



República Federativa do Brasil

Ministério do Desenvolvimento, Indústria,
Comércio e Serviços

Instituto Nacional da Propriedade Industrial

(11) PI 0906319-6 B1

(22) Data do Depósito: 01/07/2009

(45) Data de Concessão: 04/06/2024

(54) Título: DECODIFICADOR DE SINAL DE ÁUDIO, CODIFICADOR DE SINAL DE ÁUDIO, REPRESENTAÇÃO DE SINAL DE ÁUDIO MULTI-CANAL CODIFICADO E MÉTODOS

(51) Int.Cl.: G10L 19/00; G10L 19/02; G10L 21/04.

(30) Prioridade Unionista: 08/10/2008 US 61/103,820; 11/07/2008 US 61/079,873.

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

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

(86) Pedido PCT: PCT EP2009004758 de 01/07/2009

(87) Publicação PCT: WO 2010/003583 de 14/01/2010

(85) Data do Início da Fase Nacional: 30/09/2010

(57) Resumo: DECODIFICADOR DE SINAL DE ÁUDIO, CODIFICADOR DE SINAL DE ÁUDIO, REPRESENTAÇÃO DE SINAL DE ÁUDIO MULTI-CANAL CODIFICADO, MÉTODOS E PROGRAMA DE COMPUTADOR. Um decodificador de sinal de áudio para fornecer uma representação de sinal de áudio multi-canal decodificado com base em uma representação de sinal de áudio multi-canal codificado compreende um decodificador de distorção de tempo configurado para seletivamente usar contornos individuais de distorção de tempo específicos para canal de áudio ou um contorno de distorção de tempo multi-canal conjunto para uma reconstrução de uma pluralidade de canais de áudio representados pela representação de sinal de áudio multi-canal codificado. Um codificador de sinal de áudio para fornecer uma representação codificada de um sinal de áudio multi-canal compreende um provedor de representação de áudio codificado configurado para seletivamente fornecer uma representação de áudio compreendendo uma informação de contorno comum de distorção de tempo, comumente associada com uma pluralidade de canais de áudio do sinal de áudio multi-canal, ou uma representação de áudio codificada compreendendo informação individual de contorno de distorção de tempo, individualmente associada com os diferentes canais de áudio da pluralidade de canais de áudio, na dependência de uma informação descrevendo uma similaridade (...).

"DECODIFICADOR DE SINAL DE ÁUDIO, CODIFICADOR DE SINAL DE ÁUDIO, REPRESENTAÇÃO DE SINAL DE ÁUDIO MULTI-CANAL CODIFICADO E MÉTODOS"

HISTÓRICO DA INVENÇÃO

5 Configurações de acordo com a invenção são relacionadas com um decodificador de sinal de áudio. Configurações adicionais de acordo com a invenção são relacionadas com um codificador de sinal de áudio. Configurações adicionais de acordo com a invenção são relacionadas com uma representação de sinal de

10 áudio multi-canal codificado. Configurações adicionais de acordo com a invenção são relacionadas com um método para fornecer uma representação de sinal de áudio multi-canal decodificado, com um método para fornecer uma representação codificada de um sinal de áudio multi-canal, e com um programa de computador para

15 implementar os referidos métodos.

 Algumas configurações de acordo com a invenção estão relacionadas com métodos para um codificador de transformação MDCT de tempo distorcido.

 A seguir, será fornecida uma breve introdução ao

20 campo do codificador de áudio de tempo distorcido, conceitos dos quais podem ser aplicados em conjunto com algumas das configurações da invenção.

 Nos anos recentes, foram desenvolvidas técnicas para transformar um sinal de áudio em uma representação de domínio

25 de frequência, e para codificar eficientemente esta representação de domínio de frequência, por exemplo levando em conta limiares de máscara conceituais. Este conceito de codificação sinal de áudio é particularmente eficiente se os comprimentos de bloco, para o qual

um conjunto de coeficientes espectrais são transmitidos, forem longos, e se apenas um número comparativamente pequeno de coeficientes espectrais estiverem bem acima do limiar de máscara global enquanto um grande número de coeficientes espectrais estiverem próximos ou abaixo do limiar de máscara global e puderem ser desprezados (ou codificados com duração de código mínima).

Por exemplo, transformadas cíclicas moduladas baseadas em senos e cossenos são frequentemente usadas em aplicações para codificação de fonte devido às suas propriedades de compactação de energia. Isto é, para tons harmônicos com frequências fundamentais constantes (passo), elas concentram a energia sinal em um pequeno número de componentes espectrais (sub-bandas), o que leva a uma representação sinal eficiente.

Geralmente, o passo (fundamental) de um sinal deve ser entendido como sendo a menor frequência dominante distinguível de um espectro do sinal. No modelo do discurso comum, o passo é a frequência do sinal de excitação modulada pela garganta humana. Se apenas uma única frequência fundamental estivesse presente, o espectro seria extremamente simples, compreendendo a frequência fundamental e os sobretons apenas. Tal espectro poderia ser codificado com grande eficiência. Para sinais com passo variável, contudo, a energia correspondente a cada componente harmônico é espalhada por diversos coeficientes de transformada, desta forma levando a uma redução da eficiência de codificação.

A fim de contornar esta redução da eficiência de codificação, o sinal de áudio a ser codificado é efetivamente reamostrado em uma grade temporal não-uniforme. No processamento

subsequente, as posições de amostra obtidas pela reamostragem não-uniforme são processadas como se representassem valores em uma grade temporal uniforme. Esta operação é comumente denotada pela frase "distorção de tempo". Os tempos de amostra podem ser
5 escolhidos com vantagem na dependência da variação temporal do passo, tal que uma variação de passo na versão de tempo distorcido do sinal de áudio é menor do que uma variação de passo na versão original do sinal de áudio (antes da distorção de tempo). Após a distorção de tempo do sinal de áudio, a versão
10 de tempo distorcido do sinal de áudio é convertida no domínio de frequência. A distorção de tempo dependente do passo tem o efeito que a representação do domínio de frequência do sinal de áudio de tempo distorcido é tipicamente concentrada em um número muito menor de componentes espectrais do que uma representação
15 do domínio de frequência do sinal de áudio original (tempo não-distorcido).

No lado do decodificador, a representação do domínio de frequência do sinal de áudio de tempo distorcido é convertida de volta ao domínio de tempo, de tal forma que uma
20 representação do domínio de tempo do sinal de áudio de tempo distorcido está disponível no lado do decodificador. Entretanto, na representação do domínio de tempo do sinal de áudio de tempo distorcido reconstruído do lado do decodificador, as variações de passo originais do sinal de áudio de entrada no lado do
25 decodificador não são incluídas. Da mesma forma, ainda outra distorção de tempo por reamostragem da representação do sinal de áudio de tempo distorcido do domínio de tempo reconstruído pelo lado do decodificador é aplicada. A fim de obter uma boa

reconstrução do sinal de áudio de entrada pelo lado do codificador no decodificador, é desejável que a distorção de tempo no lado do decodificador seja pelo menos aproximadamente a operação inversa com respeito à distorção de tempo do lado do codificador. A fim de obter uma distorção de tempo apropriada, é desejável ter uma informação disponível no decodificador que permita um ajuste da distorção de tempo do lado do decodificador.

Como é tipicamente necessário transferir tal informação do codificador de sinal de áudio para o decodificador de sinal de áudio, é desejável manter uma taxa de bits necessária para esta transmissão pequena enquanto ainda se disponibiliza uma reconstrução confiável da necessária informação de distorção de tempo no lado do decodificador.

Em vista da discussão acima, há um desejo de ter um conceito que permita um armazenamento eficiente de taxa de bits e/ou transmissão de um sinal de áudio multi-canal.

SUMÁRIO DA INVENÇÃO

Uma configuração de acordo com a invenção cria um decodificador de sinal de áudio para fornecer uma representação de sinal de áudio multi-canal decodificado com base em uma representação de sinal de áudio multi-canal codificado. O decodificador de sinal de áudio compreende um decodificador de distorção de tempo configurado para seletivamente usar contornos de distorção de tempo específicos para canal de áudio, individuais ou um contorno de distorção de tempo multi-canal conjunto para uma reconstrução de distorção de tempo de uma pluralidade de canais de áudio representados pela representação de sinal de áudio multi-canal codificado.

Esta configuração de acordo com a invenção baseia-se no achado de que uma codificação eficiente de diferentes tipos de sinais de áudio multi-canal pode ser conseguida alternando-se entre um armazenamento e/ou transmissão de contornos de distorção de tempo específicos para canal de áudio e contornos de distorção de tempo multi-canal conjuntos. Descobriu-se que em alguns casos, uma variação de passo é significativamente diferente nos canais de um sinal de áudio multi-canal. Além disso, descobriu-se que em outros casos, a variação de passo é aproximadamente igual para múltiplos canais de um sinal de áudio multi-canal. Em vista destes diferentes tipos de sinais (ou porções de sinal de um único sinal de áudio), descobriu-se que a eficiência de codificação pode ser melhorada se o decodificador for capaz de flexivelmente (intercambiavelmente, ou seletivamente) derivar os contornos de distorção de tempo para a reconstrução dos diferentes canais do sinal de áudio multi-canal a partir de representações de contorno de distorção de tempo específicas para canal de áudio, individuais, ou a partir de uma representação de contorno de distorção de tempo multi-canal, conjunta.

Em uma configuração preferida, o decodificador de distorção de tempo é configurado para seletivamente usar um contorno de distorção de tempo multi-canal conjunto para uma reconstrução de distorção de tempo de uma pluralidade de canais de áudio para os quais está disponível informação individual codificada de domínio espectral. De acordo com um aspecto da invenção, descobriu-se que o uso de um contorno de distorção de tempo multi-canal conjunto para uma reconstrução de distorção de tempo de uma pluralidade de canais de áudio não é apenas aplicável

se os diferentes canais de áudio representarem um conteúdo de áudio similar, porém mesmo se diferentes canais de áudio representarem um conteúdo de áudio significativamente diferente. Da mesma forma, descobriu-se que é útil combinar o conceito de

5 usar um contorno de distorção de tempo multi-canal conjunto para a avaliação de informação individual codificada de domínio espectral para diferentes canais de áudio. Por exemplo, este conceito é particularmente útil se um primeiro canal de áudio representa uma primeira parte de uma peça polifônica de música, enquanto um

10 segundo canal de áudio representa uma segunda parte da peça polifônica de música. O primeiro sinal de áudio e o segundo sinal de áudio podem, por exemplo, representar o som produzido por diferentes cantores ou por diferentes instrumentos. Da mesma forma, uma representação de domínio espectral do primeiro canal de

15 áudio pode ser significativamente diferente de uma representação de domínio espectral do segundo canal de áudio. Por exemplo, as frequências fundamentais dos diferentes canais de áudio podem ser diferentes. Também, os diferentes canais de áudio podem compreender diferentes características com respeito às harmônicas

20 da frequência fundamental. Não obstante, pode haver uma significativa tendência que os passos dos diferentes canais de áudio variem aproximadamente em paralelo. Neste caso, é muito eficiente aplicar uma distorção de tempo comum (descrita pelo contorno de distorção de tempo multi-canal conjunto) aos

25 diferentes canais de áudio, muito embora os diferentes canais de áudio compreendam conteúdos de áudio significativamente diferentes (por exemplo, apresentando diferentes frequências fundamentais e diferentes espectros harmônicos). Não obstante, em outros casos, é

naturalmente desejável aplicar diferentes distorções de tempo a diferentes canais de áudio.

Em uma configuração preferida da invenção, o decodificador de distorção de tempo é configurado para receber uma primeira informação de domínio espectral codificado associada com um primeiro dos canais de áudio e para fornecer, com base nele, uma representação de domínio de tempo distorcido do primeiro canal de áudio usando uma transformação de domínio de frequências para domínio de tempo. Também, o decodificador de distorção de tempo é adicionalmente configurado para receber uma segunda informação de domínio espectral codificado, associada com um segundo dos canais de áudio, e para fornecer, com base nele, uma representação de domínio de tempo distorcido do segundo canal de áudio usando uma transformação de domínio de frequências para domínio de tempo.

Neste caso, a segunda informação de domínio espectral codificado pode ser diferente da primeira informação de domínio espectral. Também, o decodificador de distorção de tempo é configurado para reamostrar de forma variável com o tempo, com base no contorno de distorção de tempo multi-canal conjunto, a representação de domínio de tempo distorcido do primeiro canal de áudio, uma versão processada desta, para obter uma representação regularmente amostrada do primeiro canal de áudio, e para reamostrar de forma variável com o tempo, também com base no contorno de distorção de tempo multi-canal conjunto, a representação de domínio de tempo distorcido do segundo canal de áudio, ou uma segunda versão processada desta, para obter uma representação regularmente amostrada do segundo canal de áudio.

Em outra configuração preferida, o decodificador de distorção de tempo é configurado para derivar um contorno de tempo multi-canal conjunto a partir da informação de contorno de distorção de tempo multi-canal conjunto. Adicionalmente, o

5 decodificador de distorção de tempo é configurado para derivar uma primeira forma de janela individual, específica de canal associada com o primeiro dos canais de áudio com base em uma primeira informação de forma de janela codificada, e para derivar uma segunda forma de janela individual, específica de canal associada

10 com o segundo dos canais de áudio com base em uma segunda informação de forma de janela codificada. O decodificador de distorção de tempo é adicionalmente configurado para aplicar a primeira forma de janela à representação de domínio de tempo distorcido do primeiro canal de áudio, para obter uma versão

15 processada da representação de domínio de tempo distorcido do primeiro canal de áudio, e para aplicar a segunda forma de janela à representação de domínio de tempo distorcido do segundo canal de áudio, para obter uma versão processada da representação de domínio de tempo distorcido do segundo canal de áudio. Neste caso,

20 o decodificador de distorção de tempo é capaz de aplicar diferentes formas de janela às representações de domínio de tempo distorcido do primeiro e do segundo canal de áudio na dependência de uma informação de forma de janela específica de canal, individual.

25 Descobriu-se que é em alguns casos recomendável aplicar janelas de diferentes formas a diferentes sinais de áudio na preparação de uma operação de distorção de tempo, mesmo se as operações de distorção de tempo se basearem no contorno comum de

distorção de tempo. Por exemplo, pode haver uma transição entre uma estrutura, em que há um contorno comum de distorção de tempo para dois canais de áudio, e uma subsequente estrutura em que há diferentes contornos de distorção de tempo para os dois canais de áudio. Entretanto, o contorno de distorção de tempo de um dos dois canais de áudio na subsequente estrutura pode ser uma continuação não variável do contorno comum de distorção de tempo na presente estrutura, enquanto o contorno de distorção de tempo do outro canal de áudio na subsequente estrutura pode estar variando com respeito ao contorno comum de distorção de tempo na presente estrutura. Da mesma forma, uma forma de janela que esteja adaptada a uma evolução não variável do contorno de distorção de tempo pode ser usada para um dos canais de áudio, enquanto uma forma de janela adaptada a uma evolução variável do contorno de distorção de tempo pode ser aplicada para o outro canal de áudio. Desta forma, a evolução diferente dos canais de áudio pode ser levada em consideração.

Em outra configuração de acordo com a invenção, o decodificador de distorção de tempo pode ser configurado para aplicar uma escalação de tempo comum, que é determinada pelo contorno de distorção de tempo multi-canal conjunto, e diferentes formas de janela ao aplicar janela às representações de domínio de tempo do primeiro e segundo canais de áudio. Descobriu-se que mesmo se diferentes formas de janela forem usadas para aplicar janela a diferentes canais de áudio antes da respectiva distorção de tempo, a escalação de tempo do contorno de tempo deve ser adaptada em paralelo a fim de evitar uma degradação da impressão auditiva.

