 pode, por exemplo, ser configurado para receber uma informação de áudio codificada 410, em que diferentes quadros de áudio são codificados utilizando diferentes modos de codificação. Por exemplo, o decodificador áudio 400 pode ser considerado como um decodificador áudio multimodos ou um decodificador áudio “comutador”. Por exemplo, alguns dos quadros de áudio podem ser

codificados utilizando uma representação de domínio de frequência, em que a informação de áudio codificada compreende uma representação codificada de valores espectrais (por exemplo, valores FFT ou valores MDCT) e fatores de escala representam um escalonamento de diferentes bandas de frequência. Além disso, a informação de áudio codificada 410 pode compreender também uma “representação de domínio de tempo” de quadros de áudio, ou uma “representação de domínio de codificação de predição linear” de quadros de áudio múltiplos. A “representação de domínio de codificação de predição linear” (também laconicamente designada como “representação LPC”) pode, por exemplo, compreender uma representação codificada de um sinal de excitação, e uma representação codificada de parâmetros LPC (parâmetros de codificação de predição linear), em que os parâmetros de codificação de predição linear descrevem, por exemplo, um filtro de síntese de codificação de predição linear, utilizado para reconstruir um sinal de áudio baseado no sinal de excitação de domínio de tempo.

[0164] A seguir, alguns detalhes do decodificador áudio 400 serão descritos.

[0165] O decodificador áudio 400 compreende um analisador de fluxo de bits 420 que pode, por exemplo, analisar a informação de áudio codificada 410 e extraír, a partir da informação de áudio codificada 410, uma representação de domínio de frequência 422, compreendendo, por exemplo, valores espectrais codificados, fatores de escala codificados e, opcionalmente, uma informação paralela adicional. O analisador de bits 420 pode ser também configurado para extraír uma representação de domínio de predição linear 424, que pode, por exemplo, compreender uma excitação codificada 426 e coeficientes de predição linear codificados 428 (que podem ser também considerados como parâmetros de predição linear codificados). Além disso, o analisador de bits pode opcionalmente extraír informação paralela adicional, que pode ser utilizada para controlar etapas de processamento adicional, a partir da informação de áudio codificada.

[0166] O decodificador áudio 400 comprehende uma trajetória de decodificação de domínio de frequência 430, que pode, por exemplo, ser significativamente idêntica à trajetória de decodificação do decodificador áudio 300 de acordo com a Fig. 3. Por outras palavras, a trajetória de decodificação de domínio de frequência 430 pode compreender uma decodificação do valor espectral 340, uma decodificação do fator de escala 350, um escalador 360, um processamento opcional 366, um domínio de frequência para domínio de tempo por transformação 370, um pós-processamento opcional 376 e uma dissimulação de erro 380 tal como descrito em cima com referência à Fig. 3.

[0167] O decodificador áudio 400 pode compreender também uma trajetória de decodificação de domínio de predição linear 440 (que pode ser também considerada como uma trajetória de decodificação de domínio de tempo, visto que a síntese LPC é executada no domínio de tempo). A trajetória de decodificação de domínio de predição linear comprehende uma decodificação de excitação 450, que recebe a excitação codificada 426 fornecida pelo analisador do fluxo de bits 420 e fornece, baseado nisso, uma excitação decodificada 452 (que pode ter a forma de um sinal de excitação de domínio de tempo descodificado). Por exemplo, a decodificação de excitação 450 pode receber uma informação de excitação de codificação por transformação codificada, e pode fornecer, baseado nisso, um sinal de excitação de domínio de tempo descodificado. Desse modo, a decodificação de excitação 450 pode, por exemplo, executar uma funcionalidade que é executada pelo decodificador de excitação 730 descrito tendo como referência a Fig. 7. Contudo, alternativamente ou além disso, a decodificação de excitação 450 pode receber uma excitação ACELP, e pode fornecer o sinal de excitação de domínio de tempo descodificado 452 baseado na dita informação de excitação ACELP codificada.

[0168] Dever-se-á ter em atenção que existem três diferentes opções para a decodificação de excitação. Referência é feita, por exemplo, às Normas e publicações relevantes que definem os conceitos de codificação CELP, os

conceitos de codificação ACELP, modificações dos conceitos de codificação CELP e dos conceitos de codificação ACELP e o conceito de codificação TCX.

[0169] A trajetória de decodificação de domínio de predição linear 440 opcionalmente compreende um processamento 454 no qual um sinal de excitação de domínio de tempo processado 456 é derivado do sinal de excitação de domínio de tempo 452.

[0170] A trajetória de decodificação de domínio de predição linear 440 compreende também uma decodificação do coeficiente de predição linear 460, configurado para receber coeficientes de predição linear codificados recebidos e para fornecer, baseado nisso, coeficientes de predição linear descodificados 462. A decodificação do coeficiente de predição linear 460 pode utilizar diferentes representações de um coeficiente de predição linear como uma informação de entrada 428 e pode fornecer diferentes representações dos coeficientes de predição linear descodificados como informação de saída 462. Para detalhes, referência a ser feita aos diferentes documentos Normalizados nos quais uma codificação e/ou decodificação de coeficientes de predição linear é descrita.

[0171] A trajetória de decodificação de domínio de predição linear 440 compreende opcionalmente um processamento 464, que pode processar os coeficientes de predição linear descodificados e fornecer uma sua versão processada 466.

[0172] A trajetória de decodificação de domínio de predição linear 440 compreende também uma síntese LPC (síntese de codificação de predição linear) 470, configurada para receber uma excitação decodificada 452, ou a sua versão processada 456, e os coeficientes de predição linear descodificados 462, ou a sua versão processada 466, e para fornecer um sinal de áudio de domínio de tempo descodificado 472. Por exemplo, a síntese LPC 470 pode ser configurada para aplicar uma filtragem, definida pelos coeficientes de predição linear descodificados 462 (ou uma sua versão processada 466) ao sinal de excitação de domínio de tempo descodificado 452, ou à sua versão

processada, de modo que o sinal de áudio de domínio de tempo descodificado 472 seja obtido por filtragem (filtragem da síntese) do sinal de excitação de domínio de tempo 452 (ou 456). A trajetória de decodificação de domínio de predição linear 440 pode opcionalmente compreender um pós-processamento 474, que pode ser utilizado para refinar ou regular as características do sinal de áudio de domínio de tempo descodificado 472.

[0173] A trajetória de decodificação de domínio de predição linear 440 compreende também uma dissimulação de erro 480, configurada para receber os coeficientes de predição linear descodificados 462 (ou a sua versão processada 466) e o sinal de excitação de domínio de tempo descodificado 452 (ou a sua versão processada 456). A dissimulação de erro 480 pode opcionalmente receber informação adicional, como por exemplo uma informação de *pitch*. A dissimulação de erro 480 pode consequentemente fornecer uma informação de áudio de dissimulação de erro, que pode ser sob a forma de um sinal de áudio de domínio de tempo, no caso que um quadro (ou subquadro) da informação de áudio codificada 410 seja perdido. Desse modo, a dissimulação de erro 480 pode fornecer a informação de áudio de dissimulação de erro 482 de modo que as características da informação de áudio de dissimulação de erro 482 sejam significativamente adaptadas às características de um último quadro de áudio descodificado de modo adequado que antecede o quadro de áudio perdido. Dever-se-á ter em atenção que a dissimulação de erro 480 pode compreender quaisquer das características e funcionalidades descritas relativamente à dissimulação de erro 240. Além disso, dever-se-á ter em atenção que a dissimulação de erro 480 pode compreender também quaisquer das características e funcionalidades descritas relativamente à dissimulação de domínio de tempo da Fig. 6.

[0174] O decodificador áudio 400 compreende também um combinador de sinal (ou combinação de sinais 490), configurado para receber o sinal de áudio de domínio de tempo descodificado 372 (ou a sua versão pós-processada 378), a informação de áudio da dissimulação de erro 382 fornecida pela dissimulação

de erro 380, o sinal de áudio de domínio de tempo descodificado 472 (ou a sua versão pós-processada 476) e a informação de áudio de dissimulação de erro 482 fornecida pela dissimulação de erro 480. O combinador de sinal 490 pode ser configurado para combinar os ditos sinais 372 (ou 378), 382, 472 (ou 476) e 482 para desse modo obter a informação de áudio decodificada 412. Em especial, uma operação de sobreposição e adição pode ser aplicada pelo combinador de sinal 490. Por conseguinte, o combinador de sinal 490 pode fornecer transições suaves entre os quadros de áudio posteriores para os quais o sinal de áudio de domínio de tempo é fornecido pelas diferentes entidades (por exemplo, pelas diferentes trajetórias de decodificação 430, 440). Contudo, o combinador de sinal 490 pode fornecer também transições suaves se o sinal de áudio de domínio de tempo for fornecido pela mesma entidade (por exemplo, domínio de frequência ao domínio do tempo por transformação 370 ou síntese LPC 470) para quadros posteriores. Visto que alguns codecs têm algum efeito de escada na parte de sobreposição e adição que precisam ser canceladas, opcionalmente podemos criar algum efeito de escada artificial em metade de um quadro que criámos para executar a sobreposição adição. Por outras palavras, uma compensação de efeito de escada de domínio de tempo artificial (TDAC) pode ser opcionalmente utilizada.

[0175] Também, o combinador de sinal 490 pode fornecer transições suaves a e dos quadros para os quais uma informação de áudio de dissimulação de erros (tipicamente também um sinal de áudio de domínio de tempo) é fornecida.

[0176] Resumindo, o decodificador áudio 400 permite descodificar quadros de áudio que são codificados no domínio de frequência e quadros de áudio que são codificados no domínio de predição linear. Em especial, é possível mudar entre uma utilização da trajetória de decodificação de domínio de frequência e uma utilização da trajetória de decodificação de domínio de predição linear dependendo das características do sinal (por exemplo, utilizando uma informação de sinalização fornecida por um codificador de áudio). Diferentes

tipos de dissimulação de erro podem ser utilizados para fornecerem uma informação de áudio de dissimulação de erro no caso de uma perda de quadros, dependendo de se um último quadro de áudio descodificado de modo adequado foi codificado no domínio de frequência (ou, de modo equivalente, em uma representação de domínio de frequência), ou no domínio de tempo (ou de modo equivalente, em uma representação de domínio de tempo, ou, de modo equivalente, em um domínio de predição linear, ou, de modo equivalente, em uma representação de domínio de predição linear).

[0177] 5. Dissimulação de Domínio de Tempo de Acordo com a Fig. 5

[0178] A Fig. 5 ilustra um esquema de blocos de uma dissimulação de erro de acordo com um modelo desta invenção. A dissimulação de erro de acordo com a Fig. 5 é indicada na sua totalidade como 500.

[0179] A dissimulação de erro 500 é configurada para receber um sinal de áudio de domínio de tempo 510 e para fornecer, baseada nisto, uma informação de áudio de dissimulação de erro 512, que pode, por exemplo, tomar a forma de um sinal de áudio de domínio de tempo.

[0180] Dever-se-á ter em atenção que a dissimulação de erro 500 pode, por exemplo, tomar a forma da dissimulação de erro 130, de modo que essa informação de áudio de dissimulação de erro 512 possa corresponder à informação de áudio de dissimulação de erro 132. Além disso, dever-se-á ter em atenção que a dissimulação de erro 500 pode substituir a dissimulação de erro 380, de modo que o sinal de áudio de domínio de tempo 510 possa corresponder ao sinal de áudio de domínio de tempo 372 (ou ao sinal de áudio de domínio de tempo 378), e que a informação de áudio de dissimulação de erro 512 possa corresponder à informação de áudio de dissimulação de erro 382.

[0181] A dissimulação de erro 500 comprehende uma pré-ênfase 520, que pode ser considerada como opcional. A pré-ênfase recebe o sinal de áudio de domínio de tempo e fornece, baseado nisso, um sinal de áudio de domínio de tempo pré-enfatizado 522.

[0182] A dissimulação de erro 500 comprehende também uma análise LPC 530, configurada para receber um sinal de áudio de domínio de tempo 510, ou a sua versão pré-enfatizada 522, e para obter uma informação LPC 532, que pode compreender um conjunto de parâmetros LPC 532. Por exemplo, a informação LPC pode compreender um conjunto de coeficientes de filtro LPC (ou uma representação deste) e um sinal de áudio de domínio de tempo (adaptado para uma excitação de um filtro de análise LPC configurado de acordo com os coeficiente de filtro LPC, para reconstruir, pelo menos aproximadamente, o sinal de entrada da análise LPC).

[0183] A dissimulação de erro 500 comprehende também uma pesquisa de *pitch* 540, configurado para obter uma informação de *pitch* 542, por exemplo, baseado em um quadro de áudio anteriormente descodificado.

[0184] A dissimulação de erro 500 comprehende também uma extração 550, que pode ser configurada para obter um sinal de excitação de domínio de tempo extrapolado baseado no resultado da análise LPC (por exemplo, baseado no sinal de excitação de domínio de tempo determinado pela análise LPC), e possivelmente baseado no resultado da pesquisa de *pitch*.

[0185] A dissimulação de erro 500 comprehende também uma geração de ruído 560, que fornece um sinal de ruído 562. A dissimulação de erro 500 comprehende também um combinador/attenuador 570, configurado para receber o sinal de excitação de domínio de tempo extrapolado 552 e o sinal de ruído 562, e para fornecer, baseado nisso, um sinal de excitação de domínio de tempo combinado 572. O combinador/attenuador 570 pode ser configurado para combinar o sinal de excitação de domínio de tempo extrapolado 552 e o sinal de ruído 562, em que um desvanecimento pode ser executado, de modo que um contributo relativo do sinal de excitação de domínio de tempo extrapolado 552 (que determina um componente determinístico do sinal de entrada da síntese LPC) diminui ao longo do tempo enquanto um contributo relativo do sinal de ruído 562 aumenta ao longo do tempo. Contudo, uma diferente funcionalidade do combinador/attenuador é também possível.

Também, referência á feita à descrição em baixo.

[0186] A dissimulação de erro 500 comprehende também uma síntese LPC 580, que recebe o sinal de excitação de domínio de tempo combinado 572 e que fornece um sinal de áudio de domínio de tempo 582 baseado neste. Por exemplo, a síntese LPC pode receber também coeficientes de filtro LPC que descrevem um filtro de modulação LPC, aplicado ao sinal de excitação de domínio de tempo combinado 572, para derivar o sinal de áudio de domínio de tempo 582. A síntese LPC 580 pode, por exemplo, utilizar coeficientes LPC obtidos baseados em um ou mais quadros de áudio previamente descodificados (por exemplo, fornecidos pela análise LPC 530).

[0187] A dissimulação de erro 500 comprehende também uma desacentuação 584, que pode ser considerada como opcional. A desacentuação 584 pode fornecer um sinal de áudio de domínio de tempo da dissimulação de erro desacentuada 586.

[0188] A dissimulação de erro 500 comprehende também, opcionalmente, uma sobreposição e adição 590, que executa uma operação de sobreposição e adição dos sinais de áudio de domínio de tempo associados aos quadros (ou subquadros) subsequentes. Contudo, dever-se-á ter em atenção que a sobreposição e adição 590 deverão ser consideradas como opcionais, visto que a dissimulação de erro pode também utilizar uma combinação do sinal já fornecido no ambiente do decodificador áudio. Por exemplo, a sobreposição e adição 590 podem ser substituídas pela combinação do sinal 390 no decodificador áudio 300 em alguns modelos.

[0189] A seguir, alguns detalhes adicionais relativamente à dissimulação de erro 500 serão descritos.

[0190] A dissimulação de erro 500 de acordo com a Fig. 5 abrange o contexto de um codec de domínio de transformação como AAC_LC ou AAC_ELD. Por outras palavras, a dissimulação de erro 500 é bem adaptada para utilização nesse codec de domínio por transformação (e, em especial, em um decodificador de áudio de domínio por transformação). No caso de um

codec por transformação apenas (por exemplo, na ausência de uma trajetória de decodificação de domínio de predição linear), um sinal de saída de um último quadro é utilizado como um ponto de partida. Por exemplo, o sinal de áudio de domínio de tempo 372 pode ser utilizado como um ponto de partida para a dissimulação de erro. De preferência, nenhum sinal de excitação está disponível, apenas um sinal de saída de domínio de tempo de (um ou mais) quadros anteriores (coo, por exemplo, o sinal de áudio de domínio de tempo 372).