Outra configuração de acordo com a invenção cria um codificador de sinal de áudio para fornecer uma representação codificada de um sinal de áudio multi-canal. O codificador de sinal de áudio compreende um provedor de representação de áudio

5 codificado configurado para seletivamente fornecer uma representação de áudio compreendendo uma informação de contorno comum de distorção de tempo, comumente associada com uma pluralidade de canais de áudio do sinal de áudio multi-canal, ou uma representação de áudio codificada compreendendo informação

10 individual de contorno de distorção de tempo, individualmente associada com os diferentes canais de áudio da pluralidade de canais de áudio, na dependência de uma informação descrevendo uma similaridade ou diferença entre os contornos de distorção de tempo associados com os canais de áudio da pluralidade de canais de

15 áudio. Esta configuração de acordo com a invenção baseia-se no achado de que em muitos casos, múltiplos canais de um sinal de áudio multi-canal compreendem características similares de variação de passo. Da mesma forma, é em alguns casos eficiente incluir na representação codificada do sinal de áudio multi-canal

20 uma informação de contorno comum de distorção de tempo, comumente associada com uma pluralidade dos canais de áudio. Desta maneira, uma eficiência de codificação pode ser melhorada para muitos sinais. Entretanto, descobriu-se que para outros tipos de sinais (ou mesmo para outras porções de um sinal), não é recomendável

25 usar tal informação comum de distorção de tempo. Da mesma forma, uma codificação de sinal eficiente pode ser obtida se o codificador de sinal de áudio determinar a similaridade ou diferença entre contornos de tempo associados com os diferentes

canais de áudio sob consideração. Entretanto, descobriu-se que de fato vale a pena dar uma olhada nos contornos individuais de distorção de tempo, porque há muitos sinais que compreendem uma representação de domínio de tempo ou representação de domínio de frequências significativamente diferente, muito embora tenham contornos de distorção de tempo muito similares. Da mesma forma, descobriu-se que a avaliação do contorno de distorção de tempo é um novo critério para a avaliação da similaridade de sinais, que fornece uma informação extra quando comparado a uma mera avaliação das representações de domínio de tempo de múltiplos sinais de áudio ou das representações de domínio de frequência dos sinais de áudio.

Em uma configuração preferida, o provedor de representação de áudio codificado é configurado para aplicar uma informação de contorno comum de distorção de tempo para obter uma versão de tempo distorcido de um primeiro dos canais de áudio e para obter uma versão de tempo distorcido de um segundo dos canais de áudio. O provedor de representação de áudio codificado é adicionalmente configurado para fornecer uma primeira informação individual codificada de domínio espectral associada com o primeiro dos canais de áudio com base na versão de tempo distorcido do primeiro canal de áudio, e para fornecer uma segunda informação individual codificada de domínio espectral associada com o segundo canal de áudio com base na versão de tempo distorcido do segundo dos canais de áudio. Esta configuração baseia-se no achado supracitado de que canais de áudio podem ter conteúdos de áudio significativamente diferentes, mesmo se eles tiverem contorno de distorção de tempo similar. Desta forma, é

frequentemente recomendável fornecer diferente informação de domínio espectral associada com diferentes canais de áudio, mesmo se os canais de áudio forem tempo distorcidos de acordo com uma informação comum de distorção de tempo. Em outras palavras, a configuração baseia-se no achado de que não há inter-relação estrita entre uma similaridade dos contornos de distorção de tempo e uma similaridade das representações de domínio de frequências de diferentes canais de áudio.

Em outra configuração preferida, o codificador é configurado para obter a informação de contorno de tempo comum tal que o contorno de tempo comum represente uma média de contornos de tempo individuais associados ao primeiro canal de sinal de áudio e ao segundo canal de sinal de áudio.

Em outra configuração preferida, o provedor de representação de áudio codificado é configurado para fornecer uma informação lateral dentro da representação codificada do sinal de áudio multi-canal, tal que a informação lateral indique, conforme uma estrutura de áudio, se os dados de distorção de tempo estão presentes para uma estrutura e se uma informação de contorno comum de distorção de tempo está presente para uma estrutura. Fornecendo uma informação se dados de distorção de tempo estão presentes para uma estrutura, é possível reduzir uma velocidade de transferência de bits exigida para a transmissão da informação de distorção de tempo. Descobriu-se que é tipicamente exigido para transmitir uma informação descrevendo uma pluralidade de valores de contorno de distorção de tempo dentro de uma estrutura, se a distorção de tempo for usada para tal estrutura. Entretanto, descobriu-se também que há muitas estruturas para as quais a aplicação de uma

distorção de tempo não traz uma vantagem significativa. Ainda mais, descobriu-se que é mais eficiente indicar, usando, por exemplo, um bit de informação adicional, se os dados de distorção de tempo para uma estrutura estão disponíveis. Usando tal sinalização, a transmissão da extensiva informação de distorção de tempo (tipicamente compreendendo informação relativa a uma pluralidade de valores de contorno de distorção de tempo) pode ser omitida, deste modo poupando bits.

Uma configuração adicional de acordo com a invenção cria uma representação de sinal de áudio multi-canal codificado representando um sinal de áudio multi-canal. A representação de sinal de áudio multi-canal compreende uma representação de domínio de frequências codificada representando uma pluralidade de canais de áudio de tempo distorcido, tempo distorcido de acordo com uma distorção de tempo comum. A representação de sinal de áudio multi-canal também compreende uma representação codificada de uma informação de contorno comum de distorção de tempo, comumente associada com os canais de áudio e representando a distorção de tempo comum.

Em uma configuração preferida, a representação de domínio de frequências codificada compreende informação de domínio de frequências codificada de múltiplos canais de áudio que têm conteúdo de áudio diferente. Também, a representação codificada da informação de contorno de tempo comum está associada com os múltiplos canais de áudio que têm diferentes conteúdos de áudio.

Outra configuração de acordo com a invenção cria um método para fornecer uma representação de sinal de áudio multi-canal decodificado com base em uma representação de sinal de áudio

multi-canal codificado. Este método pode ser suplementado por quaisquer dos aspectos e funcionalidades descritos aqui também para o aparelho inventivo.

5 Ainda outra configuração de acordo com a invenção cria um método para fornecer uma representação codificada de um sinal de áudio multi-canal. Este método pode ser suplementado por quaisquer dos aspectos e funcionalidades descritos aqui também para o aparelho inventivo.

10 Ainda outra configuração de acordo com a invenção cria um programa de computador para implementar os métodos descritos acima.

BREVE DESCRIÇÃO DAS FIGURAS

15 Configurações de acordo com a invenção serão descritas a seguir tomando referência às figuras inclusas, nas quais:

A Figura 1 mostra um diagrama esquemático de blocos de um codificador de áudio de distorção de tempo;

A Figura 2 mostra um diagrama esquemático de blocos de um decodificador de áudio de distorção de tempo;

20 A Figura 3 mostra um diagrama esquemático de blocos de um decodificador de sinal de áudio, de acordo com uma configuração da invenção;

25 A Figura 4 mostra um fluxograma de um método para fornecer uma representação de sinal de áudio decodificado, de acordo com uma configuração da invenção;

A Figura 5 mostra um extrato detalhado de um diagrama esquemático de blocos de um decodificador de sinal de áudio de acordo com uma configuração da invenção;

A Figura 6 mostra um extrato detalhado de um fluxograma de um método para fornecer uma representação de sinal de áudio decodificado de acordo com uma configuração da invenção;

5 As Figuras 7a, 7b mostram uma representação gráfica de uma reconstrução de um contorno de distorção de tempo, de acordo com uma configuração da invenção;

A Figura 8 mostra outra representação gráfica de uma reconstrução de um contorno de distorção de tempo, de acordo com uma configuração da invenção;

10 As Figuras 9a e 9b mostram algoritmos para o cálculo do contorno de distorção de tempo;

A Figura 9c mostra uma tabela de mapeamento a partir de um índice de proporção de distorção de tempo até um valor de proporção de distorção de tempo;

15 As Figuras 10a e 10b mostram representações de algoritmos para o cálculo de um contorno de tempo, uma posição de amostra, um comprimento de transição, uma "primeira posição" e uma "última posição";

20 A Figura 10c mostra uma representação de algoritmos para um cálculo de forma de janela;

As Figuras 10d e 10e mostram uma representação de algoritmos para uma aplicação de uma janela;

A Figura 10f mostra uma representação de algoritmos para uma reamostragem de tempo variável;

25 A Figura 10g mostra uma representação gráfica de algoritmos para um processamento de estrutura pós-distorção de tempo e para uma sobreposição e soma;

As Figuras 11a e 11b mostram uma legenda;

A Figura 12 mostra uma representação gráfica de um contorno de tempo, que pode ser extraído de um contorno de distorção de tempo;

5 A Figura 13 mostra um diagrama esquemático de blocos detalhado de um aparelho para fornecer um contorno de distorção, de acordo com uma configuração da invenção;

A Figura 14 mostra um diagrama esquemático de blocos de um decodificador de sinal de áudio, de acordo com outra configuração da invenção;

10 A Figura 15 mostra um diagrama esquemático de blocos de outra calculadora de contorno de distorção de tempo de acordo com uma configuração da invenção;

15 As Figuras 16a, 16b mostram uma representação gráfica de um cômputo de valores de nó de distorção de tempo, de acordo com uma configuração da invenção;

A Figura 17 mostra um diagrama esquemático de blocos de outro codificador de sinal de áudio, de acordo com uma configuração da invenção;

20 A Figura 18 mostra um diagrama esquemático de blocos de outro decodificador de sinal de áudio, de acordo com uma configuração da invenção; e

As Figuras 19a-19f mostram representações de elementos de sintaxe de um *stream* de áudio, de acordo com uma configuração da invenção;

25

DESCRIÇÃO DETALHADA DAS CONFIGURAÇÕES

1. CODIFICADOR DE ÁUDIO DE DISTORÇÃO DE TEMPO DE ACORDO COM A FIGURA 1

Como a presente invenção está relacionada com a

codificação de áudio de distorção de tempo e decodificação de áudio de distorção de tempo, uma breve visão geral será apresentada de um protótipo de codificador de áudio de distorção de tempo e um decodificador de áudio de distorção de tempo, em que
5 a presente invenção pode ser aplicada.

A Figura 1 mostra um diagrama esquemático de blocos de um codificador de áudio de distorção de tempo, no qual alguns aspectos e configurações da invenção podem ser integrados. O codificador de sinal de áudio 100 da Figura 1 é configurado para
10 receber um sinal de entrada de áudio 110 e para fornecer uma representação codificada do sinal de entrada de áudio 110 em uma sequência de estruturas. O codificador de áudio 100 compreende um amostrador 104, que é adaptado para amostrar o sinal de áudio 110 (sinal de entrada) para derivar blocos de sinais (representações amostradas) 105 usados como base para uma transformação de domínio
15 de frequência. O codificador de áudio 100 compreende adicionalmente uma calculadora de janela de transformada 106, adaptada para derivar janelas de escalação para a saída de representações amostradas 105 do amostrador 104. Estas são
20 inseridas em um windower 108 que é adaptado para aplicar as janelas de escalação às representações amostradas 105 derivadas pelo amostrador 104. Em algumas configurações, o codificador de áudio 100 pode adicionalmente compreender um transformador de domínio de frequência 108a, a fim de derivar uma representação de domínio de
25 frequência (por exemplo, na forma de coeficientes de transformada) das representações amostradas e escaladas 105. As representações de domínio de frequência podem ser processadas ou adicionalmente transmitidas como uma representação codificada do sinal de áudio 110.

O codificador de áudio 100 adicionalmente usa um contorno de passo 112 do sinal de áudio 110, que pode ser fornecido ao codificador de áudio 100 ou que pode ser derivado pelo codificador de áudio 100. O codificador de áudio 100 pode portanto 5 opcionalmente compreender um estimador de passo para derivar o contorno de passo 112. O amostrador 104 pode operar em uma representação contínua do sinal de entrada de áudio 110. Alternativamente, o amostrador 104 pode operar em uma representação já amostrada do sinal de entrada de áudio 110. No 10 último caso, o amostrador 104 pode reamostrar o sinal de áudio 110. O amostrador 104 pode, por exemplo, pode ser adaptado para distorção de tempo próxima a blocos de áudio sobreponentes tal que a porção de sobreposição tenha um passo constante ou variação de passo reduzida dentro de cada um dos blocos de entrada após a 15 amostragem.

A calculadora de janela de transformada 106 deriva as janelas de escalação para os blocos de áudio dependendo da distorção de tempo realizada pelo amostrador 104. Para este fim, um bloco de ajuste de taxa de amostragem opcional 114 pode estar 20 presente a fim de definir uma regra de distorção de tempo usada pelo amostrador, que é então também fornecida para a calculadora de janela de transformada 106. Em uma configuração alternativa, o bloco de ajuste de taxa de amostragem 114 pode ser omitido e o contorno de passo 112 pode ser diretamente fornecido para a calculadora de janela 25 de transformada 106, que pode ela mesma realizar os cálculos apropriados. Além disso, o amostrador 104 pode comunicar a amostragem aplicada à calculadora de janela de transformada 106 a fim de habilitar o cálculo de janelas de escalação apropriadas.

A distorção de tempo é realizada tal que um contorno de passo tempo distorcido de blocos de áudio amostrados e amostrados pelo amostrador 104 não seja mais constante do que o contorno de passo do sinal de áudio original 110 dentro do bloco de entrada.

2. DECODIFICADOR DE ÁUDIO DE DISTORÇÃO DE TEMPO DE ACORDO COM A FIGURA 2

A Figura 2 mostra um diagrama esquemático de blocos de um decodificador de áudio de distorção de tempo 200 para processar um primeiro tempo distorcido e amostrado, ou simplesmente representação de tempo distorcido de uma primeira e segunda estrutura de um sinal de áudio tendo uma sequência de estruturas em que a segunda estrutura segue a primeira estrutura e para adicionalmente processar uma segunda representação de tempo distorcido da segunda estrutura e de uma terceira estrutura seguindo a segunda estrutura na sequência de estruturas. O decodificador de áudio 200 compreende uma calculadora de janela de transformada 210 adaptada para derivar uma primeira janela de escalação para a primeira representação de tempo distorcido 211a usando informação sobre um contorno de passo 212 da primeira e segunda estrutura e para derivar uma segunda janela de escalação para a segunda representação de tempo distorcido 211b usando informação sobre um contorno de passo da segunda e terceira estrutura, em que as janelas de escalação podem ter números idênticos de amostras e em que o primeiro número de amostras usadas para fechar a primeira janela de escalação podem diferir de um segundo número de amostras usadas para abrir a segunda janela de escalação. O decodificador de áudio 200 adicionalmente

compreende um windower 216 adaptado para aplicar a primeira janela de escalação à primeira representação de distorção de tempo e para aplicar a segunda janela de escalação à segunda representação de tempo distorcido. O decodificador de áudio 200 além disso

5 compreende um reamostrador 218 adaptado para distorcer inversamente o tempo na primeira representação escalada de tempo distorcido para derivar uma primeira representação amostrada usando a informação sobre o contorno de passo da primeira e segunda estrutura e para distorcer inversamente o tempo na segunda

10 representação escalada para derivar uma segunda representação amostrada usando a informação sobre o contorno de passo da segunda e terceira estrutura tal que uma porção da primeira representação amostrada correspondente à segunda estrutura compreende um contorno de passo que iguala, dentro de uma faixa de tolerância

15 predeterminada, um contorno de passo da porção da segunda representação amostrada correspondente à segunda estrutura. A fim de derivar a janela de escalação, a calculadora de janela de transformada 210 pode ou receber o contorno de passo 212 diretamente ou receber informação sobre a distorção de tempo a partir

20 de um ajustador de taxa de amostra opcional 220, que recebe o contorno de passo 212 e que deriva uma estratégia inversa de distorção de tempo de tal maneira que as posições de amostra em uma escala de tempo linear para as amostras das regiões de sobreposição sejam idênticas ou quase idênticas e regularmente espaçadas, de modo

25 que o passo se torne o mesmo nas regiões de sobreposição, e opcionalmente as durações de desvanecimento diferentes de partes de sobreposição de janela antes da distorção de tempo inversa se tornam de mesmo comprimento após a distorção de tempo inversa.