[0191] A seguir, as sub-unidades e funcionalidades da dissimulação de erro 500 serão descritas com mais detalhe.

[0192] 5.1. Análise LPC

[0193] No modelo de acordo com a Fig. 5, toda a dissimulação é feita no domínio de excitação para obter uma transição mais suave entre quadros consecutivos. Por isso, é necessário primeiro encontrar (ou, mais geralmente, obter) um conjunto adequado de parâmetros LPC. No modelo de acordo com a Fig. 5, uma análise LPC 530 é feita no sinal de domínio de tempo anteriormente pré-enfatizado 522. Os parâmetros LPC (ou os coeficientes LPC) são utilizados para efetuar a análise LPC do sinal anterior (por exemplo, baseado no sinal de áudio de domínio de tempo 510, ou baseado no sinal de domínio de tempo anteriormente pré-enfatizado 522) para obter um sinal de excitação (por exemplo, um sinal de excitação de domínio de tempo).

[0194] 5.2 Pesquisa de *Pitch*

[0195] Existem diferentes abordagens para obter o *pitch* a ser utilizado para construir o novo sinal (por exemplo, a informação de áudio de dissimulação de erro).

[0196] No contexto do codec utilizando um filtro LPC (filtro de predição a longo prazo), como AAC-LTP, se o último quadro tiver sido AAC com LTP, utilizamos este último desfasamento de *pitch* LTP recebido e o ganho correspondente para a geração da parte harmónica. Neste caso, o ganho é utilizado para decidir se deverá construir a parte harmónica no sinal ou não.

Por exemplo, se o ganho LTP for superior a 0,6 (ou qualquer outro valor pré-determinado), então a informação LTP é utilizada para construir a parte harmónica.

[0197] Se não existir qualquer informação de *pitch* disponível do quadro anterior, então existem, por exemplo, duas soluções, que serão descritas a seguir.

[0198] Por exemplo, é possível efetuar uma pesquisa de *pitch* no codificador e transmitir no fluxo de bits o desfasamento de *pitch* e o ganho. Isto é idêntico ao LTP, mas não existe qualquer filtragem aplicada (também na filtragem LTP no canal limpo).

[0199] Em alternativa, é possível efetuar uma pesquisa de *pitch* no decodificador. A pesquisa de *pitch* AMR-WB no caso de TCX é feita no domínio FFT. Em ELD, por exemplo, se o domínio MDCT foi utilizado então as fases estariam em falta. Desse modo, a pesquisa de *pitch* é de preferência feita diretamente no domínio de excitação. Isto fornece melhores resultados do que efetuar a pesquisa de *pitch* no domínio de síntese. A pesquisa de *pitch* no domínio de excitação é feita em primeiro lugar com um ciclo aberto através de uma correlação cruzada normalizada. Então, opcionalmente, refinamos a pesquisa de *pitch* efetuando uma pesquisa em ciclo fechado em redor do *pitch* de ciclo aberto com um certo delta. Devido às limitações de janelas ELD, um *pitch* errado poderia ser encontrado, pelo que verificamos também que o *pitch* encontrado é corrigido ou de outro modo descartá-lo.

[0200] Para concluir, o *pitch* do último quadro de áudio descodificado de modo adequado que antecede o quadro de áudio perdido pode ser considerado quando fornece a informação de áudio de dissimulação de erro. Em alguns casos, existe uma informação de *pitch* disponível da decodificação do quadro anterior (isto é, o último quadro que antecede o quadro de áudio perdido). Neste caso, este *pitch* pode ser reutilizado (possivelmente com alguma extração e uma consideração de uma mudança de *pitch* ao longo do tempo). Podemos também opcionalmente reutilizar o *pitch* de mais de um

quadro anterior para tentar extrapolar o *pitch* que precisamos no final do nosso quadro dissimulado.

[0201] Também, se existir uma informação (por exemplo, designada como ganho de predição a longo prazo) disponível, que descreva uma intensidade (ou intensidade relativa) de um componente de sinal determinístico (por exemplo, pelo menos aproximadamente periódico), este valor pode ser utilizado para decidir se um componente determinístico (ou harmónico) deverá ser incluído na informação de áudio de dissimulação de erro. Por outras palavras, ao comparando o dito valor (por exemplo, ganho LTP) com um valor limite pré-determinado, pode ser decidido se um sinal de excitação de domínio de tempo derivado de um quadro de áudio anteriormente descodificado deverá ser considerado para o fornecimento da informação de áudio de dissimulação de erro ou não.

[0202] Se não existir informação de *pitch* disponível do quadro anterior (ou, mais precisamente, da decodificação do quadro anterior), existem diferentes opções. A informação de *pitch* poderá ser transmitida de um codificador de áudio para um decodificador de áudio, o que iria simplificar o decodificador de áudio mas criar um cabeçalho do débito binário. Em alternativa, a informação de *pitch* pode ser determinada no decodificador áudio, por exemplo, no domínio de excitação, isto é, baseada em um sinal de excitação de domínio de tempo. Por exemplo, o sinal de excitação de domínio de tempo derivado de um quadro de áudio anterior, previamente descodificado pode ser avaliado para identificar a informação de *pitch* a ser utilizada para o fornecimento da informação de áudio de dissimulação de erro.

[0203] 5.3 Extrapolação de Excitação ou Criação da Parte Harmónica

[0204] A excitação (por exemplo, o sinal de excitação de domínio de tempo) obtida do quadro anterior (apenas calculada para quadro perdido ou já salva no quadro perdido anterior para perda de quadros múltiplos) é utilizada para construir a parte harmónica (também designada como componente determinístico ou componente aproximadamente periódico) na excitação (por

exemplo, no sinal de entrada da síntese LPC) através da cópia anterior do ciclo de *pitch* quantas vezes necessárias para obter um e metade de um quadro. Para salvar a complexidade podemos também criar um e uma metade de um quadro apenas para a primeira perda de quadro e depois deslocar o processamento para subsequente perda de quadro através de um quadro e criar apenas um quadro cada. Depois temos sempre acesso a metade de um quadro de sobreposição.

[0205] No caso da primeira perda de quadro após um bom quadro (isto é, quadro descodificado de modo adequado), o primeiro ciclo de *pitch* (por exemplo, do sinal de excitação de domínio de tempo obtido baseado no último quadro de áudio descodificado de modo adequado que antecede o quadro de áudio perdido) é filtrado de baixa frequência com um filtro dependente da taxa de amostragem (visto que ELD abrange uma combinação de taxa de amostragem mesmo muito abrangente – do núcleo AA-ELD a AAC-ELD com SBR ou SBR de taxa dupla AAC-ELD).

[0206] O *pitch* no sinal vocal está quase sempre a mudar. Por isso, a dissimulação apresentada em cima tende a criar alguns problemas (ou pelo menos distorções) na recuperação porque o *pitch* no final do sinal dissimulado (isto é, no final da informação de áudio de dissimulação de erro) muitas vezes não corresponde ao *pitch* do primeiro quadro bom. Por isso, opcionalmente, alguns modelos tenta-se predizer o *pitch* no final do quadro dissimulado para corresponder o *pitch* no início do quadro de recuperação. Por exemplo, o *pitch* no final de um quadro perdido (considerado como um quadro oculto) é de predição, em que o alvo de predição é definir o *pitch* no final do quadro perdido (quadro oculto) para aproximar o *pitch* no início do primeiro quadro descodificado de modo adequado a seguir a um ou mais quadros perdidos (cujo primeiro quadro descodificado de modo adequado é também designado de “quadro de recuperação”). Para obter ainda melhores resultados, é possível opcionalmente reutilizar algumas ferramentas convencionais e adaptá-las, tal como a ressincronização de Predição de *Pitch* e de Impulso. Para detalhes, é

feita referência, por exemplo, à referência [6] e [7].

[0207] Se uma predição a longo prazo (LTP) for utilizada em um codec de domínio de frequência, é possível utilizar o desfasamento como a informação de início sobre o *pitch*. Contudo, nalguns modelos, é também desejável ter uma melhor granularidade para poder seguir melhor o contorno do *pitch*. Desse modo, é preferível efetuar uma pesquisa de *pitch* no início e no final do último quadro bom (descodificado de modo adequado). Para adaptar o sinal ao *pitch* móvel, é desejável utilizar uma ressincronização de impulso, presente no estado da técnica.

[0208] 5.4 Ganho de Pitch

[0209] Nalguns modelos, é preferível aplicar um ganho na excitação obtida anteriormente de modo a atingir um nível desejado. O “ganho do *pitch*” (por exemplo, o ganho do componente determinístico do sinal de excitação de domínio de tempo, isto é, o ganho aplicado a um sinal de excitação de domínio de tempo derivado de um quadro de áudio previamente descodificado, de modo a obter o sinal de entrada da síntese LPC), pode, por exemplo, ser obtido através de uma correlação normalizada no domínio de tempo no final do último quadro bom (por exemplo, descodificado de modo adequado). O comprimento da correlação pode ser equivalente ao comprimento de dois subquadros, ou pode ser alterado de modo adaptável. O atraso é equivalente ao desfasamento de *pitch* utilizado para a criação da parte harmónica. Podemos também opcionalmente efetuar o cálculo do ganho apenas no primeiro quadro perdido e então depois aplicar um desvanecimento (ganho reduzido) para a seguinte perda de quadro consecutivo.

[0210] O “ganho do *pitch*” irá determinar a quantidade de tonalidade (ou a quantidade de componentes de sinal determinísticos, pelo menos aproximadamente periódicos) que irá ser criada. Contudo, é desejável adicionar algum ruído modulado para não ter apenas um tom artificial. Se obtivermos um ganho muito baixo de *pitch* então construímos um sinal que consista apenas em um ruído modulado.

[0211] Para concluir, nalguns casos o sinal de excitação de domínio de tempo obtido, por exemplo, baseado em um quadro de áudio previamente descodificado, é escalonado dependendo do ganho (por exemplo, para obter o sinal de entrada para a análise LPC). Por conseguinte, visto que o sinal de excitação de domínio de tempo determina um componente de sinal determinístico (pelo menos aproximadamente periódico), o ganho pode determinar uma intensidade relativa dos ditos componentes de sinal determinístico (pelo menos aproximadamente periódico) na informação de áudio de dissimulação de erro. Além disso, a informação de áudio de dissimulação de erro pode ser baseada em um ruído, também modulado pela síntese LPC, de modo que a energia total da informação de áudio de dissimulação de erro seja adaptada, pelo menos até um certo grau, a um quadro de áudio descodificado de modo adequado que antecede o quadro de áudio perdido e, idealmente, também a um quadro de áudio descodificado de modo adequado a seguir ao um ou mais quadros de áudio perdidos.

[0212] 5.5 Criação da Parte do Ruído

[0213] Uma “inovação” é criada por um gerador de ruído aleatório. Este ruído é opcionalmente ainda filtrado de alta frequência e opcionalmente pré-enfatizado para quadros vozeados e iniciais. Para a baixa frequência da parte harmónica, este filtro (por exemplo, o filtro de baixa frequência) é dependente da taxa de amostragem. Este ruído (fornecido, por exemplo, por uma geração de ruído 560) será modulado pela LPC (por exemplo, pela síntese LPC 580) para se aproximar o mais possível do ruído de fundo. A característica de alta frequência é também opcionalmente alterada sobre uma perda de quadro consecutiva de modo que com uma certa quantidade de perda de quadros já não há filtragem para apenas obter a banda completa de ruído modulado para obter um ruído aceitável próximo do ruído de fundo.

[0214] Um ganho de inovação (que pode, por exemplo, determinar um ganho do ruído 562 no combinador/atenuador 570, isto é, um ganho utilizando qual o sinal de ruído 562 é incluído no sinal de entrada 572 da síntese LPC) é,

por exemplo, calculado removendo o contributo previamente calculado do *pitch* (se existir) (por exemplo, uma versão escalonada, escalonada utilizando o “ganho do *pitch*”, do sinal de excitação de domínio de tempo obtido baseado no último quadro de áudio descodificado de modo adequado que antecede o quadro de áudio perdido) e efetuando uma correlação do final do último quadro de áudio. Relativamente ao ganho de *pitch*, isto poderia ser opcionalmente efetuado apenas no primeiro quadro perdido e depois desvanecido, mas neste caso o desvanecimento poderia ir até 0 resultando em uma mutação completa ou a um nível de ruído estimado presente no fundo. O comprimento da correlação é, por exemplo, equivalente ao comprimento de dois subquadros e o atraso é equivalente ao desfasamento de *pitch* utilizado para a criação da parte harmónica.

[0215] Opcionalmente, este ganho é também multiplicado por (1-“ganho do *pitch*”) para aplicar outro ganho no ruído para atingir a energia que falta se o ganho de *pitch* não o for um. Opcionalmente, este ganho é também multiplicado por um fator de ruído. Este fator de ruído provém, por exemplo, do quadro validade anterior (por exemplo, do último quadro de áudio descodificado de modo adequado que antecede o quadro de áudio perdido).

[0216] 5.6 Desvanecimento

[0217] O desvanecimento é maioritariamente utilizado para perda de múltiplos quadros. Contudo, o desvanecimento pode ser também utilizado no caso de apenas um único quadro de áudio ser perdido.

[0218] No caso de uma perda de múltiplos quadros, os parâmetros LPC não são novamente calculados. Ou o último calculado é mantido, ou a dissimulação LPC é executada convergindo a uma modulação de fundo. Neste caso, a periodicidade do sinal é convergida para zero. Por exemplo, o sinal de excitação de domínio de tempo 502 obtido baseado em um ou mais quadros de áudio que antecedem um quadro de áudio perdido utiliza ainda um ganho que é gradualmente reduzido ao longo do tempo enquanto o sinal de ruído 562 é mantido constante ou escalonado com um ganho que aumenta gradualmente

ao longo do tempo, de modo que a ponderação relativa do sinal de excitação de domínio de tempo 552 seja reduzida ao longo do tempo quando comparado a ponderação relativa do sinal de ruído 562. Consequentemente, o sinal de entrada 572 da síntese LPC 580 está cada vez mais parecido com o “tipo ruído”. Consequentemente, a “periodicidade” (ou, mais precisamente, o componente determinístico, ou pelo menos aproximadamente periódico do sinal de entrada 582 da síntese LPC 580) é reduzida ao longo do tempo.

[0219] A velocidade da convergência de acordo com a qual a periodicidade do sinal 572, e/ou a periodicidade do sinal 582, é convergida para 0 é dependente dos parâmetros do último quadro corretamente recebido (ou descodificado de modo adequado) e/ou o número de quadros consecutivos eliminados, e é controlada pelo fator de atenuação, α . O fator, α , é ainda dependente da estabilidade do filtro LP. Opcionalmente, é possível alterar o fator α na razão com o comprimento de *pitch*. Se o *pitch* (por exemplo, um comprimento de período associado ao *pitch*) for mesmo longo, então mantemos α “normal”, mas se o *pitch* for mesmo curto, é tipicamente necessário copiar muitas vezes a mesma parte da excitação passada. Isto irá rapidamente parecer muito artificial, e desse modo é preferível desvanecer mais depressa este sinal.

[0220] Ainda opcionalmente, se disponível, podemos ter em consideração a saída de predição de *pitch*. Se um *pitch* for de predição, significa que o *pitch* foi já alterado no quadro anterior e então quantos mais quadros perdemos mais distantes ficamos da verdade. Desse modo, é preferível acelerar um pouco o desvanecimento da parte tonal neste caso.