O decodificador de áudio 200 além disso compreende um adicionador opcional 230, que é adaptado para adicionar a porção da primeira representação amostrada correspondente à segunda estrutura e a porção da segunda representação amostrada correspondente à segunda estrutura para derivar uma representação reconstruída da segunda estrutura do sinal de áudio como um sinal de saída 242. A primeira representação de tempo distorcido e a segunda representação de tempo distorcido poderiam, em uma configuração, ser fornecidas como uma entrada para o decodificador de áudio 200. Em uma configuração adicional, o decodificador de áudio 200 pode, opcionalmente, compreender um transformador de domínio de frequência inversa 240, que pode derivar a primeira e segunda representações de tempo distorcido a partir de representações de domínio de frequência da primeira e segunda representações de tempo distorcido fornecidas para a entrada do transformador de domínio de frequência inversa 240..

3. DECODIFICADOR DE DISTORÇÃO DE TEMPO DE SINAL DE ÁUDIO DE ACORDO COM A FIGURA 3

A seguir, será descrito um decodificador de sinal de áudio simplificado. A Figura 3 mostra um diagrama esquemático de blocos deste decodificador de sinal de áudio simplificado 300. O decodificador de sinal de áudio 300 é configurado para receber a representação de sinal de áudio codificado 310, e para fornecer, com base nela, uma representação de sinal de áudio decodificado 312, em que a representação de sinal de áudio codificado 310 compreende uma informação de evolução de contorno de distorção de tempo. O

decodificador de sinal de áudio 300 compreende uma calculadora de contorno de distorção de tempo 320 configurada para gerar dados de contorno de distorção de tempo 322 com base na informação de evolução de contorno de distorção de tempo, informação de evolução de contorno de distorção de tempo esta que descreve uma evolução temporal do contorno de distorção de tempo, e informação de evolução de contorno de distorção de tempo esta que é composta pela representação de sinal de áudio codificado 310. Ao derivar os dados de contorno de distorção de tempo 322 a partir da informação de evolução de contorno de distorção de tempo 312, a calculadora de contorno de distorção de tempo 320 repetidamente reinicia a partir de um valor inicial de contorno de distorção de tempo predeterminado, como será descrito em detalhes a seguir. O reinício pode ter a consequência de que o contorno de distorção de tempo compreenda descontinuidades (mudanças step-wise que são maiores do que os passos codificados pela informação de evolução de contorno de distorção de tempo 312). O decodificador de sinal de áudio 300 também compreende dados de rescalador de contorno de distorção de tempo 330 que são configurados para rescalar pelo menos uma porção dos dados de contorno de distorção de tempo 322, tal que uma descontinuidade em um reinício do contorno de distorção de tempo cálculo seja evitada, reduzida ou eliminada em uma versão rescalada 332 do contorno de distorção de tempo.

O decodificador de sinal de áudio 300 também compreende um decodificador de distorção 340 configurado para prover uma representação de sinal de áudio decodificado 312 com base na representação de sinal de áudio codificado 310 e usando a versão rescalada 332 do contorno de distorção de tempo.

Para colocar o decodificador de sinal de áudio 300 no contexto de decodificação de áudio de distorção de tempo, deve-se notar que a representação de sinal de áudio codificado 310 pode compreender uma representação codificada dos coeficientes de transformada 211 e também uma representação codificada do contorno de passo 212 (também chamada de contorno de distorção de tempo). A calculadora de contorno de distorção de tempo 320 e os dados de rescalador de contorno de distorção de tempo 330 podem ser configurados para prover uma representação reconstruída do contorno de passo 212 na forma da versão rescalada 332 do contorno de distorção de tempo. O decodificador de distorção 340 pode, por exemplo, assumir a funcionalidade da janela 216, a reamostragem 218, o ajuste da taxa de amostra 220 e o ajuste da forma de janela 210. Além disso, o decodificador de distorção 340 pode, por exemplo, opcionalmente, compreender a funcionalidade da transformada inversa 240 e da sobreposição/soma 230, tal que a representação de sinal de áudio decodificado 312 pode ser equivalente ao sinal de saída de áudio 232 do decodificador de áudio de distorção de tempo 200.

Aplicando a rescalação dos dados de contorno de distorção de tempo 322, uma versão rescalada 332 contínua (ou pelo menos aproximadamente contínua) do contorno de distorção de tempo pode ser obtida, assim garantindo que um estouro ou *underflow* numérico seja evitado mesmo quando usando uma informação de evolução de contorno de distorção de tempo de variação relativa eficiente para codificar.

4. MÉTODO PARA FORNECER UMA REPRESENTAÇÃO DE SINAL DE ÁUDIO DECODIFICADO DE ACORDO COM A FIGURA 4.

A Figura 4 mostra um fluxograma de um método para fornecer uma representação de sinal de áudio decodificado com base

em uma representação de sinal de áudio codificado compreendendo uma informação de evolução de contorno de distorção de tempo, que pode ser realizada pelo aparelho 300 de acordo com a Figura 3. O método 400 compreende um primeiro passo 410 de geração de dados de contorno de distorção de tempo, reiniciando repetidamente a partir de um valor inicial de contorno de distorção de tempo predeterminado, com base em uma informação de evolução de contorno de distorção de tempo descrevendo uma evolução temporal do contorno de distorção de tempo...

10 O método 400 adicionalmente compreende um passo 420 de rescalação de pelo menos uma porção dos dados de controle da distorção de tempo, tal que uma descontinuidade em um dos reinícios seja evitada, reduzida ou eliminada em uma versão rescalada do contorno de distorção de tempo.

15 O método 400 adicionalmente compreende um passo 430 de fornecimento de uma representação de sinal de áudio decodificado com base na representação de sinal de áudio codificado usando a versão rescalada do contorno de distorção de tempo.

20 5. DESCRIÇÃO DETALHADA DE UMA CONFIGURAÇÃO DE ACORDO COM A INVENÇÃO TOMANDO REFERÊNCIA ÀS FIGURAS 5-9.

A seguir, será descrita uma configuração de acordo com a invenção em detalhe tomado referência às Figuras 5-9.

A Figura 5 mostra um diagrama esquemático de blocos de um aparelho 500 para fornecer uma informação de controle de distorção de tempo 512 com base em uma informação de evolução de contorno de distorção de tempo 510. O aparelho 500 compreende um meio 520 para fornecer um contorno reconstruído de informação de distorção de tempo 522 com base na informação de evolução de

contorno de distorção de tempo 510, e uma calculadora de informação de controle de distorção de tempo 530 para fornecer a informação de controle de distorção de tempo 512 com base no contorno reconstruído de informação de distorção de tempo 522.

5 **MEIO 520 PARA FORNECER O CONTORNO RECONSTRUÍDO DE
INFORMAÇÃO DE DISTORÇÃO DE TEMPO**

10 A seguir, serão descritas a estrutura e
funcionalidade do meio 520. O meio 520 compreende uma calculadora
de contorno de distorção de tempo 540, que é configurada para
receber a informação de evolução de contorno de distorção de tempo
510 e para fornecer, com base nela, uma nova informação de porção
de contorno de distorção 542. Por exemplo, um conjunto de
informações de evolução de contorno de distorção de tempo pode ser
transmitido para o aparelho 500 para cada estrutura do sinal de
15 áudio a ser reconstruído. Não obstante, um conjunto de informações
de evolução de contorno de distorção de tempo 510 associadas com
uma estrutura do sinal de áudio a ser reconstruído pode ser usado
para a reconstrução de uma pluralidade de estruturas do sinal de
áudio. Similarmente, uma pluralidade de conjuntos de informação de
20 evolução de contorno de distorção de tempo pode ser usada para a
reconstrução do conteúdo de áudio de uma única estrutura do sinal
de áudio, como será discutido em detalhes a seguir. Como
conclusão, pode-se afirmar que em algumas configurações, a
informação de evolução de contorno de distorção de tempo 510 pode
25 ser atualizada na mesma taxa em que conjuntos do coeficiente de
domínio de transformada do sinal de áudio a ser reconstruído ou
atualizado (uma porção de contorno de distorção de tempo por
estrutura do sinal de áudio).

A calculadora de contorno de distorção de tempo 540 compreende uma calculadora de valor de nó de distorção 544, que é configurada para computar uma pluralidade (ou sequência temporal) de valores de nó de contorno de distorção com base em uma pluralidade (ou sequência temporal) de valores de proporção de contorno de distorção de tempo (ou índices de proporção de distorção de tempo), em que os valores de proporção de distorção de tempo (ou índices) são compostos pela informação de evolução de contorno de distorção de tempo 510. Para este propósito, a calculadora de valor de nó de distorção 544 é configurada para iniciar o fornecimento dos valores de nó de contorno de distorção de tempo em um valor inicial predeterminado (por exemplo, 1) e para calcular valores subsequentes de nó de contorno de distorção de tempo usando os valores de proporção de contorno de distorção de tempo, como será discutido abaixo.

Além disso, a calculadora de contorno de distorção de tempo 540 opcionalmente compreende um interpolador 548 que é configurado para interpolar entre valores subsequentes de nó de contorno de distorção de tempo. Da mesma forma, a descrição 542 da nova porção de contorno de distorção de tempo é obtida, em que a nova porção de contorno de distorção de tempo tipicamente começa a partir do valor inicial predeterminado usado pela calculadora de valor de nó de distorção 524. Além disso, o meio 520 é configurado para considerar porções adicionais de contorno de distorção de tempo, a saber uma referida "última porção de contorno de distorção de tempo" e uma referida "porção de contorno de distorção de tempo atual" para o fornecimento de uma seção completa de contorno de distorção de tempo. Para este

propósito, o meio 520 é configurado para armazenar a referida "última porção de contorno de distorção de tempo" e a referida "porção de contorno de distorção de tempo atual" em uma memória não mostrada na Figura 5.

5 Entretanto, o meio 520 também compreende um rescalador 550, que é configurado para rescalar a "última porção de contorno de distorção de tempo" e a "porção de contorno de distorção de tempo atual" para evitar (ou reduzir, ou eliminar) quaisquer descontinuidades na seção completa de contorno de
10 distorção de tempo, que se baseia na "última porção de contorno de distorção de tempo", na "porção de contorno de distorção de tempo atual" e na "nova porção de contorno de distorção de tempo". Para este propósito, o rescalador 550 é configurado para receber a descrição armazenada da "última porção de contorno de distorção de
15 tempo" e da "porção de contorno de distorção de tempo atual" e conjuntamente rescalar a "última porção de contorno de distorção de tempo" e a "porção de contorno de distorção de tempo atual", para obter versões rescaladas da "última porção de contorno de distorção de tempo" e da "porção de contorno de distorção de tempo atual".
20 Detalhes relativos à rescalação realizada pelo rescalador 550 serão discutidos abaixo, tomando referência às Figuras 7a, 7b e 8.

 Além disso, o rescalador 550 pode também ser configurado para receber, por exemplo de uma memória não mostrada na Figura 5, um valor de soma associado com a "última porção de
25 contorno de distorção de tempo" e outro valor de soma associado com a "porção de contorno de distorção de tempo atual". Estes valores de soma são às vezes chamados de "last_warp_sum" e "cur_warp_sum", respectivamente. O rescalador 550 é configurado

para rescalarmos os valores de soma associados com as porções de contorno de distorção de tempo usando o mesmo fator de rescalação com que as correspondentes porções de contorno de distorção de tempo são rescaladas. Da mesma forma, valores de soma rescalados
5 são obtidos.

Em alguns casos, o meio 520 pode compreender um atualizador 560, que é configurado para repetidamente atualizar a entrada de porções de contorno de distorção de tempo no rescalador 550 e também a entrada de valores de soma no rescalador 550. Por exemplo, o atualizador 560 pode ser configurado para atualizar a referida informação na taxa de estrutura. Por exemplo, a "nova porção de contorno de distorção de tempo" do presente ciclo de estrutura pode servir como a "porção de contorno de distorção de tempo atual" em um próximo ciclo de estrutura. Similarmente, a
10 "porção de contorno de distorção de tempo atual" rescalada do atual ciclo de estrutura pode servir como a "última porção de contorno de distorção de tempo" em um próximo ciclo de estrutura. Da mesma forma, cria-se uma implementação eficiente de memória, porque a "última porção de contorno de distorção de tempo" do
15 atual ciclo de estrutura pode ser descartado na conclusão do atual ciclo de estrutura.

Para resumir o exposto acima, o meio 520 é configurado para prover, para cada ciclo de estrutura (com exceção de alguns ciclos de estrutura especiais, por exemplo no início de uma sequência de estrutura, ou no final de uma sequência de
25 estrutura, ou em uma estrutura na qual a distorção de tempo é inativa) uma descrição de uma seção de contorno de distorção de tempo compreendendo uma descrição de uma "nova porção de contorno

de distorção de tempo", de uma "porção rescalada de contorno de distorção de tempo atual" e de uma "última porção rescalada de contorno de distorção de tempo". Além disso, o meio 520 pode fornecer, para cada ciclo de estrutura (com exceção do ciclo de estrutura especial mencionado acima) uma representação de valores de soma de contorno de distorção, por exemplo, compreendendo um "valor de soma da nova porção de contorno de distorção de tempo", um "valor de soma de contorno de distorção de tempo atual rescalado" e um "valor de soma do último contorno de distorção de tempo rescalado".

A calculadora de informação de controle de distorção de tempo 530 é configurada para calcular a informação de controle de distorção de tempo 512 com base no contorno reconstruído de informação de distorção de tempo fornecido pelo meio 520. Por exemplo, a calculadora de informação de controle de distorção de tempo compreende uma calculadora de contorno de tempo 570, que é configurada para computar um contorno de tempo 572 com base na informação reconstruída de controle de distorção de tempo. Além disso, a calculadora de contorno de informação de distorção de tempo 530 compreende uma calculadora de posição de amostra 574, que é configurada para receber o contorno de tempo 572 e para fornecer, com base nele, uma informação de posição de amostra, por exemplo na forma de um vetor de posição de amostra 576. O vetor de posição de amostra 576 descreve a distorção de tempo realizada, por exemplo, pelo reamostrador 218.

A calculadora de informação de controle de distorção de tempo 530 também compreende uma calculadora de comprimento de transição, que é configurada para derivar uma

informação de comprimento de transição a partir da informação reconstruída de controle de distorção de tempo. A informação de comprimento de transição 582 pode, por exemplo, compreender uma informação descrevendo um comprimento de transição esquerda e uma
5 informação descrevendo um comprimento de transição direita. O comprimento de transição pode, por exemplo, depender de um comprimento de segmentos de tempo descritos pela "última porção de contorno de distorção de tempo", pela "porção de contorno de distorção de tempo atual" e pela "nova porção de contorno de
10 distorção de tempo". Por exemplo, o comprimento de transição pode ser encurtada (quando comparada com um comprimento de transição padrão) se a extensão temporal de um segmento de tempo descrito pela "última porção de contorno de distorção de tempo" for mais curta do que uma extensão temporal do segmento de tempo descrito
15 pela "porção de contorno de distorção de tempo atual", ou se a extensão temporal de um segmento de tempo descrito pela "nova porção de contorno de distorção de tempo" for mais curta do que a extensão temporal do segmento de tempo descrito pela "porção de contorno de distorção de tempo atual".