[0221] Se a predição de *pitch* falhou devido ao *pitch* ter alterado muito, significa que os valores de *pitch* não são fiáveis de todo ou que o sinal é realmente imprevisível. Desse modo, mais uma vez, é preferível desvanecer rapidamente (por exemplo, para desvanecer rapidamente o sinal de excitação de domínio de tempo 552 obtido baseado em um ou mais quadros de áudio descodificados de modo adequado que antecedem o um ou mais quadros de

áudio perdidos).

[0222] 5.7 Síntese LPC

[0223] Para regressar ao domínio de tempo, é preferível executar uma síntese LPC 580 com a totalização das duas excitações (parte tonal e parte do ruído) seguida por uma desacentuação. Por outras palavras, é preferível executar a síntese LPC 580 baseada em uma combinação ponderada de um sinal de excitação de domínio de tempo 552 obtido baseado em um ou mais quadros de áudio descodificados de modo adequado que antecedem o quadro de áudio perdido (parte tonal) e o sinal de ruído 562 (parte do ruído). Tal como referido em cima, o sinal de excitação de domínio de tempo 552 pode ser modificado quando comparado com o sinal de excitação de domínio de tempo 532 obtido pela síntese LPC 530 (para além dos coeficientes LPC que descrevem uma característica do filtro de síntese LPC utilizado para a síntese LPC 580). Por exemplo, o sinal de excitação de domínio de tempo 552 pode ser uma cópia escalonada no tempo do sinal de excitação de domínio de tempo 532 obtido pela síntese LPC 530, em que o escalonamento no tempo pode ser utilizado para adaptar o *pitch* do sinal de excitação de domínio de tempo 552 para um *pitch* desejado.

[0224] 5.8 Sobreposição e Adição

[0225] No caso de um codec por transformação apenas, para obter a melhor sobreposição-adição criamos um sinal artificial para metade de um quadro mais do que o quadro oculto e criamos nele um efeito cascata artificial. Contudo, diferentes conceitos de sobreposição-adição podem ser aplicados.

[0226] No contexto de AAC ou TCX regular, uma sobreposição-e-adição é aplicada entre a metade extra do quadro proveniente da dissimulação e a primeira parte do primeiro quadro bom (pode ser metade ou menos para janelas de atraso mais baixo como AAC-LD).

[0227] No caso especial de ELD (atraso extra baixo), para o primeiro quadro perdido, é preferível executar a análise três vezes para obter o contributo adequado das últimas três janelas e depois para o primeiro quadro

de dissimulação e todos os quais se seguem a análise é executada mais uma vez. Depois uma síntese ELD é feita para regressar ao domínio de tempo com toda a memória adequada para o quadro seguinte no domínio MDCT.

[0228] Para concluir, o sinal de entrada 572 da síntese LPC 580 (e/ou o sinal de excitação de domínio de tempo 552) pode ser fornecido para uma duração temporal mais longa do que uma duração de um quadro de áudio perdido. Por conseguinte, o sinal de saída 582 da síntese LPC 580 pode ser também fornecido para um período de tempo mais longo do que um quadro de áudio perdido.

[0229] Por conseguinte, uma sobreposição-e-adição pode ser efetuada entre a informação de áudio de dissimulação de erro (consequentemente obtida para um período de tempo mais longo do que uma extensão temporal do quadro de áudio perdido) e uma informação de áudio decodificada fornecida para um quadro de áudio descodificado de modo adequado a seguir a um ou mais quadros de áudio perdidos.

[0230] Resumindo, a dissimulação de erro 500 é bem adaptada ao caso no qual os quadros de áudio são codificados no domínio de frequência. Mesmo que os quadros de áudio sejam codificados no domínio de frequência, o fornecimento da informação de áudio de dissimulação de erro é efetuado baseado em um sinal de excitação de domínio de tempo. Diferentes modificações são aplicadas ao sinal de excitação de domínio de tempo obtido baseado em um ou mais quadros de áudio descodificados de modo adequado que antecedem um quadro de áudio perdido. Por exemplo, o sinal de excitação de domínio de tempo fornecido pela síntese LPC 530 é adaptado às alterações de *pitch*, por exemplo, utilizando um escalonamento de tempo. Além disso, o sinal de excitação de domínio de tempo fornecido pela análise LPC 530 é também modificado por um escalonamento (aplicação de um ganho), em que um desvanecimento do componente determinístico (ou tonal, ou pelo menos aproximadamente periódico) pode ser efetuado pelo escalador/atenuador 570, de modo que o sinal de entrada 572 da síntese LPC 580 compreenda ambos

um componente derivado do sinal de excitação de domínio de tempo obtido pela análise LPC e um componente de ruído baseado no sinal de ruído 562. O componente determinístico do sinal de entrada 572 da síntese LPC 580 é, contudo, tipicamente modificado (por exemplo, escalonado no tempo e/ou à escala de amplitude) relativamente ao sinal de excitação de domínio de tempo fornecido pela análise LPC 530.

[0231] Assim, o sinal de excitação de domínio de tempo pode ser adaptado às necessidades, e uma impressão de audição não natural é evitada.

[0232] 6. Dissimulação de Domínio de Tempo de Acordo com a Fig. 6

[0233] A Fig. 6 ilustra um esquema de blocos de uma dissimulação de domínio de tempo que pode ser utilizada para um codec de comutação. Por exemplo, a dissimulação de domínio de tempo 600 de acordo com a Fig. 6 pode, por exemplo, substituir a dissimulação de erro 240 ou substituir a dissimulação de erro 480.

[0234] Além disso, dever-se-á ter em atenção que o modelo de acordo com a Fig. 6 abrange o contexto (pode ser utilizado no contexto) de um codec de comutação utilizando domínio de tempo e de frequência combinados, tais como USAC (MPEG-D/MPEG-H) ou EVS (3GPP). Por outras palavras, a dissimulação de domínio de tempo 600 pode ser utilizada em decodificadores áudio nos quais existe uma comutação entre uma decodificação de domínio de frequência e uma decodificação de tempo (ou, de modo equivalente, uma decodificação baseada no coeficiente de predição linear).

[0235] Contudo, dever-se-á ter em atenção que a dissimulação de erro 600 de acordo com a Fig. 6 pode ser também utilizada em decodificadores áudio que simplesmente executam uma decodificação no domínio de tempo (ou de modo equivalente, no domínio do coeficiente de predição linear).

[0236] No caso de um codec de comutação (e mesmo no caso de um codec que simplesmente executa a decodificação no domínio do coeficiente de predição linear) habitualmente temos já o sinal de excitação (por exemplo, o sinal de excitação de domínio de tempo) proveniente de um quadro anterior

(por exemplo, um quadro de áudio descodificado de modo adequado que antecede um quadro de áudio perdido). Caso contrário (por exemplo, se o sinal de excitação de domínio de tempo não estiver disponível), é possível proceder tal como explicado no modelo de acordo com a Fig. 5, isto é, executar uma análise LPC.

[0237] Se o quadro anterior for do tipo ACELP, também temos já a informação de *pitch* dos subquadros no último quadro. Se o último quadro foi TCX (excitação codificada por transformação) com LTP (predição a longo prazo) temos também a informação de desfasamento proveniente da predição a longo prazo. E se o último quadro foi no domínio de frequência sem predição a longo prazo(LTP) então a pesquisa de *pitch* é de preferência feita diretamente no domínio de excitação (por exemplo, baseado em um sinal de excitação de domínio de tempo fornecido por uma análise LPC).

[0238] Se o decodificador utiliza já alguns parâmetros LPC no domínio de tempo, reutilizamo-los e extrapolamos um novo conjunto de parâmetros LPC. A extração dos parâmetros LPC é baseada no LPC anterior, por exemplo a média dos últimos três quadros e (opcionalmente) a forma LPC derivada durante a estimativa de ruído DTX se DTX (transmissão descontínua) existir no codec.

[0239] Toda a dissimulação é feita no domínio de excitação para obter uma transição mais suave entre os quadros consecutivos.

[0240] A seguir, a dissimulação de erro 600 de acordo com a Fig. 6 será descrita com mais detalhe.

[0241] A dissimulação de erro 600 recebe uma excitação anterior 610 e uma informação de *pitch* anterior 640. Além disso, a dissimulação de erro 600 fornece uma informação de áudio de dissimulação de erro 612.

[0242] Dever-se-á ter em atenção que a excitação anterior 610 recebida pela dissimulação de erro 600 pode, por exemplo, corresponder à saída 532 da análise LPC 530. Além disso, a informação de *pitch* anterior 640 pode, por exemplo, corresponder à informação de saída 542 da pesquisa de *pitch* 540.

[0243] A dissimulação de erro 600 compreende ainda uma extração 650, que pode corresponder à extração 550, de modo que referência é feita à discussão em cima.

[0244] Além disso, a dissimulação de erro compreende um gerador de ruído 660, que pode corresponder ao gerador de ruído 560, de modo que referência é feita à discussão em cima.

[0245] A extração 650 fornece um sinal de excitação de domínio de tempo extraído 652, que pode corresponder ao sinal de excitação de domínio de tempo extraído 552. O gerador de ruído 660 fornece um sinal de ruído 662, que corresponde ao sinal de ruído 562.

[0246] A dissimulação de erro 600 compreende também um combinador/atenuador 670, que recebe o sinal de excitação de domínio de tempo extraído 652 e o sinal de ruído 662 e fornece, baseado nisso, um sinal de entrada 672 para uma síntese LPC 680, em que a síntese LPC 680 pode corresponder à síntese LPC 580, de modo que as explicações em cima também se aplicam. A síntese LPC 680 fornece um sinal de áudio de domínio de tempo 682, que pode corresponder ao sinal de áudio de domínio de tempo 582. A dissimulação de erro compreende também (opcionalmente) uma desacentuação 684, que pode corresponder à desacentuação 584 e que fornece um sinal de áudio de domínio de tempo de uma dissimulação de erro desacentuada 686. A dissimulação de erro 600 opcionalmente compreende uma sobreposição e adição 690, que pode corresponder à sobreposição e adição 590. Contudo, as explicações em cima relativas à sobreposição e adição 590 também se aplicam à sobreposição e adição 690. Por outras palavras, a sobreposição e adição 690 pode também ser substituída pela sobreposição e adição global do decodificador áudio, de modo que o sinal de saída 682 da síntese LPC ou o sinal de saída 686 da desacentuação possa ser considerado como a informação de áudio de dissimulação de erro.

[0247] Para concluir, a dissimulação de erro 600 diverge significativamente da dissimulação de erro 500 na medida em que a dissimulação de erro 600

obtém diretamente a informação de excitação anterior 610 e a informação de *pitch* anterior 640 diretamente de um ou mais quadros de áudio descodificados de modo adequado sem necessidade de executar uma análise LPC e/ou uma análise de *pitch*. Contudo, dever-se-á ter em atenção que a dissimulação de erro 600 pode, opcionalmente, compreender uma análise LPC e/ou uma análise de *pitch* (pesquisa de *pitch*).

[0248] A seguir, alguns detalhes da dissimulação de erro 600 serão descritos com mais detalhe. Contudo, dever-se-á ter em atenção que detalhes específicos deverão ser considerados como exemplos, em vez de como características essenciais.

[0249] 6.1 Pitch Anterior da Pesquisa de Pitch

[0250] Existem diferentes abordagens para obter o *pitch* a ser utilizado para construir o novo sinal.

[0251] No contexto do codec utilizando filtro LTP, como AAC-LTP, se o último quadro (que antecede o quadro perdido) foi AAC com LTP, temos uma informação de *pitch* proveniente do último desfasamento de pitch LTP e respetivo ganho. Neste caso utilizamos o ganho para decidir se queremos construir uma parte harmónica no sinal, ou não. Por exemplo, se o ganho LTP for superior a 0,6 então utilizamos a informação LTP para construir a parte harmónica.

[0252] Se não temos qualquer informação de *pitch* disponível do quadro anterior, então existem, por exemplo, duas outras soluções.

[0253] Uma solução é efetuar uma pesquisa de *pitch* no codificador e transmitir no fluxo de bits o desfasamento de *pitch* e o ganho. Isto é idêntico à predição a longo prazo (LTP), mas não aplicamos qualquer filtragem (também nenhuma filtragem LTP no canal limpo).

[0254] Outra solução é efetuar uma pesquisa de *pitch* no decodificador. A pesquisa de *pitch* AMR-WB no caso de TCX é efetuada no domínio FFT. No TCX por exemplo, utilizamos o domínio MDCT, depois saltamos as fases. Por isso, a pesquisa de *pitch* é efetuada diretamente no domínio de excitação (por

exemplo, baseado no sinal de excitação de domínio de tempo utilizado como a entrada da síntese LPC, ou utilizado para derivar a entrada para a síntese LPC) em um modelo preferido. Isto tipicamente oferece melhores resultados do que efetuar a pesquisa de *pitch* no domínio de síntese (por exemplo, baseado em um sinal de áudio de domínio de tempo totalmente descodificado).

[0255] A pesquisa de *pitch* no domínio de excitação (por exemplo, baseado no sinal de excitação de domínio de tempo) é efetuada em primeiro lugar com um ciclo fechado através de uma correlação cruzada normalizada. Depois, opcionalmente, a pesquisa de *pitch* pode ser refinada ao efetuar-se uma pesquisa de ciclo fechado em redor de um *pitch* de ciclo aberto com um certo delta.

[0256] Em implementações preferidas, não consideramos simplesmente um valor máximo da correlação. Se tivermos uma informação de *pitch* de um quadro anterior não sujeito a erro, então selecionamos o *pitch* que corresponde a um dos cinco valores mais elevados no domínio da correlação cruzada normalizada mas o mais próximo do *pitch* do quadro anterior. Depois, é também verificado que a descoberta máxima não é um máximo errado devido à limitação da janela.

[0257] Para concluir, existem diferentes conceitos para determinar o *pitch*, e que é computacionalmente eficiente considerar o *pitch* anterior (isto é, o *pitch* associado a um quadro de áudio previamente descodificado). Em alternativa, a informação de *pitch* pode ser transmitida de um codificador áudio a um decodificador áudio. Como outra alternativa, uma pesquisa de *pitch* pode ser executada no lado do decodificador áudio, em que a determinação de *pitch* é de preferência executada baseada no sinal de excitação de domínio de tempo (isto é, no domínio de excitação).

[0258] Uma pesquisa de *pitch* de duas fases compreendendo uma pesquisa de ciclo aberto e uma pesquisa de ciclo fechado pode ser efetuada de modo a obter uma informação de *pitch* especialmente fiável e precisa. Em alternativa, ou além disso, uma informação de *pitch* proveniente de um quadro de áudio

previamente descodificado pode ser utilizada de modo a assegurar que a pesquisa de *pitch* forneça um resultado credível.

[0259] 6.2 Extrapolação da Excitação ou Criação da Parte Harmónica

[0260] A excitação (por exemplo, sob a forma de um sinal de excitação de domínio de tempo) obtida de um quadro anterior (apenas calculado para o quadro perdido ou salvo já no quadro perdido anterior para perda de múltiplos quadros) é utilizada para construir a parte harmónica na excitação (por exemplo, o sinal de excitação de domínio de tempo extrapolado 662) copiando o último ciclo de *pitch* (por exemplo, uma parte do sinal de excitação de domínio de tempo 610, uma duração temporal a qual é igual a um período de duração do *pitch*) quantas vezes as necessárias para obter, por exemplo, um e metade do quadro (perdido).

[0261] Para se obter ainda melhores resultados, é opcionalmente possível reutilizar algumas ferramentas conhecidas do estado da técnica e adaptá-las. Para detalhes, referência é feita, por exemplo, à referência [6] e [7].

[0262] Foi verificado que o *pitch* em um sinal vozeado está quase sempre a mudar. Foi verificado que, por isso, a dissimulação apresentada em cima tem tendência a criar mais problemas na recuperação porque o *pitch* no final do sinal ocultado muitas vezes não corresponde ao *pitch* do primeiro quadro bom. Desse modo, opcionalmente, tenta-se predizer o *pitch* no final do quadro ocultado para corresponder ao *pitch* no início do quadro de recuperação. Esta funcionalidade será executada, por exemplo, pela extrapolação 650.