20 Além disso, a calculadora de informação de controle de distorção de tempo 530 pode adicionalmente compreender uma calculadora de primeira e última posição 584, que é configurada para calcular uma referida "primeira posição" e uma referida "última posição" com base no comprimento de transição
25 esquerda e direita. A "primeira posição" e a "última posição" aumentam a eficiência do reamostrador, pois regiões fora destas posições são idênticas a zero após aplicação da função janela e portanto não é necessário levá-las em conta para a distorção de

tempo. Deve-se notar aqui que o vetor de posição de amostra 576 compreende, por exemplo, informação exigida pela distorção de tempo realizada pelo reamostrador 280. Além disso, o comprimento de transição esquerda e direita 582 e a "primeira posição" e "última posição" 586 constituem informação, que é, por exemplo, exigida pelo windower 216.

Da mesma forma, pode-se dizer que o meio 520 e a calculadora de informação de controle de distorção de tempo 530 podem juntas assumir a funcionalidade do ajuste de taxa de amostra 220, do ajuste de forma de janela 210 e do cálculo de posição de amostragem 219.

A seguir, será descrita a funcionalidade de um decodificador de áudio que compreende o meio 520 e a calculadora de informação de controle de distorção de tempo 530 com referência às Figuras 6, 7a, 7b, 8, 9a-9c, 10a-10g, 11a, 11b e 12.

A Figura 6 mostra um fluxograma de um método para decodificar uma representação codificada de um sinal de áudio, de acordo com uma configuração da invenção. O método 600 compreende o fornecimento de um contorno reconstruído de informação de distorção de tempo, em que o fornecimento do contorno reconstruído de informação de distorção de tempo compreende o cálculo 610 de valores de nó de distorção, interpolação 620 entre os valores de nó de distorção e rescalação 630 de um ou mais porções de contorno de distorção previamente calculadas e um ou mais valores de soma de contorno de distorção previamente calculados. O método 600 adicionalmente compreende o cálculo 640 da informação de controle de distorção de tempo usando uma "nova porção de contorno de distorção de tempo" obtida nos passos 610 e 620, as porções

rescaladas de contorno de distorção de tempo previamente calculadas ("porção de contorno de distorção de tempo atual" e "última porção de contorno de distorção de tempo") e também, opcionalmente, usando os valores rescalados de soma de contorno de distorção previamente calculados. Como resultado, uma informação de contorno de tempo, e/ou uma informação de posição de amostra, e/ou uma informação de comprimento de transição e/ou uma informação de primeira porção e de última posição podem ser obtidas no passo 640.

10 O método 600 adicionalmente compreende a realização 650 da reconstrução do sinal de tempo distorcido usando a informação de controle de distorção de tempo obtida no passo 640. Detalhes relativos à reconstrução de sinal de distorção de tempo serão descritos subsequentemente.

15 O método 600 também compreende um passo 660 de atualização de uma memória, como será descrito abaixo.

CÁLCULO DAS PORÇÕES DE CONTOURNO DE DISTORÇÃO DE TEMPO

20 A seguir, serão descritos detalhes relativos ao cálculo das porções de contorno de distorção de tempo, tomando referência às Figuras 7a, 7b, 8, 9a, 9b, 9c.

25 Será assumido que um estado inicial está presente, que é ilustrado em uma representação gráfica 710 da Figura 7a. Como pode ser visto, uma primeira porção de contorno de distorção 716 (porção de contorno de distorção 1) e uma segunda porção de contorno de distorção 718 (porção de contorno de distorção 2) estão presentes. Cada uma das porções de contorno de distorção tipicamente compreende uma pluralidade de valores

discretos de dados de contorno de distorção, que são tipicamente armazenados em uma memória. Os diferentes valores de dados de contorno de distorção estão associados com valores de tempo, em que um tempo é mostrado em uma abscissa 712. Uma magnitude dos valores de dados de contorno de distorção é mostrada em uma ordenada 714. Como pode ser visto, a primeira porção de contorno de distorção tem um valor final de 1, e a segunda porção de contorno de distorção tem um valor inicial de 1, em que o valor de 1 pode ser considerado como um "valor predeterminado". Deve-se notar que a primeira porção de contorno de distorção 716 pode ser considerada como uma "última porção de contorno de distorção de tempo" (também chamada de "last_warp_contour"), enquanto a segunda porção de contorno de distorção 718 pode ser considerada como uma "porção de contorno de distorção de tempo atual" (também chamada de "cur_warp_contour").

Começando a partir do estado inicial, uma nova porção de contorno de distorção é calculada, por exemplo, nos passos 610, 620 do método 600. Da mesma forma, valores de dados de contorno de distorção da terceira porção de contorno de distorção (também chamada de "porção de contorno de distorção 3" ou "nova porção de contorno de distorção de tempo" ou "new_warp_contour") é calculada. O cálculo pode, por exemplo, ser separado em um cálculo de valores de nó de distorção, de acordo com um algoritmo 910 mostrado na Figura 9a, e uma interpolação 620 entre os valores de nó de distorção, de acordo com um algoritmo 920 mostrado na Figura 9a. Da mesma forma, uma nova porção de contorno de distorção 722 é obtida, que começa a partir do valor predeterminado (por exemplo, 1) e que é mostrado em uma representação gráfica 720 da Figura 7a.

Como pode ser visto, a primeira porção de contorno de distorção de tempo 716, a segunda porção de contorno de distorção de tempo 718 e a terceira nova porção de contorno de distorção de tempo estão associadas com intervalos de tempo subsequentes e contíguos. Além disso, pode-se ver que há uma descontinuidade 724 entre um ponto final 718b da segunda porção de contorno de distorção de tempo 718 e um ponto inicial 722a da terceira porção de contorno de distorção de tempo.

Deve-se notar aqui que a descontinuidade 724 tipicamente compreende uma magnitude que é maior do que uma variação entre quaisquer dois valores de dados de contorno de distorção do contorno de distorção de tempo temporalmente adjacentes dentro de uma porção de contorno de distorção de tempo. Isso se deve ao fato de que o valor inicial 722a da terceira porção de contorno de distorção de tempo 722 é forçado ao valor predeterminado (e.g. 1), independente do valor final 718b da segunda porção de contorno de distorção de tempo 718. Deve-se notar que a descontinuidade 724 é portanto maior do que a variação inevitável entre dois valores de dados de contorno de distorção adjacentes, discretos.

Não obstante, esta descontinuidade entre a segunda porção de contorno de distorção de tempo 718 e a terceira porção de contorno de distorção de tempo 722 seria prejudicial para o uso posterior dos valores de dados de contorno de distorção de tempo.

Da mesma forma, a primeira porção de contorno de distorção de tempo e a segunda porção de contorno de distorção de tempo são conjuntamente rescaladas no passo 630 do método 600. Por

exemplo, os valores de dados de contorno de distorção de tempo da primeira porção de contorno de distorção de tempo 716 e os valores de dados de contorno de distorção de tempo da segunda porção de contorno de distorção de tempo 718 são rescalados por multiplicação com um fator de rescalação (também chamado de "norm_fac"). Da mesma forma, uma versão rescalada 716' da primeira porção de contorno de distorção de tempo 716 é obtida, e também uma versão rescalada 718' da segunda porção de contorno de distorção de tempo 718 é obtida. Em contraste, a terceira porção de contorno de distorção de tempo é tipicamente não afetada neste passo de rescalação, como pode ser visto em uma representação gráfica 730 da Figura 7a. A rescalação pode ser realizada tal que o ponto final rescalado 718b' compreende, pelo menos aproximadamente, os mesmos valores de dados que o ponto inicial 722a da terceira porção de contorno de distorção de tempo 722. Da mesma forma, a versão rescalada 716' da primeira porção de contorno de distorção de tempo, a versão rescalada 718' da segunda porção de contorno de distorção de tempo e a terceira porção de contorno de distorção de tempo 722 juntas formam uma seção de contorno de distorção de tempo (aproximadamente) contínua. Em particular, a rescalação pode ser realizada tal que uma diferença entre os valores de dados do ponto final rescalado 718b' e o ponto inicial 722a não é maior do que um máximo da diferença entre quaisquer dois valores de dados adjacentes das porções de contorno de distorção de tempo 716', 718', 722.

Da mesma forma, a seção aproximadamente contínua de contorno de distorção de tempo compreendendo as porções rescaladas de contorno de distorção de tempo 716', 718' e a porção

original de contorno de distorção de tempo 722 é usada para o cálculo da informação de controle de distorção de tempo, que é realizada no passo 640. Por exemplo, a informação de controle de distorção de tempo pode ser computada para uma estrutura de áudio temporalmente associada com a segunda porção de contorno de distorção de tempo 718.

Entretanto, no cálculo da informação de controle de distorção de tempo no passo 640, uma reconstrução de sinal de tempo distorcido pode ser realizada em um passo 650, que será explicado em mais detalhes abaixo.

Subsequentemente, é necessário obter a informação de controle de distorção de tempo para uma próxima estrutura de áudio. Para este propósito, a versão rescalada 716' da primeira porção de contorno de distorção de tempo pode ser descartada para poupar memória, porque não é mais necessária. Entretanto, a versão rescalada 716' pode naturalmente também ser salva para qualquer propósito. Mais ainda, a versão rescalada 718' da segunda porção de contorno de distorção de tempo toma o lugar da "última porção de contorno de distorção de tempo" para o novo cálculo, como pode ser visto em uma representação gráfica 740 da Figura 7b. Depois, a terceira porção de contorno de distorção de tempo 722, que tomou o lugar da "nova porção de contorno de distorção de tempo" no cálculo anterior, assume o papel da "porção de contorno de distorção de tempo atual" para um próximo cálculo. A associação é mostrada na representação gráfica 740.

Subsequente a esta atualização da memória (passo 660 do método 600), uma nova porção de contorno de distorção de tempo 752 é calculada, como pode ser visto na representação

gráfica 750. Para este propósito, os passos 610 and 620 do método 600 podem ser re-executados com novos dados de entrada. A quarta porção de contorno de distorção de tempo 752 assume o papel da "nova porção de contorno de distorção de tempo" agora. Como pode ser visto, há tipicamente uma descontinuidade entre um ponto final 722b da terceira porção de contorno de distorção de tempo e um ponto inicial 752a da quarta porção de contorno de distorção de tempo 752. Esta descontinuidade 754 é reduzida ou eliminada por uma subsequente rescalação (passo 630 do método 600) da versão rescalada 718' da segunda porção de contorno de distorção de tempo e da versão original da terceira porção de contorno de distorção de tempo 722. Da mesma forma, uma versão rescalada duas vezes 718'' da segunda porção de contorno de distorção de tempo e uma versão rescalada uma vez 722' da terceira porção de contorno de distorção de tempo são obtidas, como pode ser visto a partir de uma representação gráfica 760 da Figura 7b. Como pode ser visto, as porções de contorno de distorção de tempo 718'', 722', 752 formam uma seção de contorno de distorção de tempo pelo menos aproximadamente contínua, que pode ser usada para o cálculo de informação de controle de distorção de tempo em uma re-execução do passo 640. Por exemplo, uma informação de controle de distorção de tempo pode ser calculada com base nas porções de contorno de distorção de tempo 718'', 722', 752, cuja informação de controle de distorção de tempo é associada a uma estrutura de tempo de sinal de áudio centrado na segunda porção de contorno de distorção de tempo.

Deve-se notar que em alguns casos é desejável ter um valor de soma de contorno de distorção associado para cada uma das porções de contorno de distorção de tempo. Por exemplo, um

primeiro valor de soma de contorno de distorção pode ser associado com a primeira porção de contorno de distorção de tempo, um segundo valor de soma de contorno de distorção pode ser associado com a segunda porção de contorno de distorção de tempo, e assim por diante. Os valores de soma de contorno de distorção podem, por exemplo, ser usados para o cálculo da informação de controle de distorção de tempo no passo 640.

Por exemplo, o valor de soma de contorno de distorção pode representar uma soma dos valores de dados de contorno de distorção de uma respectiva porção de contorno de distorção de tempo. Entretanto, como as porções de contorno de distorção de tempo são escaladas, é às vezes desejável também escalar o valor de soma de contorno de distorção de tempo, tal que o valor de soma de contorno de distorção de tempo siga a característica de sua porção de contorno de distorção de tempo associada. Da mesma forma, um valor de soma de contorno de distorção associado com a segunda porção de contorno de distorção de tempo 718 pode ser escalado (por exemplo, pelo mesmo fator de escalação) quando a segunda porção de contorno de distorção de tempo 718 é escalada para obter a versão escalada 718' desta. Similarmente, o valor de soma de contorno de distorção associado com a primeira porção de contorno de distorção de tempo 716 pode ser escalado (por exemplo, com o mesmo fator de escalação) quando a primeira porção de contorno de distorção de tempo 716 é escalada para obter a versão escalada 716' desta, se desejado.

Além disso, uma re-associação (ou re-alocação de memória) pode ser realizada ao proceder à consideração de uma nova porção de contorno de distorção de tempo. Por exemplo, o valor de

soma de contorno de distorção associado com a versão escalada 718' da segunda porção de contorno de distorção de tempo, que assume o papel de um "valor atual de soma de contorno de distorção de tempo" para o cálculo da informação de controle de distorção de tempo associada com as porções de contorno de distorção de tempo 716', 718', 722 pode ser considerado como um "último valor de soma de distorção de tempo" para o cálculo de uma informação de controle de distorção de tempo associada com as porções de contorno de distorção de tempo 718'', 722', 752. Similarmente, o valor de soma de contorno de distorção associado com a terceira porção de contorno de distorção de tempo 722 pode ser considerado como um "novo valor de soma de contorno de distorção" para o cálculo da informação de controle de distorção de tempo associada com as porções de contorno de distorção de tempo 716', 718', 722 e pode ser mapeado para agir como "atual valor de soma de contorno de distorção" para o cálculo da informação de controle de distorção de tempo associada com as porções de contorno de distorção de tempo 718'', 722', 752. Além disso, o recém calculado valor de soma de contorno de distorção da quarta porção de contorno de distorção de tempo 752 pode assumir o papel do "novo valor de soma de contorno de distorção" para o cálculo da informação de controle de distorção de tempo associada com as porções de contorno de distorção de tempo 718'', 722', 752.

EXEMPLO DE ACORDO COM A FIGURA 8

A Figura 8 mostra uma representação gráfica ilustrando um problema que é resolvido pelas configurações de acordo com a invenção. Uma primeira representação gráfica 810 mostra uma evolução temporal de um passo relativo reconstruído ao

longo do tempo, que é obtido em algumas configurações convencionais. Uma abscissa 812 descreve o tempo, uma ordenada 814 descreve o passo relativo. Uma curva 816 mostra a evolução temporal do passo relativo ao longo do tempo, que poderia ser

5 reconstruída a partir da informação de passo relativo. No que concerne à reconstrução do contorno de passo relativo, deve-se notar que para a aplicação da transformada de cossenos discreta modificada (MDCT) de tempo distorcido apenas o conhecimento da variação relativa do passo dentro da estrutura real é necessário.

10 A fim de entender isto, faz-se referência aos passos de cálculo para obter o contorno de tempo a partir do contorno de passo relativo, que leva a um contorno de tempo idêntico para versões escaladas do mesmo contorno de passo relativo. Portanto, é suficiente apenas codificar o valor de passo relativo ao invés de

15 um valor absoluto, que aumenta a eficiência da codificação. Para adicionalmente aumentar a eficiência, o valor quantizado real não é o passo relativo mas a mudança relativa no passo, isto é, a proporção do atual passo relativo sobre o passo relativo anterior (como será discutido em detalhe a seguir). Em algumas estruturas,

20 onde, por exemplo, o sinal não exibe estrutura harmônica alguma, não poderia ser desejada distorção de tempo. Em tais casos, um *flag* adicional pode opcionalmente indicar um contorno plano de passo ao invés de codificar este contorno plano com o método mencionado anteriormente. Como em sinais do mundo real a

25 quantidade de tais estruturas é tipicamente alta suficiente, o balanço entre o bit adicional adicionado em todas as vezes e os bits salvos para estruturas não distorcidas está a favor da economia de bits.

O valor inicial para o cálculo da variação de passo (contorno de passo relativo, ou contorno de distorção de tempo) pode ser escolhido arbitrariamente ou até diferir no codificador e decodificador. Devido à natureza da MDCT de tempo
5 distorcido (TW-MDCT) diferentes valores iniciais da variação de passo ainda dão as mesmas posições de amostra e formas de janela adaptadas para realizar a TW-MDCT.

Por exemplo, um codificador (de áudio) obtém um contorno de passo para cada nó que é expresso como atraso de passo
10 real em amostras em conjunto com uma especificação vocalizado/não vocalizado opcional, que era, por exemplo, obtida aplicando-se uma estimativa de passo e decisão vocalizado/não vocalizado que se sabia oriunda de codificação da fala. Se para o atual nó a classificação é estabelecida vocalizado, ou não há decisão
15 vocalizado/não vocalizado disponível, o codificador calcula a proporção entre o atraso de passo real e o quantiza, ou apenas estabelece a proporção para 1 se não vocalizado. Outro exemplo poderia ser que a variação de passo é estimada diretamente por um método apropriado (por exemplo, estimativa de variação de sinal).

20 No decodificador, o valor inicial para o primeiro passo relativo no início do áudio codificado é estabelecido para um valor arbitrário, por exemplo para 1. Portanto, o contorno de passo relativo decodificado não está mais na mesma variação absoluta do codificador de contorno de passo, porém uma versão
25 escalada deste. Ainda, como descrito acima, o algoritmo da TW-MDCT leva às mesmas posições de amostra e formas de janela. Além disso, o codificador poderia decidir, se as proporções de passo codificadas resultarem em um contorno de passo plano, não enviar o

contorno totalmente codificado, mas ao invés disso estabelecer o *flag* *activePitchData* em 0, economizando bits nesta estrutura (por exemplo, salvando bits *numPitchbits* * *numPitches* nesta estrutura).

A seguir, serão discutidos os problemas que ocorrem na ausência da renormalização do contorno de passo inventivo. Como mencionado acima, para a TW-MDCT, apenas a mudança de passo relativo dentro de um certo intervalo de tempo limitado em volta do bloco atual é necessária para o cômputo da distorção de tempo e a adaptação da forma de janela correta (vide explicações acima). A distorção de tempo segue o contorno decodificado para segmentos onde uma mudança de passo foi detectada, e permanece constante em todos os outros casos (vide a representação gráfica 810 da Figura 8). Para o cálculo de janela e posições de amostragem de um bloco, três segmentos consecutivos de contorno de passo relativo (por exemplo, três porções de contorno de distorção de tempo) são necessários, em que o terceiro é o recentemente transmitido na estrutura (designada como "nova porção de contorno de distorção de tempo") e os outros dois são amortecidos a partir do passado (por exemplo, designado como "última porção de contorno de distorção de tempo" e "porção de contorno de distorção de tempo atual").

Para ter um exemplo, faz-se referência, por exemplo, às explicações que foram feitas com referência às Figuras 7a e 7b, e também às representações gráficas 810, 860 da Figura 8. Para calcular, por exemplo, as posições de amostragem da janela para (ou associada com) a estrutura 1, que se estende da estrutura 0 para a estrutura 2, os contornos de passo de (ou associados com) a estrutura 0, 1 e 2 são necessários. No bit *stream*, apenas a

informação de passo para a estrutura 2 é enviada na atual estrutura, e as outras duas são tomadas a partir do passado. Como explicado aqui, o contorno de passo pode ser continuado aplicando-se a primeira proporção de passo relativo decodificada ao último

5 passo da estrutura 1 para obter o passo no primeiro nó da estrutura 2, e assim por diante. Agora é possível, devido à natureza do sinal, que se o contorno de passo for simplesmente continuado (isto é, se a parte recém transmitida do contorno for anexada às duas partes existentes sem qualquer modificação), que

10 um estouro de variação no formato de número interno do codificador ocorra após um certo tempo. Por exemplo, um sinal poderia começar com um segmento de características harmônicas fortes e um valor de passo alto no início que é decrescente ao longo de todo o segmento, levando a um passo relativo decrescente. Então, um

15 segmento sem informação de passo pode seguir, a fim de que o passo relativo mantenha-se constante. Então novamente, uma seção harmônica pode começar com um passo absoluto que é maior do que o último passo absoluto do segmento anterior, e novamente seguindo para baixo. Entretanto, se simplesmente continua-se o passo

20 relativo, é o mesmo que no final do último segmento harmônico e irá baixar mais ainda, e assim por diante. Se o sinal é forte o suficiente e tem em seus segmentos harmônicos uma tendência global de subir ou descer (como mostrado na representação gráfica 810 da Figura 8), cedo ou tarde o passo relativo alcança a fronteira de

25 uma gama do formato de número interno. É bem conhecido a partir de codificação da fala que sinais da voz de fato exibem tal característica. Portanto não é surpresa, que a codificação de um conjunto concatenado de sinais do mundo real incluindo a voz de

fato excederam a gama dos valores de flutuação usados para o passo relativo após um tempo relativamente curto ao se usar o método convencional descrito acima.

Para resumir, para um segmento (ou estrutura) de
5 sinal de áudio para o qual um passo pode ser determinado, poderia ser determinada uma evolução apropriada do contorno de passo relativo (ou contorno de distorção de tempo). Para segmentos de
10 sinal de áudio (ou sinais de estrutura de áudio) para os quais um passo não possa ser determinado (por exemplo, porque os segmentos de sinal de áudio são semelhantes a ruídos) o contorno de passo relativo (ou contorno de distorção de tempo) poderia ser mantido constante. Da mesma forma, se houvesse um desequilíbrio entre
15 segmentos de áudio com passo crescente e passo decrescente, o contorno de passo relativo (ou contorno de distorção de tempo) tenderia a um *underflow* numérico ou estouro numérico.

Por exemplo, na representação gráfica 810 um
contorno de passo relativo é mostrado para o caso em que há uma pluralidade de porções de contorno de passo relativo 820a, 820a,
820c, 820d com passo decrescente e alguns segmentos de áudio 822a,
20 822b sem passo, porém sem segmentos de áudio com passo crescente. Da mesma forma, pode-se ver que o contorno de passo relativo 816 tende a um *underflow* numérico (pelo menos sob circunstâncias muito adversas).

A seguir, será descrita uma solução para este
25 problema. Para prevenir os problemas supracitados, em particular o *underflow* ou estouro numérico, uma renormalização periódica do contorno de passo relativo foi introduzida de acordo com um aspecto da invenção. Como o cálculo do contorno de tempo

distorcido e as formas de janela confiam apenas nas mudanças relativas ao longo dos três segmentos de contorno de passo relativo supracitados (também designados como "porções de contorno de distorção de tempo"), como explicado aqui, é possível
5 normalizar este contorno (por exemplo, o contorno de distorção de tempo, que pode ser composto de três partes de "porções de contorno de distorção de tempo") para cada estrutura (por exemplo, do sinal de áudio) de uma nova maneira com o mesmo resultado.

Para isto, a referência foi, por exemplo,
10 escolhida para ser a última amostra do segundo segmento de contorno (também designado como "porção de contorno de distorção de tempo"), e o contorno está agora normalizado (por exemplo, multiplicativamente no domínio linear) de tal forma que esta amostra tem um valor de 1,0 (vide a representação gráfica 860 da
15 Figura 8).

A representação gráfica 860 da Figura 8 representa a normalização do contorno de passo relativo. Uma abscissa 862 mostra o tempo, subdividido em estruturas (estruturas 0, 1, 2). Uma ordenada 864 descreve o valor do contorno de passo
20 relativo.

Um contorno de passo relativo antes da normalização é designado com 870 e cobre duas estruturas (por exemplo, estrutura número 0 e estrutura número 1). Um novo segmento de contorno de passo relativo (também designado como
25 "porção de contorno de distorção de tempo") começando a partir do valor inicial predeterminado do contorno de passo relativo (ou valor inicial de contorno de distorção de tempo) é designado com 874. Como pode ser visto, o reinício do novo segmento de contorno

de passo relativo 874 a partir do valor inicial predeterminado do contorno de passo relativo (por exemplo, 1) traz uma descontinuidade entre o segmento de contorno de passo relativo 870 que precede o ponto temporal de reinício e o novo segmento de

5 contorno de passo relativo 874, que é designado com 878. Esta descontinuidade traria um grave problema para a derivação de qualquer informação de controle de distorção de tempo a partir do contorno e possivelmente resultará em distorções de áudio. Portanto, um segmento de contorno de passo relativo 870

10 previamente obtido que precede o ponto temporal de reinício é rescalado (ou normalizado), para obter um segmento rescalado de contorno de passo relativo 870'. A normalização é realizada tal que a última amostra do segmento de contorno de passo relativo 870 é escalada para um valor inicial predeterminado do contorno de

15 passo relativo (por exemplo, de 1,0).

DESCRIÇÃO DETALHADA DO ALGORITMO

A seguir, serão descritos em detalhe alguns dos algoritmos realizados por um decodificador de áudio de acordo com uma configuração da invenção. Para este propósito, serão feitas

20 referências às Figuras 5, 6, 9a, 9b, 9c e 10a-10g. Adicionalmente, faz-se referência à legenda de elementos de dados, elementos de ajuda e constantes das Figuras 11a e 11b..

Falando de forma geral, pode-se dizer que o método descrito aqui pode ser usado para decodificar um *stream* de

25 áudio que esteja codificada de acordo com uma transformada de cossenos discreta modificada de tempo distorcido. Assim, quando a TW-MDCT é habilitada para o *stream* de áudio (que pode ser indicada por um *flag*, por exemplo chamado de *flag* "twMdct", que pode ser

composta em uma informação de configuração específica), um banco de filtros de tempo distorcido e troca de bloco podem substituir um banco de filtros padrão e troca de bloco. Adicionalmente à transformada inversa de cossenos discreta modificada (IMDCT), o

5 banco de filtros de tempo distorcido e troca de bloco contém um domínio de tempo para mapeamento de domínio de tempo a partir de uma grade de tempo arbitrariamente espaçado para a grade de tempo regularmente espaçado normal e uma adaptação correspondente de formas de janela.

10 A seguir, será descrito o processo de codificação. Em um primeiro passo, o contorno de distorção é decodificado. O contorno de distorção pode ser, por exemplo, codificado usando índices de livro de código de nós de contorno de

15 distorção. Os índices de livro de código dos nós de contorno de distorção são decodificados, por exemplo, usando o algoritmo mostrado em uma representação gráfica 910 da Figura 9a. De acordo com o referido algoritmo, valores de proporção de distorção (*warp_value_tbl*) são derivados a partir de índices do livro de

20 códigos para proporção de distorção (*tw_ratio*), por exemplo usando um mapeamento definido por uma tabela de mapeamento 990 da Figura 9c. Como pode ser visto a partir do algoritmo mostrado como numeral de referência 910, os valores de nó de distorção podem ser estabelecidos para um valor constante predeterminado, se um *flag* (*tw_data_present*) indicar que dados de distorção de tempo não

25 estão presentes. Em contraste, se o *flag* indicar que dados de distorção de tempo estão presentes, um primeiro valor de nó de distorção pode ser estabelecido para o valor predeterminado inicial de contorno de distorção de tempo (por exemplo, 1).

Valores subsequentes de nó de distorção (de uma porção de contorno de distorção de tempo) podem ser determinados com base em uma formação de um produto de múltiplos valores de proporção de distorção de tempo. Por exemplo, um valor de nó de distorção de um

5 nó imediatamente após o primeiro nó de distorção ($i=0$) pode ser igual a um primeiro valor de proporção de distorção (se o valor inicial for 1) ou igual a um produto do primeiro valor de proporção de distorção e o valor inicial. Valores subsequentes de nó de distorção de tempo ($i=2,3,\dots, \text{num_tw_nodes}$) são computados

10 formando-se um produto de múltiplos valores de proporção de distorção de tempo (opcionalmente levando em consideração o valor inicial, se o valor inicial diferir de 1). Naturalmente, a ordem da formação de produto é arbitrária. Entretanto, é vantajoso derivar um $(i+1)$ -ésimo valor de modo de distorção a partir de um

15 i -ésimo valor de nó de distorção multiplicando o i -ésimo valor de nó de distorção com um único valor de proporção de distorção descrevendo uma proporção entre dois valores subsequentes de nó do contorno de distorção de tempo...

Como pode ser visto a partir do algoritmo

20 mostrado no numeral de referência 910, pode haver múltiplos índices do livro de códigos para proporção de distorção para uma única porção de contorno de distorção de tempo ao longo de uma única estrutura de áudio (em que pode haver uma correspondência 1-a-1 entre porções de contorno de distorção de tempo e estruturas

25 de áudio).

Para resumir, uma pluralidade de valores de nó de distorção de tempo pode ser obtida para uma determinada porção de contorno de distorção de tempo (ou uma determinada estrutura de

áudio) no passo 610, por exemplo usando a calculadora de valor de
 nó de distorção 544. Subsequentemente, uma interpolação linear
 pode ser realizada entre os valores de nó de distorção de tempo
 (warp_node_values[i]). Por exemplo, para obter os valores de dados
 5 de contorno de distorção de tempo da "nova porção de contorno de
 distorção de tempo" (new_warp_contour) o algoritmo mostrado no
 numeral de referência 920 na Figura 9a pode ser usado. Por
 exemplo, o número de amostras da nova porção de contorno de
 distorção de tempo é igual a metade do número das amostras de
 10 domínio de tempo de uma transformada inversa de cossenos discreta
 modificada. Com relação a esta questão, deve-se notar que sinais
 de estrutura de áudio adjacentes são tipicamente trocados (pelo
 menos aproximadamente) por metade do número de amostras de domínio
 de tempo da MDCT ou IMDCT. Em outras palavras, para obter o modo
 15 de amostra (amostras N_long) new_warp_contour[], os
 warp_node_values[] são interpolados linearmente entre os nós
 igualmente espaçados (separação interp_dist) usando o algoritmo
 mostrado no numeral de referência 920.

A interpolação pode, por exemplo, ser realizada
 20 pelo interpolador 548 do aparelho da Figura 5, ou no passo 620 do
 algoritmo 600.

Antes de obter o completo contorno de distorção
 para esta estrutura (isto é para a estrutura presentemente sob
 consideração) os valores amortecidos a partir do passado são
 25 rescalados para que o último valor de distorção do
 past_warp_contour[] seja igual a 1 (ou qualquer outro valor
 predeterminado, que é preferivelmente igual ao valor inicial da
 nova porção de contorno de distorção de tempo).

Deve-se notar aqui que o termo "contorno de
 distorção passado" preferivelmente compreende a "última porção
 de contorno de distorção de tempo" descrita acima e a "porção
 de contorno de distorção de tempo atual" descrita acima. Deve-
 5 se também notar que o "contorno de distorção passado"
 tipicamente compreende uma extensão que é igual a um número de
 amostras de domínio de tempo da IMDCT, tal que valores do
 "contorno de distorção passado" são designados com índices
 entre 0 e $2 \cdot n_{\text{long}} - 1$. Assim, "past_warp_contour[$2 \cdot n_{\text{long}} - 1$]"
 10 designa um último valor de distorção do "contorno de distorção
 passado". Da mesma forma, um fator de normalização "norm_fac"
 pode ser calculado de acordo com uma equação mostrada no
 numeral de referência 930 na Figura 9a. Desta forma, o contorno
 de distorção passado (compreendendo a "última porção de
 15 contorno de distorção de tempo" e a "porção de contorno de
 distorção de tempo atual") pode ser multiplicativamente
 rescalado de acordo com a equação mostrada no numeral de
 referência 932 na Figura 9a. Além disso, o "último valor de
 soma de contorno de distorção" (last_warp_sum) e o "atual valor
 20 de soma de contorno de distorção" (cur_warp_sum) podem ser
 multiplicativamente rescalados, como mostrado nos numerais de
 referência 934 e 936 na Figura 9a. A rescalação pode ser
 realizada pelo rescalador 550 da Figura 5, ou no passo 630 do
 método 600 da Figura 6.