[0263] Se LTP em TCX for utilizada, o desfasamento pode ser utilizado como a informação inicial sobre o *pitch*. Contudo, é desejável ter uma melhor granularidade para poder seguir melhor o contorno do *pitch*. Desse modo, uma pesquisa de *pitch* é opcionalmente efetuada no início e no final do último quadro bom. Para adaptar o sinal ao *pitch* móvel, uma resincronização de impulso, presente no estado da técnica, pode ser utilizada.

[0264] Para concluir, a extrapolação (por exemplo, um sinal de excitação de domínio de tempo associado a, ou obtido baseado nisso, um último quadro de

áudio descodificado de modo adequado que antecede o quadro perdido) pode compreender uma cópia de uma parte de tempo do dito sinal de excitação de domínio de tempo associado a um quadro de áudio anterior, em que a parte de tempo copiada pode ser modificada dependente de um cálculo, ou estimativa, de uma alteração de *pitch* (esperada) durante o quadro de áudio perdido. Diferentes conceitos encontram-se disponíveis para determinarem a alteração de *pitch*.

[0265] 6.3 Gанho do Pitch

[0266] Em um modelo de acordo com a fig. 6, um ganho é aplicado à excitação previamente obtida de modo a atingir um nível desejado. O ganho do *pitch* é obtido, por exemplo, ao efetuar-se uma correlação normalizada no domínio de tempo no final do último quadro bom. Por exemplo, o comprimento da correlação pode ser equivalente ao comprimento de dois quadros e o atraso pode ser equivalente ao desfasamento de *pitch* utilizado para a criação da parte harmónica (por exemplo, para copiar o sinal de excitação de domínio de tempo). Foi verificado que ao efetuar-se o cálculo do ganho no domínio de tempo oferece um ganho muito mais fiável do que fazê-lo no domínio de excitação. As LPC alteram todos os quadros e depois aplicar um ganho, calculado no quadro anterior, em um sinal de excitação que irá ser processado por um outro conjunto LPC, não irá produzir uma energia esperada no domínio de tempo.

[0267] O ganho do *pitch* determina a quantidade de tonalidade que irá ser criada, mas algum ruído modulado irá também ser adicionado para não ter apenas um tom artificial. Se um ganho de *pitch* muito baixo for obtido, então um sinal pode ser construído compreendendo apenas um ruído modulado.

[0268] Para concluir, um ganho aplicado à escala do sinal de excitação de domínio de tempo obtido baseado no quadro anterior (ou um sinal de excitação de domínio de tempo obtido para um quadro previamente descodificado, ou associado ao quadro previamente descodificado) é regulado para desse modo determinar uma ponderação de um componente tonal (ou determinístico, ou

pelo menos aproximadamente periódico) no sinal de entrada da síntese LPC 680 e, consequentemente, na informação de áudio de dissimulação de erro. O dito ganho pode ser determinado baseado em uma correlação, aplicada ao sinal de áudio de domínio de tempo obtido por uma decodificação do quadro previamente descodificado (em que o dito sinal de áudio de domínio de tempo pode ser obtido utilizando uma síntese LPC efetuada no decorrer da decodificação).

[0269] 6.4 Criação da Parte do Ruído

[0270] Uma inovação é criada por um gerador de ruído aleatório 660. Este ruído é adicionalmente filtrado de alta frequência e opcionalmente desacentuado para quadros vozeados e iniciais. A filtragem de alta frequência e a desacentuação, que pode ser efetuada de modo seletivo para quadros vozeados e iniciais, não se encontram ilustradas explicitamente na Fig. 6, mas podem ser efetuadas, por exemplo, no gerador de ruído 660 ou no combinador/atenuador 670.

[0271] O ruído será modulado (por exemplo, após combinação com o sinal de excitação de domínio de tempo 652 obtido pela extração 650) pela LPC para ficar o mais próximo possível do ruído de fundo.

[0272] Por exemplo, o ganho de inovação pode ser calculado removendo o contributo previamente calculado do *pitch* (se existir) e efetuando uma correlação no final do último quadro bom. O comprimento da correlação pode ser equivalente ao comprimento de dois quadros e o atraso pode ser equivalente ao desfasamento de *pitch* utilizado para a criação da parte harmônica.

[0273] Opcionalmente, este ganho pode ser também multiplicado por (ganho de *pitch*-1) para aplicar o maior ganho possível ao ruído para atingir uma energia perdida se o ganho do *pitch* não for um. Opcionalmente, este ganho é também multiplicado por um fator de ruído. Este fator de ruído pode ser proveniente de um quadro previamente válido.

[0274] Para concluir, um componente de ruído da informação de áudio de

dissimulação de erro é obtido através da modulação de ruído fornecida pelo gerador de ruído 660 utilizando a síntese LPC 680 (e, possivelmente, a desacentuação 684). Além disso, uma filtragem de alta frequência adicional e/ou pré-accentuação pode ser aplicada. O ganho do contributo do ruído para o sinal de entrada 672 da síntese LPC 680 (também designada como “ganho de inovação”) pode ser calculado baseado no último quadro de áudio descodificado de modo adequado que antecede o quadro de áudio perdido, em que um componente determinístico (ou pelo menos aproximadamente periódico) pode ser removido do quadro de áudio que antecede o quadro de áudio perdido, e em que uma correlação pode então ser executada para determinar uma intensidade (ou ganho) do componente de ruído no sinal de domínio de tempo do quadro de tempo que antecede o quadro de áudio perdido.

[0275] Opcionalmente, algumas modificações adicionais podem ser aplicadas ao ganho do componente do ruído.

[0276] 6.5 Desvanecimento

[0277] O desvanecimento é maioritariamente utilizado para perda de múltiplos quadros. Contudo, o desvanecimento pode ser também utilizado no caso de apenas um único quadro de áudio ter sido perdido.

[0278] No caso de perda de múltiplos quadros, os parâmetros LPC não são recalculados. Ou o último calculado é mantido ou uma dissimulação LPC é efetuada tal como explicado em cima.

[0279] Uma periodicidade do sinal é convergida para zero. A velocidade da convergência é dependente dos parâmetros do último quadro corretamente recebido (ou corretamente descodificado) e do número de quadros eliminados (ou perdidos) consecutivos, e é controlada por um fator de atenuação, α . O fator, α , é ainda dependente da estabilidade do filtro LP. Opcionalmente, o fator α pode ser alterado na razão com o comprimento de *pitch*. Por exemplo, se o *pitch* for mesmo longo, então α pode ser mantido normal, mas se o *pitch* for mesmo curto, pode ser desejável (ou necessário) copiar muitas vezes a mesma

parte da excitação passada. Visto que foi verificado que isto irá rapidamente parecer demasiado artificial, o sinal é pois rapidamente desvanecido.

[0280] Ainda opcionalmente, é possível ter em consideração a saída de predição de *pitch*. Se um *pitch* for de predição, significa que o *pitch* foi já alterado no quadro anterior e então quantos mais quadros forem perdidos mais distantes ficamos da verdade. Desse modo, é desejável acelerar um pouco o desvanecimento da parte tonal neste caso.

[0281] Se a predição de *pitch* falhou devido ao *pitch* ter alterado muito, significa que os valores de *pitch* não são fiáveis de todo ou que o sinal é realmente imprevisível. Desse modo, mais uma vez devemos desvanecer rapidamente.

[0282] Para concluir, o contributo do sinal de excitação de domínio de tempo extrapolado 652 ao sinal de entrada 672 da síntese LPC 680 é tipicamente reduzido ao longo do tempo. Isto pode ser obtido, por exemplo, reduzindo um valor de ganho, aplicado ao sinal de excitação de domínio de tempo extrapolado 652, ao longo do tempo. A velocidade utilizada para reduzir gradualmente o ganho aplicada à escala o sinal de excitação de domínio de tempo 552 obtido baseado em um ou mais quadros de áudio que antecedem um quadro de áudio perdido (ou um ou mais cópias deste) é regulada dependendo de um ou mais parâmetros de um ou mais quadros de áudio (e/ou dependendo de um número de quadros de áudio consecutivos perdidos). Em especial, o comprimento do *pitch* e/ou a taxa na qual o *pitch* altera ao longo do tempo, e/ou a questão se uma predição de *pitch* falha ou é bem-sucedida, pode ser utilizado para regular a dita velocidade.

[0283] 6.6 Síntese LPC

[0284] Para regressar ao domínio de tempo, uma síntese LPC 680 é efetuada com a totalização (ou regra geral, combinação ponderada) das duas excitações (parte tonal e parte do ruído) seguida por uma desacentuação. Por outras palavras, o resultado da combinação (desvanecimento) ponderada do sinal de excitação de domínio de tempo extrapolado 652 e do sinal de ruído

662, forma um sinal de excitação de domínio de tempo combinado e entra em uma síntese LPC 680, que pode, por exemplo, efetuar uma filtragem de síntese baseada no dito sinal de excitação de domínio de tempo combinado 672 dependendo dos coeficientes LPC que descrevem o filtro de síntese.

[0285] 6.7 Sobreposição e Adição

[0286] Visto que não é sabido durante a dissimulação que será o modo do próximo quadro (por exemplo, ACELP, TCX ou FD), é preferível preparar diferentes sobreposições com antecedência. Para obter a melhor sobreposição e adição se o próximo quadro se encontrar no domínio de transformação (TCX ou FD) um sinal artificial (por exemplo, uma informação de áudio de dissimulação de erro) pode, por exemplo, ser criado para metade de um quadro mais do que o quadro (perdido) oculto. Além disso, o efeito cascata artificial pode ser criado nele (em que o efeito cascate artificial pode, por exemplo, ser adaptado à sobreposição e adição MDCT).

[0287] Para obter uma boa sobreposição e adição e nenhuma descontinuidade no quadro futuro no domínio de tempo (ACELP), executamos tal como em cima, mas sem o efeito cascata, para poder aplicar-se longas janelas de sobreposição e adição ou se quisermos utilizar uma janela quadrada, a resposta de entrada zero (ZIR) é calculada no final da síntese tampão.

[0288] Para concluir, em um decodificador áudio de comutação (que pode, por exemplo, trocar entre uma decodificação ACELP, uma decodificação TCX e uma decodificação de domínio de frequência (decodificação FD)), uma sobreposição e adição pode ser executada entre a informação de áudio de dissimulação de erro fornecida em primeiro lugar para um quadro de áudio perdido, mas também para uma certa parte de tempo a seguir ao quadro de áudio perdido, e a informação de áudio decodificada fornecida para o primeiro quadro de áudio previamente decodificado a seguir a uma sequência de um ou mais quadros de áudio perdidos. De modo a obter uma sobreposição e adição adequada mesmo para modos de decodificação que trazem um efeito

cascate de domínio de tempo em uma transição entre quadros de áudio subsequentes, uma informação de cancelamento de efeito cascata (por exemplo, designado como efeito cascata artificial) pode ser fornecida. Por conseguinte, uma sobreposição e adição entre a informação de áudio de dissimulação de erro e a informação de áudio de domínio de tempo obtida baseada no primeiro quadro de áudio descodificado de modo adequado a seguir a um quadro de áudio perdido, resulta em um cancelamento do efeito cascata.

[0289] Se o primeiro quadro de áudio descodificado de modo adequado a seguir a uma sequência de um ou mais quadros de áudio perdidos é codificado no modo ACELP, uma informação de sobreposição específica pode ser calculada, podendo ser baseada em uma resposta de entrada de zero (ZIR) de um filtro LPC.

[0290] Para concluir, a dissimulação de erro 600 é bem adequada para a utilização em um codec de áudio de comutação. Contudo, a dissimulação de erro 600 pode ser também utilizada no codec de áudio que simplesmente descodifica um conteúdo áudio codificado em um modo TCX ou em um modo ACELP.

[0291] 6.8 Conclusão

[0292] Dever-se-á ter em atenção que uma dissimulação de erro especialmente boa é obtida pelo conceito referido em cima para extrapolar um sinal de excitação de domínio de tempo, para combinar o resultado da extração com um sinal de ruído utilizando um desvanecimento (por exemplo, desvanecimento cruzado) e para executar uma análise LPC baseada em um resultado de um desvanecimento cruzado.

[0293] 7. Decodificador Áudio de Acordo com a Fig. 11

[0294] A Fig. 11 ilustra um esquema de blocos de um decodificador áudio 1100, de acordo com um modelo desta invenção.

[0295] Dever-se-á ter em atenção que o decodificador áudio 1100 pode fazer parte de um decodificador de comutação de áudio. Por exemplo, o

decodificador áudio 1100 pode substituir a trajetória de decodificação de domínio de predição linear 440 no decodificador áudio 400.

[0296] O decodificador áudio 1100 é configurado para receber uma informação de áudio codificada 1110 e para fornecer, baseado nisso, uma informação de áudio decodificada 1112. A informação de áudio codificada 1110 pode, por exemplo, corresponder à informação de áudio codificada 410 e a informação de áudio decodificada 1112 pode, por exemplo, corresponder à informação de áudio decodificada 412.

[0297] O decodificador áudio 1100 comprehende um analisador do fluxo de bits 1120, configurado para extrair uma representação codificada 1122 de um conjunto de coeficientes espectrais e uma representação codificada dos coeficientes de codificação de predição linear 1124 da informação de áudio codificada 1110. Contudo, o analisador do fluxo de bits 1120 pode opcionalmente extrair informação adicional da informação de áudio codificada 1110.

[0298] O decodificador áudio 1100 comprehende também uma decodificação de valor espectral 1130, configurada para fornecer um conjunto de valores espectrais descodificados 1132 baseados nos coeficientes espectrais codificados 1122. Qualquer conceito de decodificação conhecido para decodificação de coeficientes espectrais pode ser utilizado.

[0299] O decodificador áudio 1100 comprehende também uma conversão de coeficientes de codificação de predição linear para fatores de escala 1140 configurada para fornecer um conjunto de fatores de escala 1142 baseado em uma representação codificada 1124 de coeficientes de codificação de predição linear. Por exemplo, a conversão de coeficientes de codificação de predição linear para fatores de escala 1140 pode efetuar uma funcionalidade descrita na norma USAC. Por exemplo, a representação codificada 1124 dos coeficientes de codificação de predição linear pode compreender uma representação polinomial, decodificada e convertida em um conjunto de fatores de escala através da conversão de coeficientes de codificação de predição linear para

fatores de escala 1142.

[0300] O decodificador áudio 1100 comprehende também um escalar 1150, configurado para aplicar os fatores de escala 1142 aos valores espectrais descodificados 1132, para desse modo obter valores espectrais descodificados por escalonamento 1152. Além disso, o decodificador áudio 1100 comprehende, opcionalmente, um processamento 1160, que pode, por exemplo, corresponder ao processamento 366 descrito em cima, em que valores espectrais decodificadores por escalonamento processados 1162 são obtidos por um processamento opcional 1160. O decodificador áudio 1100 comprehende também uma transformação de domínio de frequência para domínio de tempo 1170, configurado para receber valores espectrais descodificados por escalonamento 1152 (que podem corresponder aos valores espectrais descodificados por escalonamento 362), ou os valores espectrais decodificadores por escalonamento processados 1162 (que podem corresponder aos valores espectrais descodificados por escalonamento processados 368) e fornecer, baseado nisso, uma representação de domínio de tempo 1172, que pode corresponder à representação de domínio de tempo 372 descrita em cima. O decodificador áudio 1100 comprehende também um primeiro pós-processamento opcional 1174, e um segundo pós-processamento opcional 1178, que podem, por exemplo, corresponder, pelo menos parcialmente, ao pós-processamento opcional 376 mencionado em cima. Por conseguinte, o decodificador áudio 1100 obtém (opcionalmente) uma versão pós-processada 1179 da representação de áudio de domínio de tempo 1172.

[0301] O decodificador áudio 1100 comprehende também um bloco de dissimulação de erro 1180 configurado para receber uma representação de áudio de domínio de tempo 1172, ou uma versão pós-processada desta, e os coeficientes de codificação de predição linear (em forma codificada ou em forma decodificada) e fornece, baseado nisso, uma informação de áudio de dissimulação de erro 1182.