25 Deve-se notar que a normalização descrita aqui,
 por exemplo no numeral de referência 930, então poderia ser
 modificada, por exemplo, substituindo o valor inicial de "1" por
 qualquer outro valor predeterminado desejado.

Aplicando-se a normalização, um "full warp_contour[]" também designado como uma "seção de contorno de distorção de tempo" é obtido concatenando-se o "past_warp_contour" e o "new_warp_contour". Desta forma, três porções de contorno de distorção de tempo ("última porção de contorno de distorção de tempo", "porção de contorno de distorção de tempo atual", and "nova porção de contorno de distorção de tempo") formam o "contorno de distorção completo", que pode ser aplicado em passos adicionais do cálculo.

10 Além disso, um valor de soma de contorno de distorção (new_warp_sum) é calculado, por exemplo, como uma soma de todos os valores de "new_warp_contour[]". Por exemplo, um novo valor de soma de contorno de distorção pode ser calculado de acordo com os algoritmos mostrados no numeral de referência 940 na Figura 9a.

15 Após os cálculos descritos acima, a informação de entrada exigida pela calculadora de informação de controle de distorção de tempo 330 ou pelo passo 640 do método 600 está disponível. Da mesma forma, o cálculo 640 da informação de controle de distorção de tempo pode ser realizado, por exemplo
20 pela calculadora de informação de controle de distorção de tempo 530. Também, a reconstrução de sinal de tempo distorcido 650 pode ser realizada pelo decodificador de áudio. Ambos, o cálculo 640 e a reconstrução de sinal de tempo distorcido 650 serão explicados em maiores detalhes abaixo.

25 Entretanto, é importante notar que o presente algoritmo procede iterativamente. É portanto computacionalmente eficiente atualizar a memória. Por exemplo, é possível descartar informação sobre a última porção de contorno de distorção de tempo.

Adicionalmente, é recomendável usar a presente "porção de contorno de distorção de tempo atual" como uma "última porção de contorno de distorção de tempo" em um próximo ciclo de cálculo. Adicionalmente, é recomendável usar a presente "nova porção de contorno de distorção de tempo" como uma "porção de contorno de distorção de tempo atual" em um próximo ciclo de cálculo. Esta atribuição pode ser feita usando a equação mostrada no numeral de referência 950 na Figura 9b, (em que $\text{warp_contour}[n]$ descreve a presente "nova porção de contorno de distorção de tempo" para $2 \cdot n_{\text{long}} \leq n < 3 \cdot n_{\text{long}}$).

10 Atribuições apropriadas podem ser vistas nos numerais de referência 952 e 954 na Figura 9b.

Em outras palavras, buffers de memória usados para decodificar a próxima estrutura podem ser atualizados de acordo com as equações mostradas nos numerais de referência 950, 952 e 954.

Deve-se notar que a atualização de acordo com as equações 950, 952 e 954 não fornece um resultado razoável, se a informação apropriada não estiver sendo gerada para uma estrutura prévia. Da mesma forma, antes de decodificar a primeira estrutura ou se a última estrutura foi codificada com um tipo diferente de codificador (por exemplo, um codificador de domínio LPC) no contexto de um codificador trocado, os estados de memória podem ser estabelecidos de acordo com as equações mostradas nos numerais de referência 960, 962 e 964 da Figura 9b.

25 **CÁLCULO DE INFORMAÇÃO DE CONTROLE DE DISTORÇÃO DE TEMPO**

A seguir, será descrito brevemente como a informação de controle de distorção de tempo pode ser calculada

com base no contorno de distorção de tempo (compreendendo, por exemplo, três porções de contorno de distorção de tempo) e com base nos valores de soma de contorno de distorção.

Por exemplo, deseja-se reconstruir um contorno de tempo usando o contorno de distorção de tempo. Para este propósito, pode ser usado um algoritmo que é mostrado nos numerais de referência 1010, 1012 na Figura 10a. Como pode ser visto, o contorno de tempo mapeia um índice i ($0 \leq i \leq 3 \cdot n_long$) em um correspondente valor de contorno de tempo. Um exemplo de tal mapeamento é mostrado na Figura 12.

Com base no cálculo do contorno de tempo, é tipicamente necessário calcular uma posição de amostra (`sample_pos[]`), que descreve posições amostras de tempo 'distorcido em uma escala de tempo linear. Tal cálculo pode ser realizado usando um algoritmo, que é mostrado no numeral de referência 1030 na Figura 10b. No algoritmo 1030, podem ser usadas funções auxiliares, que são mostradas nos numerais de referência 1020 e 1022 na Figura 10a. Da mesma forma, uma informação sobre o tempo de amostra pode ser obtida.

Além disso, algumas durações de transições de tempo distorcido (`warped_trans_len_left`; `warped_trans_len_right`) são calculadas, por exemplo usando um algoritmo 1032 mostrado na Figura 10b. Opcionalmente, as durações de transição de distorção de tempo pode ser adaptada dependendo de um tipo de janela ou um comprimento de transformada, por exemplo usando um algoritmo mostrado no numeral de referência 1034 na Figura 10b. Além disso, uma assim chamada "primeira posição" e uma assim chamada "última posição" podem ser computadas com base nas informações de durações

de transição, por exemplo usando um algoritmo mostrado no numeral de referência 1036 na Figura 10b. Para resumir, será feito um ajuste de posições de amostra e durações de janela, que pode ser realizado pelo aparelho 530 ou no passo 640 do método 600. A partir de "warp_contour()", pode ser computado um vetor das posições de amostra ("sample_pos()") das amostras de tempo distorcido em uma escala de tempo linear. Para isto, primeiro o contorno de tempo pode ser gerado usando o algoritmo mostrado nos numerais de referência 1010, 1012. Com as funções auxiliares "warp_in_vec()" e "warp_time_inv()", que são mostradas nos numerais de referência 1020 e 1022, o vetor de posição de amostra ("sample_pos()") e as durações de transição ("warped_trans_len_left" e "warped_trans_len_right") são computados, por exemplo usando os algoritmos mostrados nos numerais de referência 1030, 1032, 1034 e 1036. Da mesma forma, a informação de controle de distorção de tempo 512 é obtida.

RECONSTRUÇÃO DE SINAL DE TEMPO DISTORCIDO

A seguir, a reconstrução de sinal de tempo distorcido, que pode ser realizada com base na informação de controle de distorção de tempo será brevemente discutida para colocar o cômputo do contorno de distorção de tempo no contexto próprio.

A reconstrução de um sinal de áudio compreende a execução de uma transformada inversa de cossenos discreta modificada, que não é descrita aqui em detalhe, porque é bem conhecida para qualquer técnico no assunto. A execução da transformada inversa de cossenos discreta modificada permite reconstruir amostras de domínio de tempo distorcido com base em um conjunto de coeficientes de domínio de frequência. A execução da

IMDCT pode, por exemplo, ser feita na forma de estrutura, que significa, por exemplo, que uma estrutura de 2048 amostras de domínio de tempo distorcido é reconstruída com base em um conjunto de 1024 coeficientes de domínio de frequência. Para a correta
 5 reconstrução é necessário que não mais do que duas janelas subsequentes se sobreponham. Devido à natureza da TW-MDCT poderia ocorrer que uma porção de tempo inversamente distorcido de uma estrutura se estenda até uma estrutura não-vizinha, desta forma violando o pré-requisito afirmado acima. Portanto a duração de
 10 desvanecimento da forma de janela precisa ser encurtada calculando-se os valores apropriados de `warped_trans_len_left` e `warped_trans_len_right` mencionados acima.

Uma janela e troca de bloco 650b é então aplicado às amostras de domínio de tempo obtidas a partir da IMDCT. A
 15 janela e troca de bloco podem ser aplicados às amostras de domínio de tempo distorcido fornecidas pela IMDCT 650a na dependência da informação de controle de distorção de tempo, para obter amostras de domínio de tempo distorcido em janela. Por exemplo, dependendo de uma informação, ou elemento, de `"window_shape"`, diferentes
 20 protótipos de janela de transformada sobreamostrada podem ser usados, em que o comprimento das janelas sobreamostradas pode ser dada pela equação mostrada no numeral de referência 1040 na Figura 10c. Por exemplo, para um primeiro tipo de forma de janela (por exemplo, `window_shape==1`), os coeficientes de janela são dados por
 25 uma janela derivada de "Kaiser-Bessel" (KBD) de acordo com a definição mostrada no numeral de referência 1042 na Figura 10c, em que W' , a "função de janela central de Kaiser-Bessel", é definida como mostrado no numeral de referência 1044 na Figura 10c.

De outra forma, ao usar uma forma de janela diferente (por exemplo, se `window_shape==0`), uma janela de seno pode ser empregada de acordo com a definição no numeral de referência 1046. Para todos os tipos de sequências de janela ("window_sequences"), o protótipo usado para a parte de janela esquerda é determinado pela forma de janela do bloco anterior. A fórmula mostrada no numeral de referência 1048 na Figura 10c expressa este fato. De maneira similar, o protótipo para forma de janela direita é determinado pela fórmula mostrada no numeral de referência 1050 na Figura 10c.

A seguir, será descrita a aplicação das janelas acima descritas às amostras de domínio de tempo distorcido fornecidas pela IMDCT. Em algumas configurações, a informação para uma estrutura pode ser fornecida por uma pluralidade de sequências curtas (por exemplo, oito sequências curtas). Em outras configurações, a informação para uma estrutura pode ser fornecida usando blocos de diferentes durações, em que um tratamento especial pode ser necessário para sequências de início, sequências de parada e/ou sequências de durações não-padrão. Entretanto, como o comprimento transicional pode ser determinada com descrito acima, pode ser suficiente diferenciar entre estruturas que foram codificadas usando oito sequências curtas (indicadas por uma informação de tipo de estrutura apropriada "eight_short_sequence") e todas as outras estruturas.

Por exemplo, em uma estrutura descrita por uma sequência curta de oito, um algoritmo mostrado como numeral de referência 1060 na Figura 10d pode ser aplicado para a aplicação da janela. Em contraste, para estruturas codificadas usando outras

informações, um algoritmo é mostrado no numeral de referência 1064 na Figura 10e pode ser aplicado. Em outras palavras, a porção semelhante a código C mostrada no numeral de referência 1060 na Figura 10d descreve a aplicação da janela e sobreposição-soma interna de uma assim chamada "oito sequências curtas". Em contraste, a porção semelhante a código C mostrada no numeral de referência 1064 na Figura 10d descreve a aplicação de janela em outros casos.

REAMOSTRAGEM

A seguir, será descrita a distorção de tempo inversa 650c das amostras de domínio de tempo distorcido em janela na dependência da informação de controle de distorção de tempo, pela qual amostras de domínio de tempo amostradas regularmente, ou simplesmente amostras de domínio de tempo, são obtidas por reamostragem variável com o tempo. Na reamostragem variável com o tempo, o bloco em janela $z[]$ é reamostrado de acordo com as posições amostradas, por exemplo usando uma resposta de impulso mostrada no numeral de referência 1070 na Figura 10f. Antes da reamostragem, o bloco em janela pode ser preenchido com zeros em ambas as extremidades, como mostrado no numeral de referência 1072 na Figura 10f. A reamostragem em si é descrita pela seção de pseudo código mostrada no numeral de referência 1074 na Figura 10f.

PROCESSAMENTO DE ESTRUTURA PÓS-REAMOSTRADOR

A seguir, será descrito um pós-processamento opcional 650d das amostras de domínio de tempo. Em algumas configurações, o processamento de estrutura pós-reamostragem pode ser realizado na dependência de um tipo da sequência de janela. Dependendo do parâmetro "window_sequence", certos passos de processamento adicionais podem ser aplicados.

Por exemplo, se a sequência de janela é uma assim chamada "EIGHT_SHORT_SEQUENCE", uma assim chamada "LONG_START_SEQUENCE", uma assim chamada "STOP_START_SEQUENCE", uma assim chamada "STOP_START_1152_SEQUENCE" seguida por uma assim chamada LPD_SEQUENCE, pode ser realizado um pós-processamento como mostrado nos numerais de referência 1080a, 1080b, 1082.

Por exemplo, se a próxima sequência de janela for uma assim chamada "LPD_SEQUENCE", uma janela de correção $W_{corr}(n)$ pode ser calculada como mostrada no numeral de referência 1080a, levando em conta as definições mostradas no numeral de referência 1080b. Também. A janela de correção $W_{corr}(n)$ pode ser aplicada como mostrado no numeral de referência 1082 na Figura 10g.

Para todos os outros casos, nada pode ser feito, como pode ser visto no numeral de referência 1084 na Figura 10g.

SOBREPOSIÇÃO E SOMA COM SEQUÊNCIAS DE JANELA ANTERIORES

Além disso, uma sobreposição-e-soma das amostras de domínio de tempo atuais com uma ou mais amostras anteriores de domínio de tempo pode ser realizada. A sobreposição e soma pode ser a mesma para todas as sequências e pode ser descrita matematicamente como mostrado no numeral de referência 1086 na Figura 10g.

LEGENDA

Com relação às explicações dadas, também se faz referência à legenda, que é mostrada nas Figuras 11a e 11d. Em particular, o comprimento de janela de síntese N para a transformada inversa é tipicamente uma função do elemento de sintaxe "window_sequence" e do contexto algorítmico. Pode ser por

exemplo definida como mostrado no numeral de referência 1190 da Figura 11b.

CONFIGURAÇÃO DE ACORDO COM A FIGURA 13

A Figura 13 mostra um diagrama esquemático de blocos de um meio 1300 para fornecer um contorno reconstruído de informação de distorção de tempo que assume a funcionalidade do meio 520 descrita com referência à Figura 5. Entretanto, os caminhos e buffers de dados são mostrados em maiores detalhes. O meio 1300 compreende uma calculadora de valor de nó de distorção 1344, que assume a função da calculadora de valor de nó de distorção 544. A calculadora de valor de nó de distorção 1344 recebe um índice de livro de código "tw_ratio[]" da proporção de distorção como uma informação de proporção de distorção codificada. A calculadora de valor de nó de distorção compreende uma tabela de valor de distorção representando, por exemplo, o mapeamento de um índice de proporção de distorção de tempo sobre um valor de proporção de distorção de tempo representado na Figura 9c. A calculadora de valor de nó de distorção 1344 pode adicionalmente compreender um multiplicador para realizar o algoritmo representado no numeral de referência 910 da Figura 9a. Da mesma forma, a calculadora de valor de nó de distorção fornece valores de nó de distorção "warp_node_values[i]". Adicionalmente, o meio 1300 compreende um interpolador de contorno de distorção 1348, que assume a função do interpolador 540a, e que pode ser figurado para realizar o algoritmo mostrado no numeral de referência 920 na Figura 9a, desta forma obtendo valores do novo contorno de distorção ("new_warp_contour"). O meio 1300 adicionalmente compreende um novo buffer de contorno de distorção

1350, que armazena os valores do novo contorno de distorção (isto é `warp_contour [i]`, com $2 \cdot n_long \leq i < 3 \cdot n_long$). O meio 1300 adicionalmente compreende um buffer/atualizador de contorno de distorção passado 1360, que armazena a "última porção de contorno de distorção de tempo" e a "porção de contorno de distorção de tempo atual" e atualiza os conteúdos da memória em resposta a uma rescalação e em resposta a uma conclusão do processamento da atual estrutura. Desta forma, o buffer/atualizador de contorno de distorção passado 1360 pode estar em cooperação com o rescalador de contorno de distorção passado 1370, tal que o buffer/atualizador de contorno de distorção passado e o rescalador de contorno de distorção passado juntos cumprem a funcionalidade dos algoritmos 930, 932, 934, 936, 950, 960. Opcionalmente, o buffer/atualizador de contorno de distorção passado 1360 pode também assumir a funcionalidade dos algoritmos 932, 936, 952, 954, 962, 964.