[0302] O bloco de dissimulação de erro 1180 é configurado para fornecer a

informação de áudio de dissimulação de erro 1182 para dissimular uma perda de um quadro de áudio a seguir a um quadro de áudio codificado em uma representação de domínio de frequência utilizando um sinal de excitação de domínio de tempo, e desse modo é idêntico à dissimulação de erro 480, e também à dissimulação de erro 500 e à dissimulação de erro 600.

[0303] Contudo, o bloco de dissimulação de erro 1180 compreende uma análise LPC 1184, significativamente idêntica à análise LPC 530. Contudo, a análise LPC 1184 pode, opcionalmente, utilizar os coeficientes LPC 1124 para facilitar a análise (quando comparado com a análise LPC 530). A análise LPC 1134 fornece um sinal de excitação de domínio de tempo 1186, significativamente idêntico ao sinal de excitação de domínio de tempo 532 (e também ao sinal de excitação de domínio de tempo 610). Além disso, o bloco de dissimulação de erro 1180 compreende uma dissimulação de erro 1188, que podem por exemplo, executar a funcionalidade dos blocos, 540, 550, 560, 570, 580, 584 da dissimulação de erro 500, ou que pode, por exemplo, executar a funcionalidade dos blocos 640, 650, 660, 670, 680, 684 da dissimulação de erro 600. Contudo, o bloco de dissimulação de erro 1180 diverge ligeiramente da dissimulação de erro 500 e também da dissimulação de erro 600. Por exemplo, o bloco da dissimulação de erro 1180 (compreendendo uma análise LPC 1184) diverge da dissimulação de erro 500 na medida em que os coeficientes LPC (utilizados para a síntese LPC 580) não são determinados pela análise LPC 530, mas não (opcionalmente) recebidos do fluxo de bits. Além disso, o bloco de dissimulação de erro 1188, compreendendo a análise LPC 1184, diverge da dissimulação de erro 600 na medida em que a “excitação passada” 610 é obtida pela análise LPC 1184, em vez de se encontrar diretamente disponível.

[0304] O decodificador áudio 1100 compreende também uma combinação de sinal 1190, configurada para receber a representação de áudio de domínio de tempo 1172, ou uma versão pós-processada desta, e também uma informação de áudio de dissimulação de erro 1182 (naturalmente, para quadros

de áudio posteriores) e combina os ditos sinais, de preferência utilizando uma operação de sobreposição e adição, para desse modo obter a informação de áudio decodificada 1112.

[0305] Para detalhes adicionais, referência é feita às explicações em cima.

[0306] 8. Método de Acordo com a Fig. 9

[0307] A Fig. 9 ilustra um fluxograma de um método para fornecer uma informação de áudio decodificada baseada em uma informação de áudio codificada. O método de acordo com a Fig. 9 compreende o fornecimento 910 de uma informação de áudio de dissimulação de erro para dissimular uma perda de um quadro de áudio a seguir a um quadro de áudio codificado em uma representação de domínio de frequência utilizando um sinal de excitação de domínio de tempo. O método 900 de acordo com a Fig. 9 é baseado nas mesmas considerações que o decodificador áudio de acordo com a Fig. 1. Além disso, dever-se-á ter em atenção que o método 900 pode ser complementado por qualquer uma das características e funcionalidades aqui descritas, individualmente ou combinadas.

[0308] 9. Método de Acordo com a Fig. 10

[0309] A Fig. 10 ilustra um fluxograma de um método para fornecer uma informação de áudio decodificada baseada em uma informação de áudio codificada. O método 1000 compreende o fornecimento 1010 de uma informação de áudio de dissimulação de erro para dissimulação de uma perda de um quadro de áudio, em que um sinal de excitação de domínio de tempo obtido para (ou baseado em) um ou mais quadros de áudio que antecedem um quadro de áudio perdido é modificado de modo a obter informação de áudio de dissimulação de erro.

a. O método 1000 de acordo com a Fig. 10 é baseado nas mesmas considerações que o decodificador áudio mencionado em cima de acordo com a Fig. 2.

[0310] Além disso, dever-se-á ter em atenção que o método de acordo com a fig. 10 pode ser complementado por qualquer uma das características e

funcionalidades aqui descritas, individualmente ou combinadas.

[0311] 10. Observações Adicionais

[0312] Nos modelos descritos em cima, perda de múltiplos quadros pode ser tratada de diferentes maneiras. Por exemplo, se dois ou mais quadros estiverem perdidos, a parte periódica do sinal de excitação de domínio de tempo para o segundo quadro perdido pode derivar de (ou ser igual a) uma cópia da parte tonal do sinal de excitação de domínio de tempo associada ao primeiro quadro perdido. Em alternativa, o sinal de excitação de domínio de tempo para o segundo quadro perdido pode ser baseado em uma análise LPC do sinal de síntese do quadro perdido anterior. Por exemplo, em um codec a LPC pode estar a mudar todos os quadros perdidos, depois tem sentido refazer a análise para todos os quadros perdidos.

[0313] 11. Alternativas de Implementação

[0314] Apesar de alguns aspectos terem sido aqui descritos no contexto de um aparelho, está claro que estes aspectos representam também uma descrição do método respetivo, em que um bloco ou dispositivo corresponde a uma etapa do método ou uma característica de uma etapa do método. Do mesmo modo, aspectos descritos no contexto de uma etapa do método representam também uma descrição de um bloco respetivo ou item ou característica de um aparelho respetivo. Algumas ou todas as etapas do método podem ser executadas por (ou utilizar) um aparelho de hardware, como por exemplo, um microprocessador, um computador programável ou um circuito eletrônico. Nalguns modelos, algumas ou mais das etapas importantes do método podem ser executadas por esse aparelho.

[0315] Dependendo de certos requisitos de implementação, os modelos da invenção podem ser implementados em hardware ou em software. A implementação pode ser efetuada utilizando um suporte de memória digital, por exemplo, um disquete, um DVD, um CD, uma ROM, uma PROM, uma EPROM, uma EEPROM ou uma memória FLASH, dotado de sinais de controlo lidos eletronicamente aí armazenados, cooperando (ou capazes de cooperar) com

um sistema de computador programável de modo que o respetivo método seja executado.

[0316] Algumas versões de acordo com a invenção compreendem um suporte de dados com sinais de controlo de leitura eletrônica, que são capazes de cooperar com um sistema de computador programável, de modo a que seja executado um dos métodos aqui descritos.

[0317] De um modo geral, os modelos da presente invenção podem ser implementados como um produto de programa de computador com um código de programa, sendo o código de programa operativo para executar um dos métodos quando o produto do programa de computador corre em um computador. O código de programa pode, por exemplo, ser guardado em um suporte de leitura em máquina.

[0318] Outros modelos compreendem o programa de computador para executar um dos métodos aqui descritos, guardados em um suporte de leitura em máquina.

[0319] Por outras palavras, um modelo do método da invenção é, por isso, um programa de computador com um código de programa para executar um dos métodos aqui descritos, quando o programa de computador é executado em um computador.

[0320] Outro modelo dos métodos da invenção é, por isso, um suporte de dados (ou um suporte de armazenamento digital ou um suporte de leitura em computador) compreendendo, aí gravados, o programa de computador para executar um dos métodos aqui descritos. O suporte de dados, o suporte de armazenamento digital ou os meios de registo são tipicamente tangíveis e/ou não transitórios.

[0321] Outro modelo do método da invenção é, por isso, um fluxo de dados ou uma sequência de sinais que representam o programa de computador para executar um dos métodos aqui descritos. O fluxo de dados ou a sequência de sinais pode, por exemplo, ser configurado para ser transferido através de uma ligação de comunicação de dados, por exemplo via Internet.

[0322] Outro modelo comprehende um meio de processamento, por exemplo, um computador, ou um dispositivo lógico programável, configurado ou adaptado para executar um dos métodos aqui descritos.

[0323] Outro modelo comprehende um computador com o programa de computador instalado para executar um dos métodos aqui descritos.

[0324] Outro modelo de acordo com a invenção comprehende um aparelho ou um sistema configurado para transferir (por exemplo, eletronicamente ou opcionalmente) um programa de computador para executar um dos métodos aqui descritos para um receptor. O receptor pode, por exemplo, ser um computador, um dispositivo móvel, um dispositivo de memória ou idêntico. O aparelho ou o sistema pode, por exemplo, compreender um servidor de ficheiro para transferir o programa de computador ao receptor.

[0325] Em alguns modelos, pode ser utilizado um dispositivo programável lógico (por exemplo uma rede de portas lógicas programáveis) para executar algumas ou todas as funcionalidades dos métodos aqui descritos. Em alguns modelos, uma rede de portas lógicas programáveis pode cooperar com um microprocessador para executar um dos métodos aqui descritos. De um modo geral, os métodos são preferencialmente executados por qualquer aparelho de hardware.

[0326] O aparelho aqui descrito pode ser implementado utilizando um aparelho de hardware, ou utilizando um computador, ou utilizando uma combinação de um aparelho de hardware e um computador.

[0327] Os métodos aqui descritos podem ser executados utilizando um aparelho de hardware, ou utilizando um computador, ou utilizando uma combinação de um aparelho de hardware e um computador.

[0328] Os modelos acima descritos são meramente ilustrativos para os princípios da presente invenção. Compreende-se que as modificações e variações das disposições e dos detalhes descritos serão evidentes aos profissionais da matéria. Pretende-se, por isso, que seja limitado apenas pelo

âmbito das reivindicações impendentes da patente e não pelos detalhes específicos da descrição e explicação dos modelos aqui constantes.

[0329] 12. Conclusões

[0330] Para concluir, enquanto alguma dissimulação para codecs de domínio por transformação tem sido descrita na área, modelos de acordo com a invenção superam codecs convencionais (ou decodificadores). Modelos de acordo com a invenção utilizam uma alteração de domínio para dissimulação (domínio de frequência para domínio de tempo ou excitação). Por conseguinte, modelos de acordo com a invenção criam uma dissimulação de fala de alta qualidade para decodificadores de domínio por transformação.

[0331] O modo de decodificação por transformação é idêntico ao do USAC (ver, por exemplo, referência [3]). Utiliza a modificação de transformada discreta de coseno (MDCT) como uma transformação e a modulação de ruído espectral é obtida aplicando a envolvente espectral LPC ponderada no domínio de frequência (também conhecido como FDNS “modulação de ruído de domínio de frequência”). Por outras palavras, modelos de acordo com a invenção podem ser utilizados em um decodificador áudio, que utiliza os conceitos de decodificação descritos na norma USAC. Contudo, o conceito de dissimulação de erro aqui divulgado pode ser também utilizado em um decodificador áudio que é do tipo “AAC” ou em qualquer codec (ou decodificador) da família AAC.

[0332] O conceito de acordo com esta invenção aplica-se a um codec de comutação tal como USAC assim como um codec de domínio de frequência pura. Em ambos casos, a dissimulação é efetuada no domínio de tempo ou no domínio de frequência.

[0333] A seguir, algumas vantagens e características da dissimulação de domínio de tempo (ou da dissimulação de domínio de excitação) serão descritas.

[0334] A dissimulação TCX convencional, tal como descrita, por exemplo, tendo como referência as Figs. 7 e 8, também designada substituição de ruído,

não é muito adequada para sinais do tipo fala ou mesmo sinais tonais. Modelos de acordo com a invenção criam uma nova dissimulação para um codec de domínio por transformação aplicado no domínio de tempo (ou domínio de excitação de um decodificador de codificação de predição linear). É idêntica a uma dissimulação do tipo ACELP e aumenta a qualidade da dissimulação. Foi verificado que a informação de *pitch* é vantajosa (ou mesmo exigida, nalguns casos) para uma dissimulação do tipo ACELP. Assim, modelos de acordo com esta invenção são configurados para verificar a credibilidade dos valores de *pitch* para o quadro anterior codificado no domínio de frequência.

[0335] Diferentes partes e detalhes foram explicados em cima, por exemplo baseados nos modelos de acordo com as Figs. 5 e 6.

[0336] Para concluir, modelos de acordo com a invenção criam uma dissimulação de erro que ultrapassa em desempenho as soluções convencionais.

[0337] Bibliografia:

[1] 3GPP, “Audio codec processing functions; Extended Adaptive Multi-Rate – Wideband (AMR-WB+) codec; Transcoding functions,” 2009, 3GPP TS 26.290.

[2] “MDCT-BASED CODER FOR HIGHLY ADAPTIVE SPEECH AND AUDIO CODING”; Guillaume Fuchs & al.; EUSIPCO 2009.

[3] ISO_IEC_DIS_23003-3_(E); Information technology - MPEG audio technologies - Part 3: Unified speech and audio coding.

[4] 3GPP, “General Audio Codec audio processing functions; Enhanced aacPlus general audio codec; Additional decoder tools,” 2009, 3GPP TS 26.402.

[5] “Audio decoder and coding error compensating method”, 2000, EP 1207519 B1.

[6] “Apparatus and method for improved concealment of the adaptive codebook in ACELP-like concealment employing improved pitch lag estimation”, 2014, PCT/EP2014/062589.

[7] "Apparatus and method for improved concealment of the adaptive codebook in ACELP-like concealment employing improved pulse resynchronization", 2014, PCT/EP2014/062578.

REIVINDICAÇÕES

1. Decodificador de áudio (200;400) para fornecer uma informação de áudio descodificada (220;412) baseada em informação de áudio codificada (210;410), caracterizado por o decodificador de áudio comprehende:

uma dissimulação de erro (240;480;600) configurada para fornecer uma informação de áudio de dissimulação de erro (242;482;612) para dissimular uma perda de um quadro de áudio,

em que a dissimulação de erro está configurada para modificar um sinal de excitação no domínio de tempo (452,456;610) obtido para um ou mais quadros de áudio que antecedem um quadro de áudio perdido, de modo a obter a informação de áudio de dissimulação de erro;

em que a dissimulação de erro (240;480,600) está configurada para regular uma velocidade utilizada para reduzir gradualmente um ganho aplicado para escalonamento do sinal de excitação no domínio de tempo (452,456;610) obtido para um ou mais quadros de áudio que antecedem um quadro de áudio perdido, ou uma ou mais cópias deste, dependendo de um comprimento de um período de *pitch* do sinal de excitação no domínio de tempo, de modo que um componente determinístico de entrada do sinal de excitação no domínio de tempo (672) na síntese LPC (680) seja desvanecido mais rapidamente para os sinais com um comprimento mais curto do período de *pitch* quando comparado aos sinais com um comprimento maior do período de *pitch*.

2. Decodificador de áudio (200;400) para fornecer uma informação de áudio descodificada (220;412) baseada em informação de áudio codificada (210;410), caracterizado por o decodificador de áudio comprehende:

uma dissimulação de erro (240;480;600) configurada para fornecer uma informação de áudio de dissimulação de erro (242;482;612) para dissimular uma perda de um quadro de áudio,

em que a dissimulação de erro está configurada para modificar um sinal de excitação no domínio de tempo (452,456;610) obtido para um ou mais quadros de áudio que antecedem um quadro de áudio perdido, de modo a

obter a informação de áudio de dissimulação de erro;

em que a dissimulação de erro (240;480,600) está configurada para regular uma velocidade utilizada para reduzir gradualmente um ganho aplicado para escalonamento do sinal de excitação no domínio de tempo (452,456;610) obtido para um ou mais quadros de áudio que antecedem um quadro de áudio perdido, ou uma ou mais cópias deste, dependendo de um resultado de uma análise de *pitch* ou uma predição de *pitch*,

de modo que um componente determinístico de uma entrada do sinal de excitação no domínio de tempo (572) na síntese LPC (580) seja desvanecido mais rapidamente para os sinais com uma alteração de *pitch* maior por unidade de tempo quando comparado aos sinais com uma alteração de *pitch* mais pequena por unidade de tempo, e/ou

de modo que um componente determinístico de uma entrada do sinal de excitação no domínio de tempo (572) na síntese LPC (580) seja desvanecido mais rapidamente para os sinais para os quais uma predição de *pitch* falha quando comparado aos sinais para os quais a predição de *pitch* é bem-sucedida.

3. Método (1000) para fornecer uma informação de áudio descodificada baseada em informação de áudio codificada, caracterizado por o método compreender:

fornecimento (1010) de uma informação de áudio de dissimulação de erro para dissimular uma perda de um quadro de áudio,

em que um sinal de excitação no domínio de tempo obtido baseado em um ou mais quadros de áudio que antecedem um quadro de áudio perdido é modificado de modo a obter a informação de áudio de dissimulação de erro;

em que o método compreende a regulação da velocidade utilizada para reduzir gradualmente um ganho aplicado ao escalonamento do sinal de excitação no domínio de tempo (452,456;610) obtido para um ou mais quadros de áudio que antecedem um quadro de áudio perdido, ou a uma ou mais cópias deste, dependendo de um comprimento de um período de *pitch* do sinal

de excitação no domínio de tempo, de modo que o componente determinístico da entrada do sinal de excitação no domínio de tempo (672) na síntese (680) seja desvanecida mais rapidamente para sinais com um comprimento mais curto do período de *pitch* quando comparado aos sinais com um comprimento maior do período de *pitch*.

4. Método (1000) para fornecer uma informação de áudio descodificada baseada em informação de áudio codificada, caracterizado por o método compreender:

fornece (1010) de uma informação de áudio de dissimulação de erro para dissimular uma perda de um quadro de áudio,

em que um sinal de excitação no domínio de tempo obtido baseado em um ou mais quadros de áudio que antecedem um quadro de áudio perdido é modificado de modo a obter a informação de áudio de dissimulação de erro;

em que o método compreende a regulação da velocidade utilizada para reduzir gradualmente um ganho aplicado ao escalonamento do sinal de excitação no domínio de tempo (452,456;610) obtido para um ou mais quadros de áudio que antecedem um quadro de áudio perdido, ou a uma ou mais cópias deste, dependendo de um resultado de uma análise de *pitch* ou uma predição de *pitch*;

de modo que um componente determinístico de uma entrada do sinal de excitação no domínio de tempo (572) na síntese LPC (580) seja desvanecido mais rapidamente para os sinais com uma alteração de *pitch* maior por unidade de tempo quando comparado aos sinais com uma menor alteração de *pitch* por unidade de tempo, e/ou

de modo que um componente determinístico de uma entrada do sinal de excitação no domínio de tempo (572) na síntese LPC (580) seja desvanecido mais rapidamente para os sinais para os quais uma predição de *pitch* falha quando comparado aos sinais para os quais a predição de *pitch* é bem-sucedida.

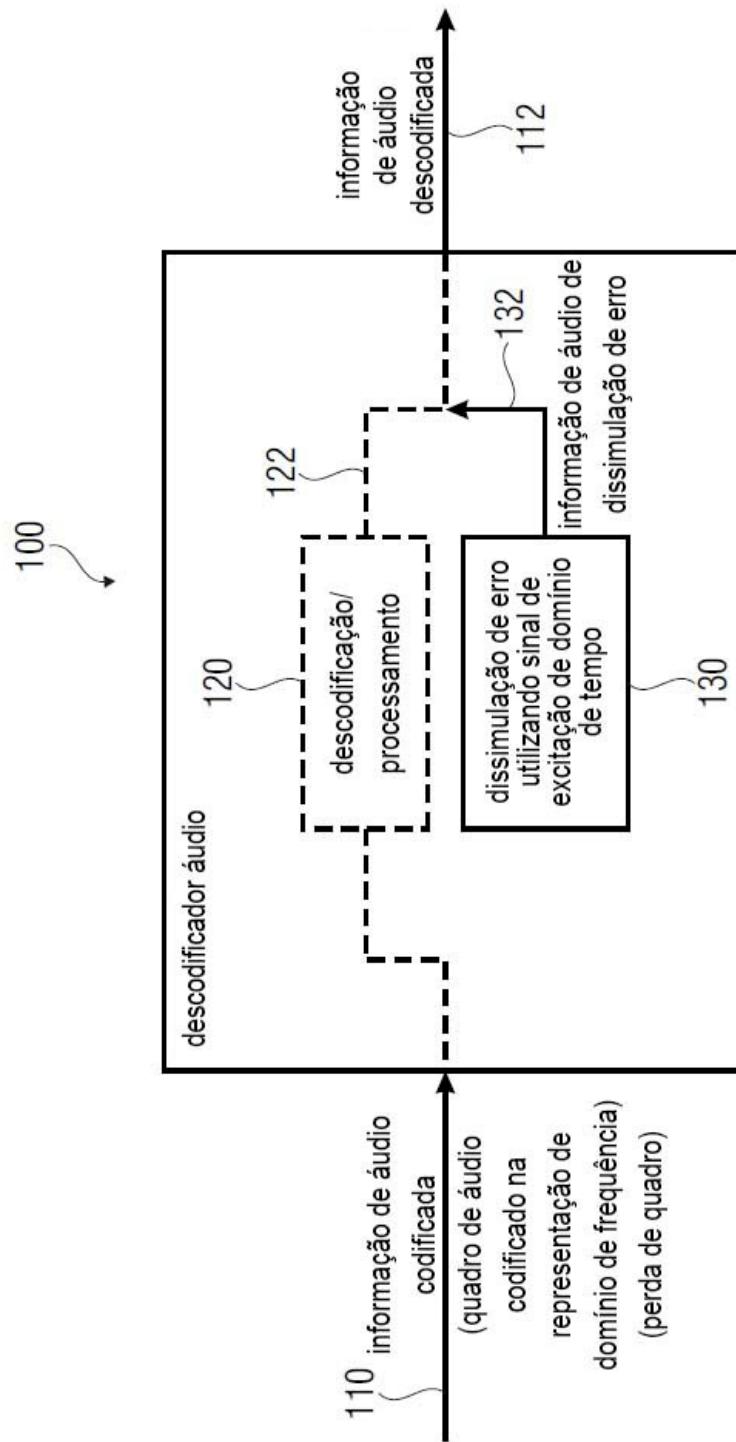


FIG 1

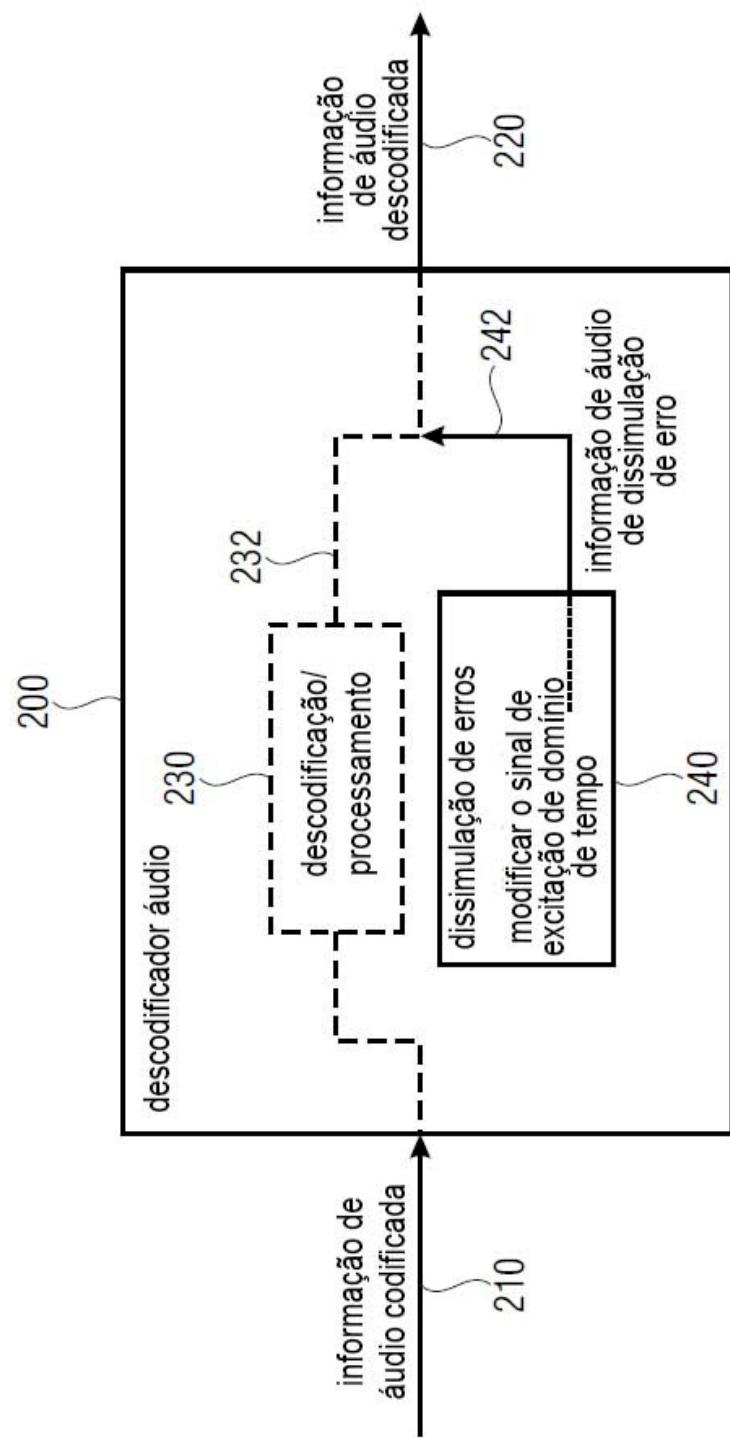


FIG 2

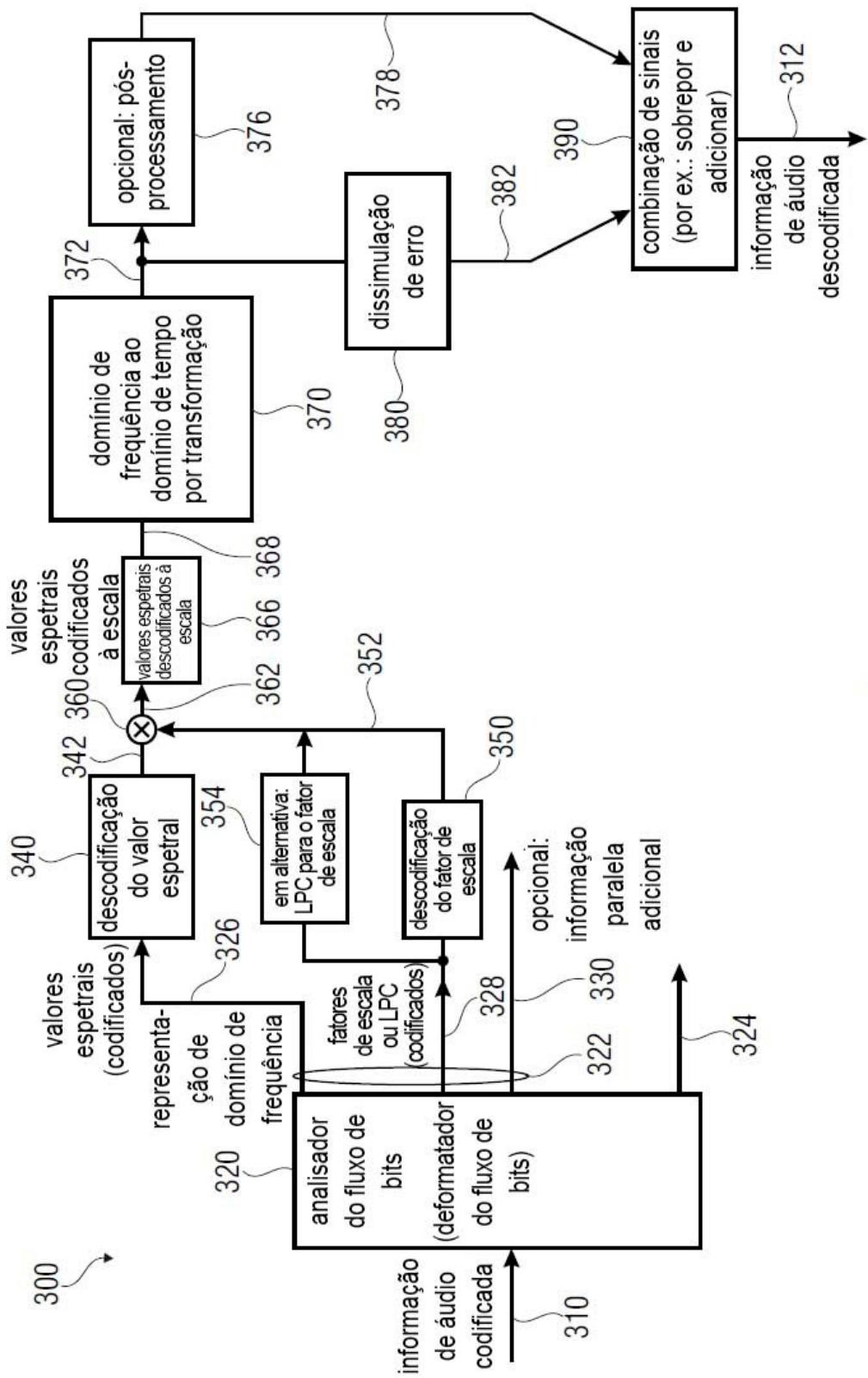


FIG 3

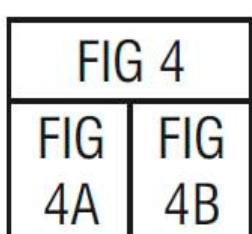
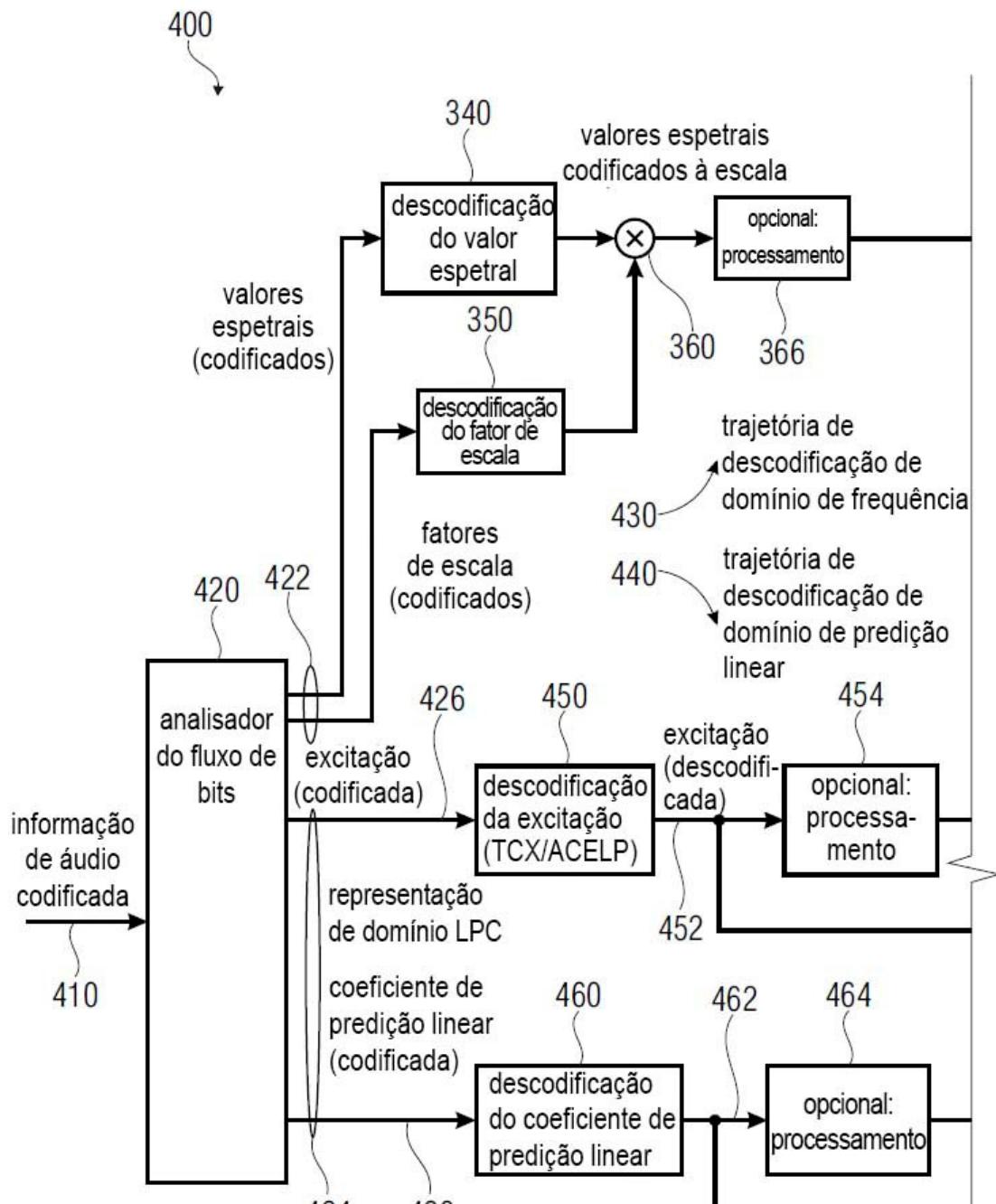