Desta forma, o meio 1300 fornece o contorno de distorção ("warp_contour") e otimamente também fornece os valores de soma de contorno de distorção.

20 **CODIFICADOR DE SINAL DE ÁUDIO DE ACORDO COM A**

FIGURA 14

A seguir, será descrito um codificador de sinal de áudio de acordo com um aspecto da invenção. O codificador de sinal de áudio da Figura 14 é designado em sua plenitude com 1400.

25 O codificador de sinal de áudio 1400 é configurado para receber um sinal de áudio 1410 e, opcionalmente, uma informação de contorno de distorção 1412 fornecida externamente e associada com o sinal de áudio 1410. Adicionalmente, o codificador de sinal de áudio

1400 é configurado para prover uma representação codificada 1440 do sinal de áudio 1410.

O codificador de sinal de áudio 1400 compreende um codificador de contorno de distorção de tempo 1420, configurado para receber um contorno de informação de distorção de tempo 1422 associado com o sinal de áudio 1410 e para fornecer um contorno de informação de distorção de tempo codificado 1424 com base nele.

O codificador de sinal de áudio 1400 adicionalmente compreende um processador de sinal de distorção de tempo (ou codificador de sinal de distorção de tempo) 1430 que é configurado para receber o sinal de áudio 1410 e para fornecer, com base nele, uma representação codificada de distorção de tempo 1432 do sinal de áudio 1410, levando em conta uma distorção de tempo descrita pela informação de distorção de tempo 1422. A representação codificada 1414 do sinal de áudio 1410 compreende o contorno de informação de distorção de tempo codificado 1424 e a representação codificada 1432 do espectro do sinal de áudio 1410.

Opcionalmente, o codificador de sinal de áudio 1400 compreende uma calculadora informação de contorno de distorção 1440, que é configurada para prover o contorno de informação de distorção de tempo 1422 com base no sinal de áudio 1410. Alternativamente, entretanto, o contorno de informação de distorção de tempo 1422 pode ser fornecido com base na informação de contorno de distorção fornecida externamente 1412.

O codificador de contorno de distorção de tempo 1420 pode ser configurado para computar uma proporção entre valores subsequentes de nó do contorno de distorção de tempo descrito pelo contorno de informação de distorção de tempo 1422.

Por exemplo, os valores de nó podem ser valores de amostra do contorno de distorção de tempo representados pelo contorno de informação de distorção de tempo. Por exemplo, se o contorno de informação de distorção de tempo compreende uma pluralidade de valores para cada estrutura do sinal de áudio 1410, os valores de nó de distorção de tempo podem ser um subconjunto verdadeiro deste contorno de informação de distorção de tempo. Por exemplo, os valores de nó de distorção de tempo podem ser um subconjunto verdadeiro periódico dos valores de contorno da distorção de tempo. Um valor de nó de contorno de distorção de tempo pode estar presente conforme N das amostras de áudio, em que N pode ser maior do que ou igual a 2.

A calculadora de proporção de valor de nó de contorno de tempo pode ser configurada para computar uma proporção entre valores subsequentes de nó de distorção de tempo do contorno de distorção de tempo, desta forma fornecendo uma informação que descreve uma proporção entre valores subsequentes de nó do contorno de distorção de tempo. Um codificador de proporção do codificador de contorno de distorção de tempo pode ser configurado para codificar a proporção entre valores subsequentes de nó do contorno de distorção de tempo. Por exemplo, o codificador de proporção pode mapear diferentes proporções para diferentes índices do livro de códigos. Por exemplo, um mapeamento pode ser escolhido tal que as proporções fornecidas pela calculadora de proporção de valor de contorno de distorção de tempo estejam dentro de uma faixa entre 0,9 e 1,1, ou mesmo entre 0,95 e 1,05. Da mesma forma, o codificador de proporção pode ser configurado para mapear esta faixa para diferentes índices de livro de código.

Por exemplo, as correspondências mostradas na tabela da Figura 9c podem agir como pontos de apoio neste mapeamento, tal que, por exemplo, uma proporção de 1 seja mapeada sobre um índice de livro de código de 3, enquanto uma proporção de 1,0057 seja mapeada para um índice de livro de código de 4, e assim por diante (comparar Figura 9c). Valores de proporção entre aqueles mostrados na tabela da Figura 9c podem ser mapeados para apropriados índices de livro de código, por exemplo para o índice de livro de código do mais próximo valor de proporção para o qual o índice de livro de código é mostrado na tabela da Figura 9c.

Naturalmente, diferentes codificações podem ser usadas tal que, por exemplo, um número de índices de livro de código disponíveis pode ser escolhido maior ou menor do que o mostrado aqui. Também, a associação entre valores de nó de contorno de distorção e índices dos valores do livro de códigos podem ser escolhidos apropriadamente. Também, os índices de livro de código podem ser codificados, por exemplo, usando uma codificação binária, opcionalmente usando uma codificação de entropia.

Da mesma forma, são obtidas as proporções codificadas 1424.

O processador de sinal de distorção de tempo 1430 compreende um conversor de distorção de tempo do domínio temporal para o domínio de frequências 1434, que é configurado para receber o sinal de áudio 1410 e um contorno de informação de distorção de tempo 1422a associado com o sinal de áudio (ou uma versão codificada deste), e para fornecer, com base nele, uma representação de domínio espectral (domínio de frequências) 1436.

O contorno de informação de distorção de tempo 1422a pode preferivelmente ser obtido a partir da informação codificada 1424 fornecida pelo codificador de contorno de distorção de tempo 1420 usando um decodificador de distorção 1425.

5 Desta forma, pode-se perceber que o codificador (em particular o seu processador de sinal de distorção de tempo 1430) e o decodificador (que recebe a representação codificada 1414 do sinal de áudio) opera nos mesmos contornos de distorção, a saber o contorno de distorção (de tempo) decodificado. Entretanto, em uma
10 configuração simplificada, o contorno de informação de distorção de tempo 1422a usado pelo processador de sinal de distorção de tempo 1430 pode ser idêntico ao contorno de informação de distorção de tempo 1422 inserido no codificador de contorno de distorção de tempo 1420.

15 O conversor de distorção de tempo do domínio temporal para o domínio de frequências 1434 pode, por exemplo, considerar uma distorção de tempo ao formar a representação de domínio espectral 1436, por exemplo usando uma operação de reamostragem variável com o tempo do sinal de áudio 1410.
20 Alternativamente, entretanto, a reamostragem variável com o tempo e a conversão de domínio temporal para domínio de frequências podem ser integradas em um único passo de processamento. O processador de sinal de distorção de tempo também compreende um codificador de valor espectral 1438, que é configurado para
25 codificar a representação de domínio espectral 1346. O codificador de valor espectral 1438 pode, por exemplo, ser configurado pra levar em consideração o mascaramento perceptivo. Também, o codificador de valor espectral 1438 pode ser configurado para

adaptar a exatidão de codificação à relevância perceptiva das bandas de frequência e para aplicar uma codificação de entropia. Da mesma forma, obtém-se a representação codificada 1432 do sinal de áudio 1410.

5 **CALCULADORA DE CONTORNO DE DISTORÇÃO DE TEMPO DE ACORDO COM A FIGURA 15**

10 A Figura 15 mostra o diagrama esquemático de blocos de uma calculadora de contorno de distorção de tempo, de acordo com outra configuração da invenção. A calculadora de contorno de distorção de tempo 1500 é configurada para receber uma informação de proporção de distorção codificada 1510 para fornecer, com base nela, uma pluralidade de valores de nó de distorção 1512. A calculadora de contorno de distorção de tempo 1500 compreende, por exemplo, um decodificador de proporção de
15 distorção 1520, que é configurado para derivar uma sequência de valores de proporção de distorção 1522 a partir da informação de proporção de distorção codificada 1510. A calculadora de contorno de distorção de tempo 1500 também compreende uma calculadora de contorno de distorção 1530, que é configurada para derivar a
20 sequência de valores de nó de distorção 1512 a partir da sequência de valores de proporção de distorção 1522. Por exemplo, a calculadora de contorno de distorção pode ser configurada para obter os valores de nó de contorno de distorção começando a partir de um valor inicial de contorno de distorção, em que proporções
25 entre o valor inicial de contorno de distorção, associado com um nó inicial de contorno de distorção, e os valores de nó de contorno de distorção são determinados pelos valores de proporção de distorção 1522. A calculadora de valor de nó de distorção é

também configurada para computar um valor de nó de contorno de
 distorção 1512 de um determinado nó de contorno de distorção que é
 espaçado a partir do nó inicial de contorno de distorção por um nó
 de contorno de distorção intermediário, com base em uma formação
 5 de produto compreendendo uma proporção entre o valor inicial de
 contorno de distorção (por exemplo, 1) e o valor de nó de contorno
 de distorção do nó de contorno de distorção intermediário e uma
 proporção entre o valor de nó de contorno de distorção do nó de
 contorno de distorção intermediário e o valor de nó de contorno de
 10 distorção do determinado nó de contorno de distorção como fatores.

A seguir, a operação da calculadora de contorno
 de distorção de tempo 1500 será brevemente discutida tomando
 referência às Figuras 16a e 16b...

A Figura 16a mostra uma representação gráfica de
 15 um cálculo sucessivo de um contorno de distorção de tempo. Uma
 primeira representação gráfica 1610 mostra uma sequência de
 índices de livro de código de proporção de distorção de tempo 1510
 (index=0, index=1, index=2, index=3, index=7). Adicionalmente, a
 representação gráfica 1610 mostra uma sequência de valores de
 20 proporção de distorção (0.983, 0.988, 0.994, 1.000, 1.023)
 associados com os índices de livro de código. Adicionalmente,
 pode-se ver que um primeiro valor de nó de distorção 1621 (i=0) é
 escolhido para ser 1 (em que 1 é um valor inicial). Como pode ser
 visto, um segundo valor de nó de distorção 1622 (i=1) é obtido
 25 multiplicando-se o valor inicial de 1 com o primeiro valor de
 proporção de 0,983 (associado com o primeiro índice 0). Também se
 pode ser visto que o terceiro valor de nó de distorção 1623 é obtido
 ao se multiplicar o segundo valor de nó de distorção 1622 de 0,983

pelo segundo valor de proporção de distorção de 0,988 (associado com o segundo índice de 1). Da mesma maneira, o quarto valor de nó de distorção 1624 é obtido ao se multiplicar o terceiro valor de nó de distorção 1623 pelo terceiro valor de proporção de distorção de 0,994
5 (associado com um terceiro índice de 2).

Da mesma forma, uma sequência de valores de nó de distorção 1621, 1622, 1623, 1624, 1625, 1626 são obtidos.

Um respectivo valor de nó de distorção é efetivamente obtido tal que seja um produto do valor inicial (por exemplo, 1) e todos os valores intermediários de proporção de
10 distorção entre os nós de distorção inicial 1621 e o respectivo valor de nó de distorção 1622 to 1626.

Uma representação gráfica 1640 ilustra uma interpolação linear entre os valores de nó de distorção. Por exemplo, os valores interpolados 1621a, 1621b, 1621c poderiam ser
15 obtidos em um decodificador de sinal de áudio entre dois valores adjacentes de nó de distorção de tempo 1621, 1622, por exemplo fazendo uso de uma interpolação linear.

A Figura 16b mostra uma representação gráfica de
20 uma reconstrução de contorno de distorção de tempo usando um reinício periódico a partir de um valor inicial predeterminado, que pode opcionalmente ser implementado na calculadora de contorno de distorção de tempo 1500. Em outras palavras, o reinício repetido ou periódico não é uma característica essencial, desde
25 que um estouro numérico possa ser evitado por qualquer outra medida apropriada no lado do codificador ou no lado do decodificador. Como pode ser visto, uma porção de contorno de distorção pode começar a partir de um nó inicial 1660 em que os

nós de contorno de distorção 1661, 1662, 1663, 1664 podem ser determinados. Para este propósito, valores de proporção de distorção (0.983, 0.988, 0.965, 1.000) podem ser considerados, tal que os nós adjacentes de contorno de distorção 1661 a 1664 da primeira porção de contorno de distorção de tempo são separados por proporções determinadas por estes valores de proporção de distorção. Entretanto, uma segunda porção de contorno de distorção de tempo, adicional, pode ser iniciada após um nó de final 1664 da primeira porção de contorno de distorção de tempo (compreendendo os nós 1660-1664) ter sido alcançado. A segunda porção de contorno de distorção de tempo pode começar a partir de um nó inicial 1665, que pode assumir o valor inicial predeterminado, independente de quaisquer valores de proporção de distorção. Da mesma forma, valores de nó de distorção da segunda porção de contorno de distorção de tempo podem ser computados começando a partir do nó inicial 1665 da segunda porção de contorno de distorção de tempo com base nos valores de proporção de distorção da segunda porção de contorno de distorção de tempo. Posteriormente, uma terceira porção de contorno de distorção de tempo pode começar a partir de um correspondente nó inicial 1670, que pode novamente assumir o valor inicial predeterminado independente de quaisquer valores de proporção de distorção. Da mesma forma, um reinício periódico das porções de contorno de distorção de tempo é obtido. Opcionalmente, uma repetida renormalização pode ser aplicada, como descrita em detalhes acima.

O CODIFICADOR DE SINAL DE ÁUDIO DE ACORDO COM A **FIGURA 17**

A seguir, um codificador de sinal de áudio de acordo com outra configuração da invenção será brevemente

descrito, tomando referência à Figura 17. O codificador de sinal de áudio 1700 é configurado para receber um sinal de áudio multi-canal 1710 e para fornecer uma representação codificada 1712 do sinal de áudio multi-canal 1710. O codificador de sinal de áudio 5 1700 compreende um provedor de representação de áudio codificado 1720, que é configurado para seletivamente fornecer uma representação de áudio compreendendo uma informação comum de contorno de distorção, comumente associada com uma pluralidade de canais de áudio do sinal de áudio multi-canal, ou uma 10 representação de áudio codificado compreendendo informação individual de contorno de distorção, individualmente associados com os diferentes canais de áudio da pluralidade de canais de áudio, dependentes de uma informação que descreve uma similaridade ou diferença entre contornos de distorção associados com os canais 15 de áudio da pluralidade de canais de áudio.

Por exemplo, o codificador de sinal de áudio 1700 compreende uma calculadora de similaridade de contorno de distorção ou calculadora de diferença de contorno de distorção 1730 configurada para prover a informação 1732 que descreve a 20 similaridade ou diferença entre contornos de distorção associados com os canais de áudio. O provedor de representação de áudio codificado compreende, por exemplo, um codificador seletivo de contorno de distorção de tempo 1722 configurado para receber o contorno de informação de distorção de tempo 1724 (que pode ser 25 fornecido externamente ou que pode ser fornecido por uma calculadora de contorno de informação de distorção de tempo opcional 1734) e a informação 1732. Se a informação 1732 indicar que os contornos de distorção de tempo de dois ou mais canais de

áudio são suficientemente semelhantes, o codificador seletivo de contorno de distorção de tempo 1722 pode ser configurado para prover uma informação conjunta de contorno de distorção de tempo codificado. A informação conjunta de contorno de distorção pode, por exemplo, se basear em uma média da informação de contorno de distorção de dois ou mais canais. Entretanto, alternativamente a informação conjunta de contorno de distorção pode se basear em uma única informação de contorno de distorção de um único canal de áudio, porém conjuntamente associados com uma pluralidade de canais.

Entretanto, se a informação 1732 indicar que os contornos de distorção de múltiplos canais de áudio não são suficientemente similares, o codificador seletivo de contorno de distorção de tempo 1722 pode fornecer informação codificada separada dos diferentes contornos de distorção de tempo.