FIG 4A

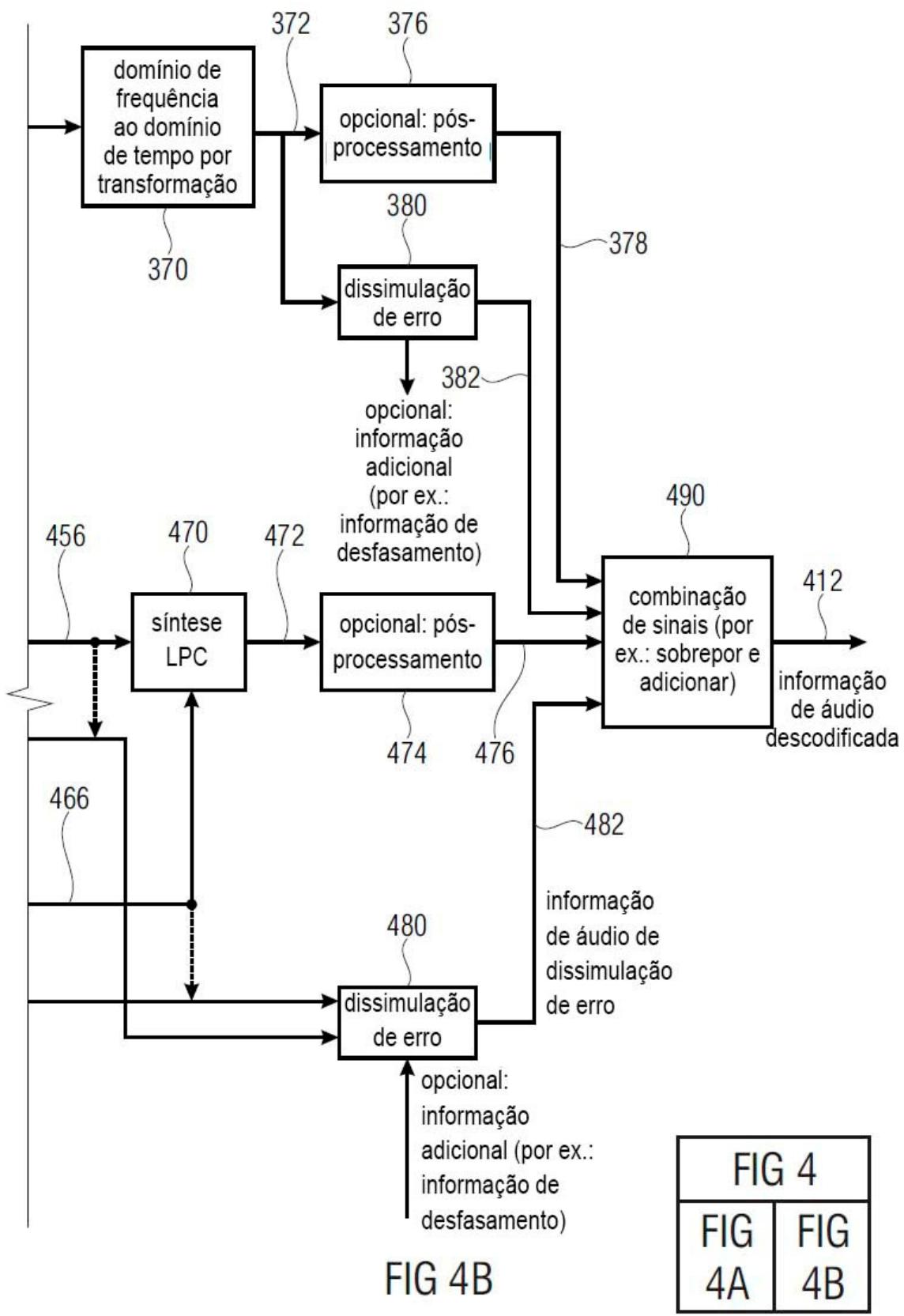
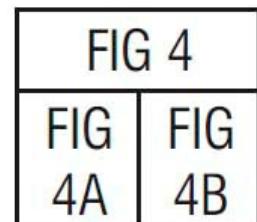
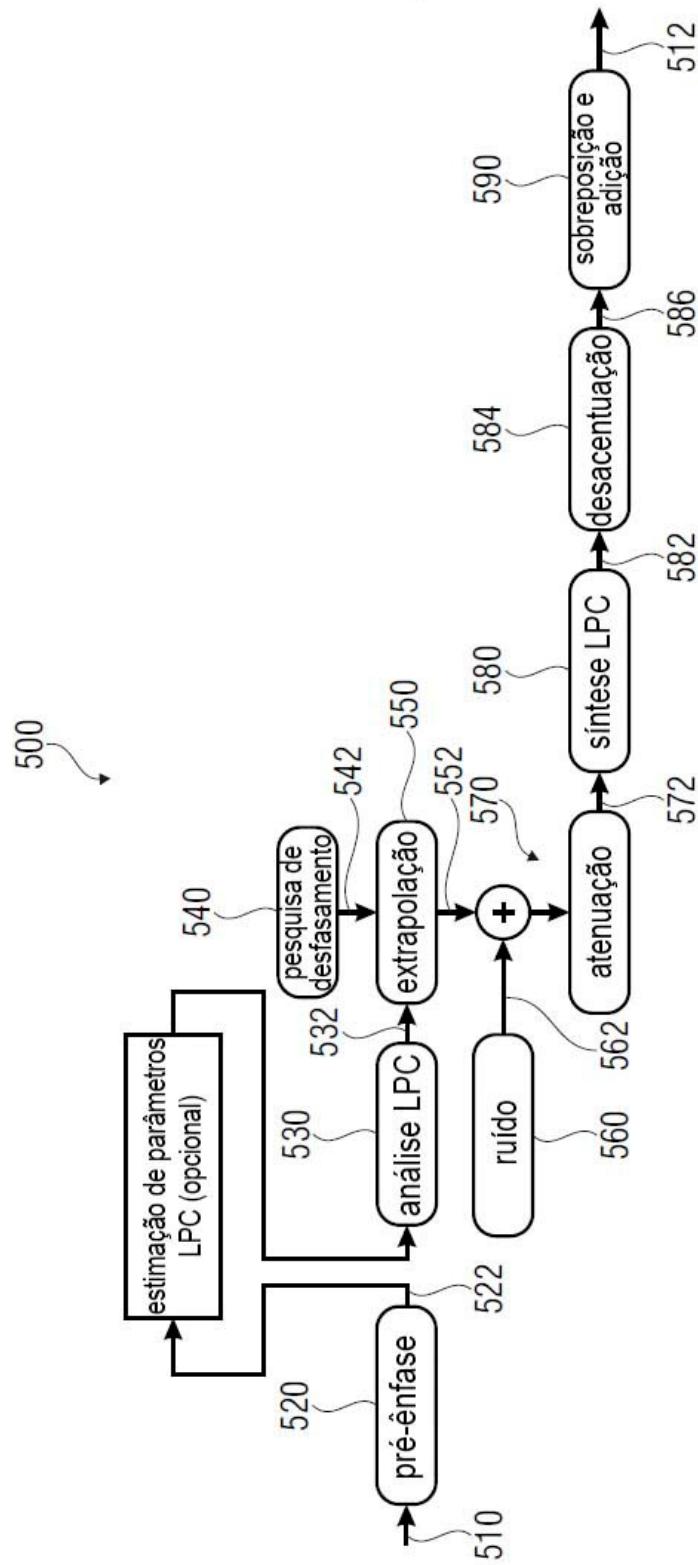


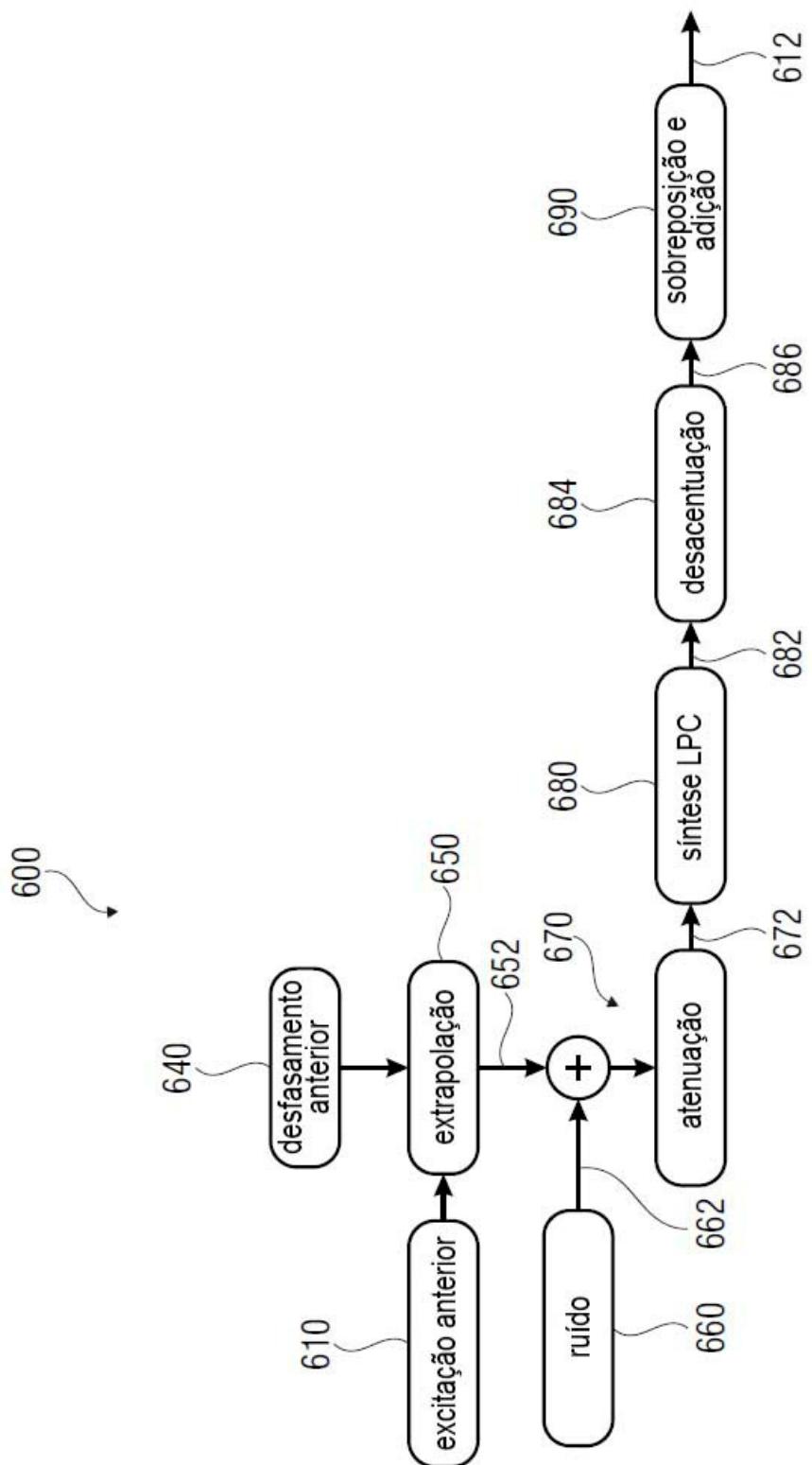
FIG 4B





Visão global da dissimulação de domínio de tempo para descodificador por transformação

FIG 5



Visão global da dissimulação de domínio de tempo para comutação do codec

FIG6

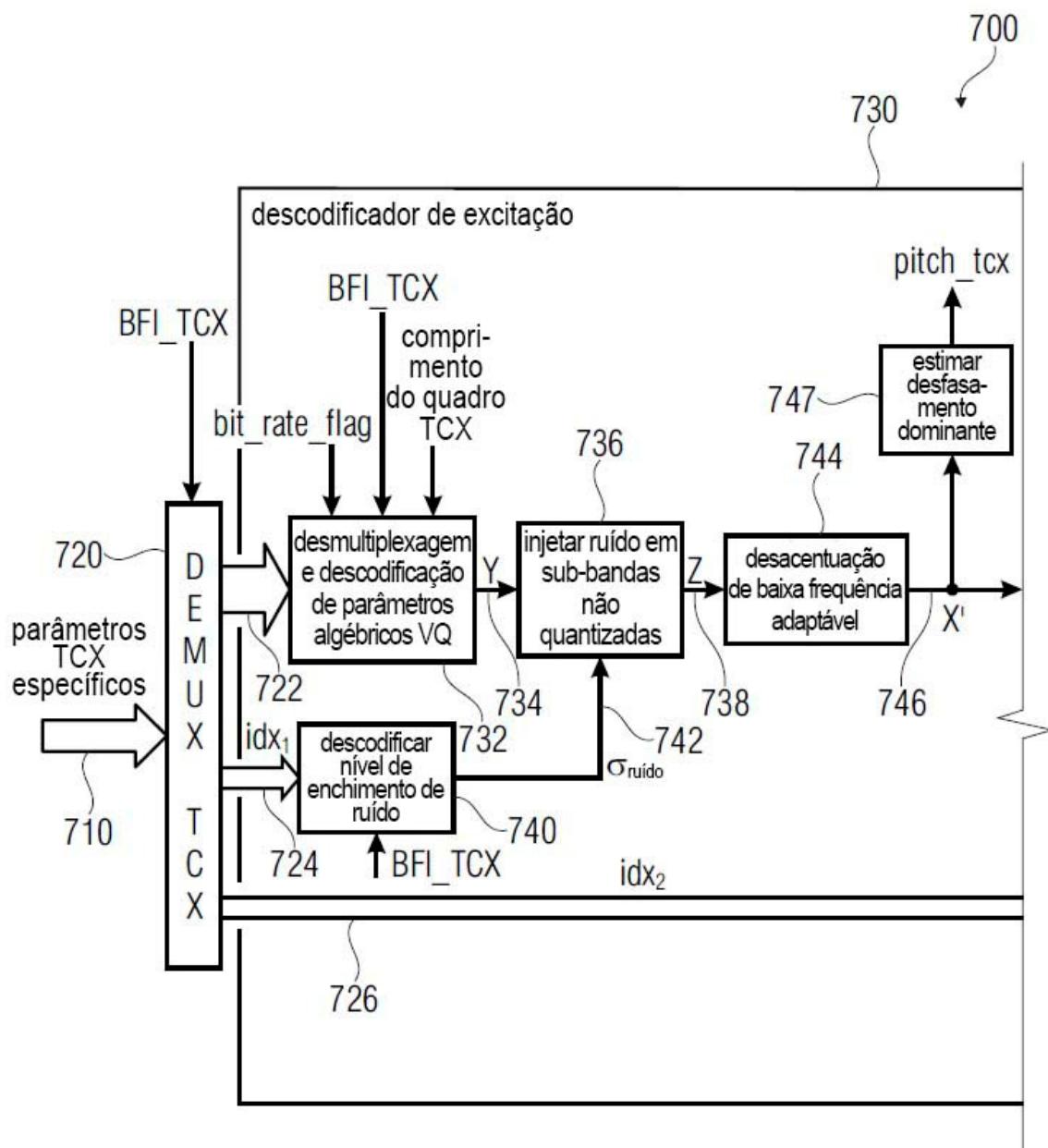


Diagrama de blocos do descodificador TCX

FIG 7

FIG
7A

FIG 7A

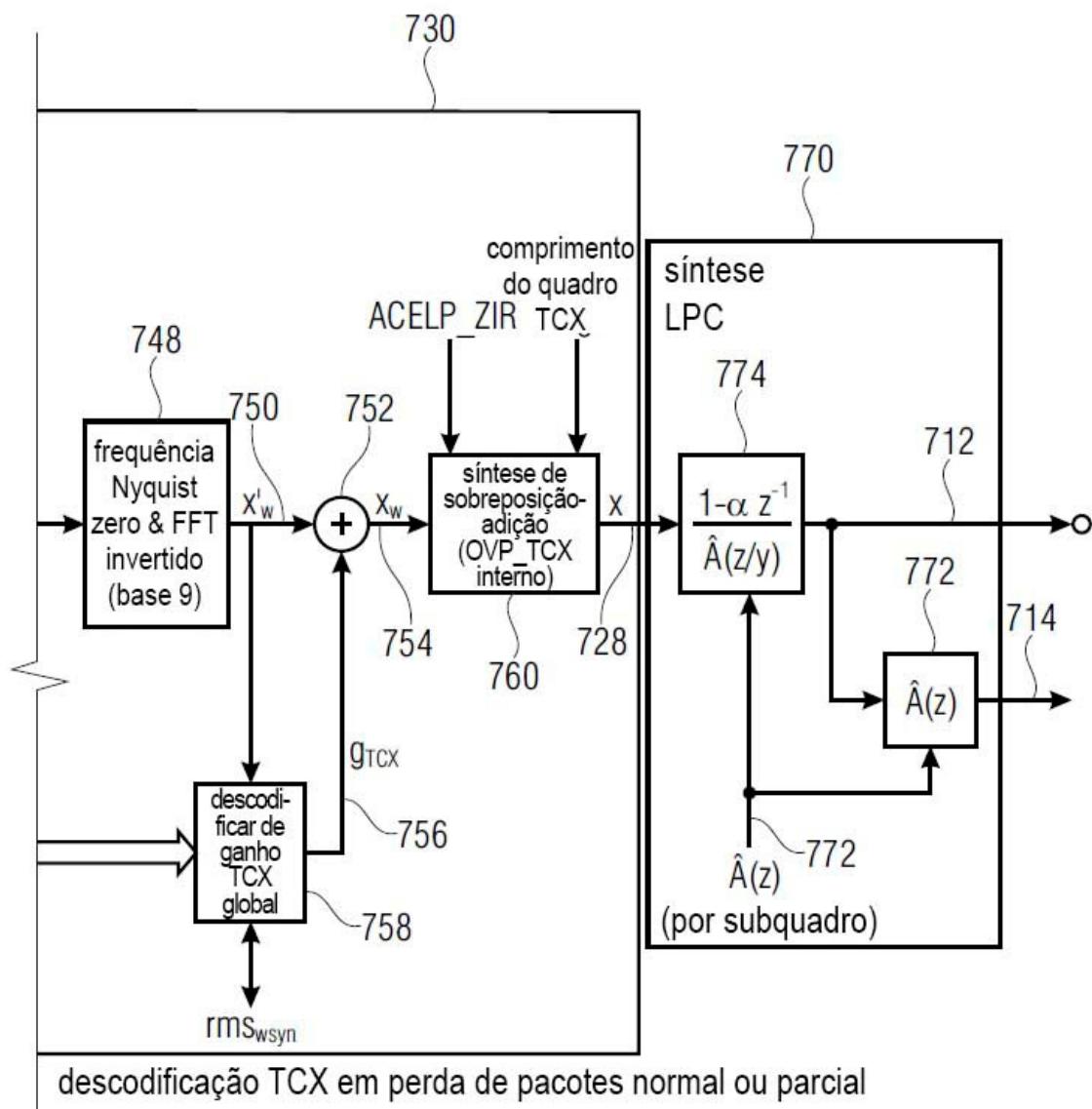
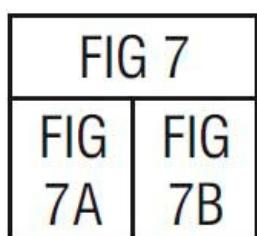


Diagrama de blocos do descodificador TCX

FIG 7B



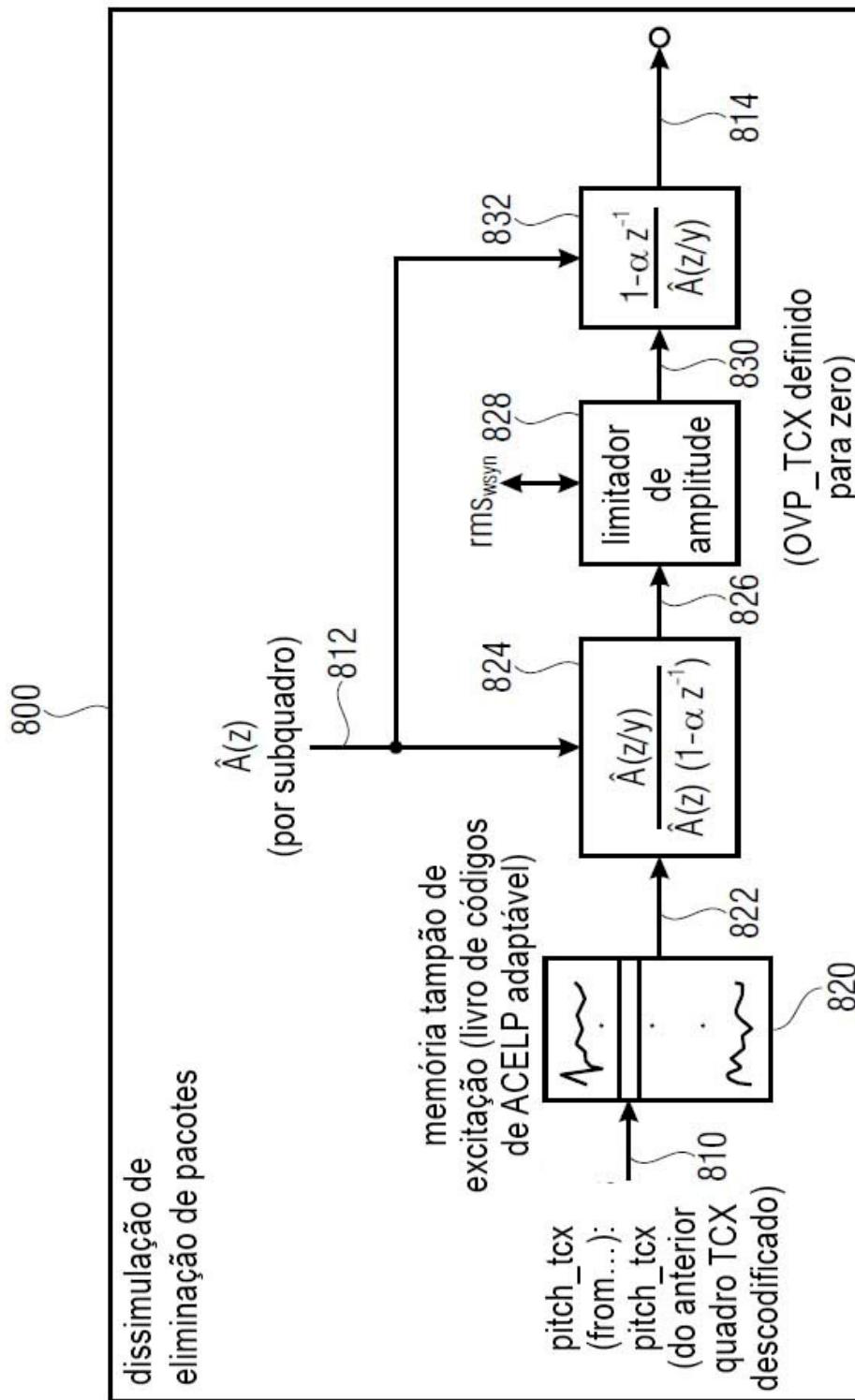
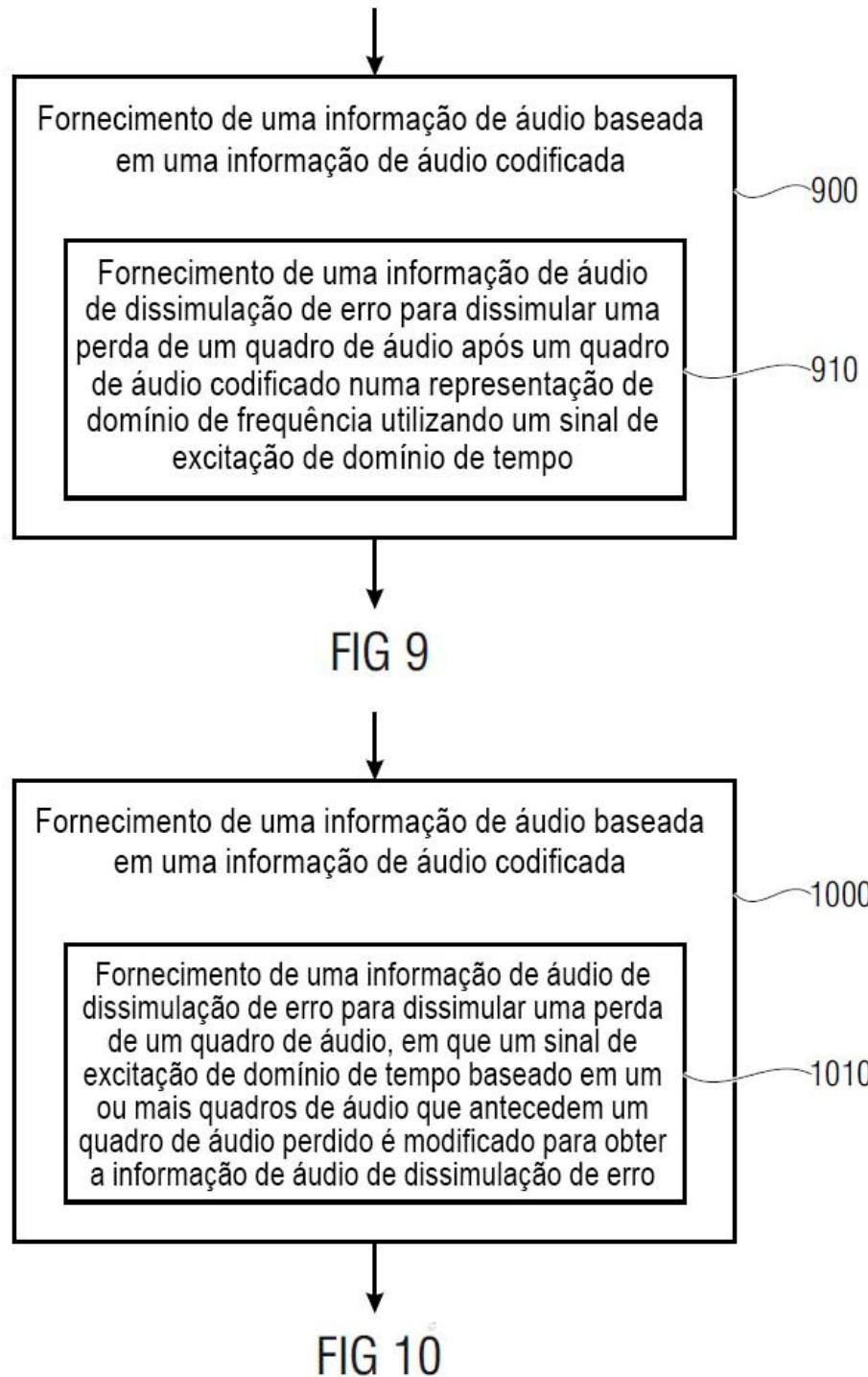


Diagrama de blocos do descodificador TCX
FIG 8



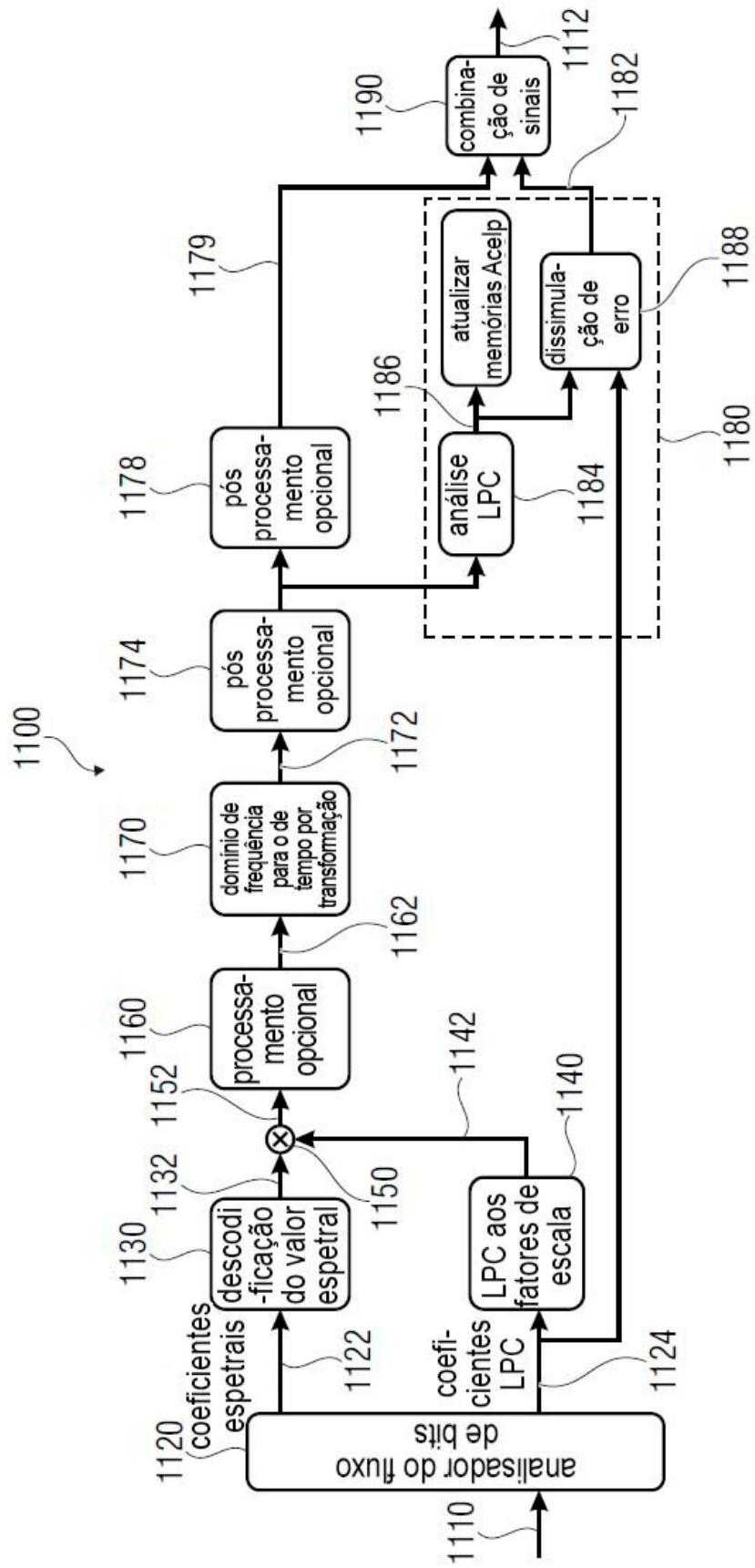


FIG 11