O provedor de representação de áudio codificado 1720 também compreende um processador de sinal de distorção de tempo 1726, que é também configurado para receber o contorno de informação de distorção de tempo 1724 e o sinal de áudio multi-canal 1710. O processador de sinal de distorção de tempo 1726 é configurado para codificar os múltiplos canais do sinal de áudio 1710. O processador de sinal de distorção de tempo 1726 pode compreender diferentes modos de operação. Por exemplo, o processador de sinal de distorção de tempo 1726 pode ser configurado para seletivamente codificar canais de áudio individualmente ou codificá-los conjuntamente, explorando similaridades entre os canais. Em alguns casos, é preferível que o processador de sinal de distorção de tempo 1726 seja capaz de

codificar múltiplos canais de áudio da maneira comum tendo um contorno comum de informação de distorção de tempo. Há casos em que um canal de áudio esquerdo e um canal de áudio direito exibem a mesma evolução de passo relativo, porém têm características de sinal diferentes em outros aspectos, por exemplo diferentes frequências fundamentais absolutas ou diferentes envelopes espectrais. Neste caso, não é desejável codificar o canal de áudio esquerdo e o canal de áudio direito conjuntamente, por causa da significativa diferença entre o canal de áudio esquerdo e o canal de áudio direito. Não obstante, a evolução de passo relativo no canal de áudio esquerdo e no canal de áudio direito pode ser paralela, tal que a aplicação de uma distorção de tempo comum é uma solução muito eficiente. Um exemplo de tal sinal de áudio é uma música polifônica, em que os conteúdos de múltiplos canais de áudio exibem uma diferença significativa (por exemplo, são dominados por diferentes cantores ou instrumentos musicais), porém exibem similar variação de passo. Desta forma, a eficiência de codificação pode ser significativamente melhorada ao fornecer a possibilidade de ter uma codificação conjunta dos contornos de distorção de tempo para múltiplos canais de áudio enquanto se mantém a opção de codificar separadamente os espectros de frequência dos diferentes canais de áudio para os quais é fornecido um contorno comum de informação de passo.

O provedor de representação de áudio codificado 1720 opcionalmente compreende um codificador de informação lateral 1728, que é configurado para receber a informação 1732 e para fornecer uma informação lateral indicando se um contorno de distorção comum codificado está fornecido para múltiplos canais de

áudio ou se contornos de distorção individuais codificados são fornecidos para os múltiplos canais de áudio. Por exemplo, tal informação lateral pode ser fornecida na forma de um *flag* de 1 bit chamado "common_tw".

5 Para resumir, o codificador seletivo de contorno de distorção de tempo 1722 seletivamente fornece representações individuais codificadas dos contornos de distorção de tempo de áudio associados com múltiplos sinais de áudio, ou uma representação codificada conjunta de contorno de distorção de

10 tempo representando um único conjunto de contorno de distorção de tempo associado com os múltiplos canais de áudio. O codificador de informação lateral 1728 opcionalmente fornece uma informação lateral indicando se são fornecidas representações individuais de contorno de distorção de tempo ou uma representação de conjunto de

15 contorno de distorção de tempo. O processador de sinal de distorção de tempo 1726 fornece representações codificadas dos múltiplos canais de áudio. Opcionalmente, uma informação codificada comum pode ser fornecida para múltiplos canais de áudio. Entretanto, tipicamente é até possível fornecer

20 representações individuais codificadas de múltiplos canais de áudio, para os quais uma representação comum de contorno de distorção de tempo está disponível, tal que diferentes canais de áudio com diferentes conteúdos de áudio, porém distorções de tempo idênticas são apropriadamente representados. Consequentemente, a

25 representação codificada 1712 compreende informação codificada fornecida pelo codificador seletivo de contorno de distorção de tempo 1722, e o processador de sinal de distorção de tempo 1726 e, opcionalmente, o codificador de informação lateral 1728.

DECODIFICADOR DE SINAL DE ÁUDIO DE ACORDO COM A

FIGURA 18

A Figura 18 mostra um diagrama esquemático de blocos de um decodificador de sinal de áudio de acordo com uma configuração da invenção. O decodificador de sinal de áudio 1800 é configurado para receber uma representação de sinal de áudio codificado 1810 (por exemplo, a representação codificada 1712) e para fornecer, com base nela, uma representação decodificada 1812 do sinal de áudio multi-canal. O decodificador de sinal de áudio 1800 compreende um extrator de informação lateral 1820 e um decodificador de distorção de tempo 1830. O extrator de informação lateral 1820 é configurado para extrair uma informação de aplicação de contorno de distorção de tempo 1822 e uma informação de contorno de distorção 1824 a partir da representação de sinal de áudio codificado 1810. Por exemplo, o extrator de informação lateral 1820 pode ser configurado para reconhecer se um único, contorno comum de informação de distorção de tempo está disponível para múltiplos canais do sinal de áudio codificado, ou se o contorno de informação de distorção de tempo separado está disponível para os múltiplos canais. Da mesma forma, o extrator de informação lateral pode fornecer tanto a informação de aplicação de contorno de distorção de tempo 1822 (indicando se está disponível contorno de informação de distorção de tempo conjunto ou individual) quanto o contorno de informação de distorção de tempo 1824 (descrevendo uma evolução temporal do contorno comum (conjunto) de distorção de tempo ou dos contornos de distorção de tempo individuais). O decodificador de distorção de tempo 1830 pode ser configurado para reconstruir a representação decodificada

do sinal de áudio multi-canal com base na representação de sinal de áudio codificado 1810, levando em consideração a distorção de tempo descrita pela informação 1822, 1824. Por exemplo, o decodificador de distorção de tempo 1830 pode ser configurado para

5 aplicar um contorno comum de distorção de tempo para decodificar diferentes canais de áudio, para os quais está disponível informação individual codificada sobre domínio de frequências. Da mesma forma, o decodificador de distorção de tempo 1830 pode, por exemplo, reconstruir diferentes canais do sinal de áudio multi-

10 canal, que compreendem distorção de tempo similar ou idêntica, porém passo diferente.

STREAM DE ÁUDIO DE ACORDO COM AS FIGURAS 19A A 19E

A seguir, será descrito um *stream* de áudio, que compreende uma representação codificada de um ou mais sinais de

15 canal de áudio e um ou mais contornos de distorção de tempo.

A Figura 19a mostra uma representação gráfica de um assim chamado elemento de *stream* de dados "USAC_raw_data_block" que pode compreender um elemento de canal único (SCE), um elemento de dois canais (CPE) ou uma combinação de um ou mais elementos de

20 canal único e/ou um ou mais elementos de dois canais.

O "USAC_raw_data_block" pode tipicamente compreender um bloco de dados de áudio codificados, enquanto informação adicional sobre contorno de distorção de tempo pode ser fornecida em um elemento separado de *stream* de dados. Não

25 obstante, é usualmente possível codificar alguns dados de contorno de distorção de tempo no "USAC_raw_data_block".

Como pode ser visto a partir da Figura 19b, um elemento de canal único tipicamente compreende um *stream* de canal

de domínio de frequências ("fd_channel_stream"), que será explicado em detalhes com referência à Figura 9d.

Como pode ser visto a partir da Figura 19c, um elemento de dois canais ("channel_pair_element") tipicamente
5 compreende uma pluralidade de *streams* de canal de domínio de frequências. Também, o elemento de dois canais pode compreender informação de distorção de tempo. Por exemplo, um *flag* de ativação de distorção de tempo ("tw_MDCT") que pode ser transmitido em um elemento de *stream* de dados de configuração ou no
10 "USAC_saw_data_block" determina se a informação de distorção de tempo está incluída no elemento de dois canais. Por exemplo, se o *flag* "tw_MDCT" indicar que a distorção de tempo está ativa, o elemento de dois canais pode compreender um *flag* ("common_tw") que indica se há uma distorção de tempo comum para os canais de áudio
15 do elemento de dois canais. Se o referido *flag* (common_tw) indicar que há uma distorção de tempo comum para múltiplo dos canais de áudio, então uma informação comum de distorção de tempo (tw_data) é incluída no elemento de dois canais, por exemplo, separado dos *streams* de canal de domínio de frequências.

20 Tomando referência agora à Figura 19d, o *stream* de canal de domínio de frequências é descrito. Como pode ser visto a partir da Figura 19d, o *stream* de canal de domínio de frequências, por exemplo, compreende uma informação de ganho global. Também, o *stream* de canal de domínio de frequências
25 compreende dados de distorção de tempo, se a distorção de tempo estiver ativa (*flag* "tw_MDCT" ativo) e se não houver informação comum de distorção de tempo para canal de sinal de áudio múltiplo (o *flag* "common_tw" está inativo).

Adicionalmente, um *stream* de canal de domínio de frequências também compreende dados de fator de escala ("scale_factor_data") e dados espectrais codificados (por exemplo, dados espectrais aritmeticamente codificados "ac_spectral_data").

5 Tomando referência agora à Figura 19e, a sintaxe dos dados de distorção de tempo será brevemente discutida. Os dados de distorção de tempo podem por exemplo, opcionalmente, compreender um *flag* (por exemplo, "tw_data_present" ou "Dados de Passo ativos") indicando se os dados de distorção de tempo estão presentes. Se os
10 dados de distorção de tempo estiverem presentes, (isto é o contorno de distorção de tempo não está plano) os dados de distorção de tempo podem compreender uma sequência de uma pluralidade de valores de proporção de distorção de tempo codificados (por exemplo, "tw_ratio[i]" ou "pitchIdx[i]"), que podem, por exemplo, ser condicionados de
15 acordo com a tabela do livro de códigos da Figura 9c.

Desta forma, os dados de distorção de tempo podem compreender um *flag* indicando que não há dados de distorção de tempo disponíveis, que pode ser estabelecido por um codificador de sinal de áudio, se o contorno de distorção de tempo for constante
20 (as proporções de distorção de tempo são aproximadamente iguais a 1,000). Em contraste, se o contorno de distorção de tempo estiver variando, proporções entre nós subsequentes de contorno de distorção de tempo podem ser codificadas usando os índices de livro de código que compõem a informação "tw_ratio".

25 CONCLUSÃO

Resumindo o texto acima, as configurações de acordo com a invenção trazem diferentes melhorias no campo de distorção de tempo.

Os aspectos da invenção descritos aqui estão no contexto de um codificador de tempo distorcido por transformada MDCT (vide, por exemplo, a referência [1]). Configurações de acordo com a invenção fornecem métodos para um desempenho melhorado de um
5 codificador de tempo distorcido por transformada MDCT.

De acordo com um aspecto da invenção, é fornecido um formato de *bitstream* particularmente eficiente. A descrição do formato de *bitstream* baseia-se em e aperfeiçoa a sintaxe de *bitstream* MPEG-2 AAC (vide, por exemplo, a referência [2]), porém
10 é naturalmente aplicável a todos os formatos de *bitstream* com um cabeçalho de descrição geral no início de um *stream* e uma sintaxe de informação na forma de estrutura individual.

Por exemplo, a seguinte informação lateral pode ser transmitida no *bitstream*:

15 Em geral, um *flag* de um bit (por exemplo, a saber "tw_MDCT") pode apresentar na configuração específica de áudio geral (GASC), indicando se a distorção de tempo está ativa ou não. Dados de passo podem ser transmitidos usando a sintaxe mostrada na Figura 19e ou a sintaxe mostrada na Figura 19f. Na sintaxe mostrada na
20 Figura 19f, o número de passos ("numPitches") pode ser igual a 16, e o número de bits de passo em ("numPitchBits") pode ser igual a 3. Em outras palavras, pode haver 16 valores de proporção de distorção codificados por porção de contorno de distorção de tempo (ou por sinal de estrutura de áudio), e cada valor de proporção de contorno
25 de distorção pode ser codificado usando 3 bits.

Além disso, em um elemento de canal único (SCE) os dados de passo (*pitch_data*[]) podem ser localizados antes dos dados de seção no canal individual, se a distorção estiver ativa.

Em um elemento de dois canais (CPE), um *flag* de passo comum sinaliza se houver dados de um passo comum para ambos os canais, que segue após isto, se não, os contornos de passo individuais são encontrados nos canais individuais.

5 A seguir, será dado um exemplo para um elemento de dois canais. Um exemplo poderia ser um sinal de uma única fonte de sons harmônicos, disposta dentro do panorama estéreo. Neste caso, os contornos de passo relativo para o primeiro canal e o segundo canal serão iguais ou difeririam apenas ligeiramente devido a alguns pequenos erros na estimativa da variação. Neste caso, o codificador pode decidir que ao invés de enviar dois contornos de passo codificados separados para cada canal, enviar apenas um contorno de passo que é uma média dos contornos de passo do primeiro e segundo canal, e usar o mesmo contorno ao aplicar a TW-MDCT em ambos os canais. Por outro lado, poderia haver um sinal onde a estimativa do contorno de passo gera diferentes resultados para o primeiro e segundo canal respectivamente. Neste caso, os contornos de passo individualmente codificados são enviados dentro do canal correspondente.

20 A seguir, será descrita uma decodificação vantajosa de dados de passo de contorno, de acordo com um aspecto da invenção. Por exemplo, se o *flag* "active PitchData" for 0, o contorno de passo é estabelecido para 1 para todas as amostras na estrutura, do contrário os nós de contorno de passo individuais são computados como segue:

- há $\text{numPitches} + 1$ nós,
- nó [0] é sempre 1,0;
- $\text{nó}[i] = \text{nó}[i-1] \cdot \text{relChange}[i]$ ($i=1..\text{numPitches}+1$),

onde o relChange é obtido por quantização inversa do pitchIdx[i].

O contorno de passo é então gerado pela interpolação linear entre os nós, onde as posições de amostra de nós são 0:frameLen/numPitches:frameLen.

5

ALTERNATIVAS DE IMPLEMENTAÇÃO

Dependendo de certos requisitos de implementação, as configurações da invenção podem ser implementadas em hardware ou em software. A implementação pode ser realizada usando uma mídia de armazenamento digital, por exemplo um disquete, um DVD, um CD, um ROM, um PROM, um EPROM, um EEPROM ou uma memória FLASH, tendo sinais de controle eletronicamente legíveis armazenados nela, que cooperam (ou são capazes de cooperar) com um sistema de computador programável tal que o respectivo método seja executado.

15

Albumas configurações de acordo com a invenção compreendem um carregador de dados tendo sinais de controle eletronicamente legíveis, que são capazes de cooperar com um sistema de computador programável, tal que um dos métodos descritos aqui seja executado.

20

Geralmente, as configurações da presente invenção podem ser implementadas como um produto de programa de computador com um código de programa, o código de programa sendo operante para executar um dos métodos quando o produto do programa de computador roda em em um computador. O código de programa pode, por exemplo, ser armazenado em um carregador legível pela máquina.

25

Outras configurações compreendem o programa de computador para executar um dos métodos descritos aqui, armazenado em um carregador legível pela máquina.

Em outras palavras, uma configuração do método inventivo é, portanto, um programa de computador que tem um código de programa para executar um dos métodos descritos aqui, quando o programa de computador roda em um computador.

5 Uma configuração adicional dos métodos inventivos é, portanto, um carregador de dados (ou uma mídia de armazenamento digital, ou uma mídia legível por computador) compreendendo, nela registrado, o programa de computador para executar um dos métodos descritos aqui.

10 Uma configuração adicional do método inventivo é, portanto, um *stream* de dados ou uma sequência de sinais representando o programa de computador para executar um dos métodos descritos aqui. O *stream* de dados ou a sequência de sinais pode, por exemplo, ser configurado para ser transferido por meio de uma conexão para
15 comunicação de dados, por exemplo, por meio da Internet.

 Uma configuração adicional compreende um meio de processamento, por exemplo um computador, ou um dispositivo lógico programável, configurado para ou adaptado para executar um dos métodos descritos aqui. Al

20 Uma configuração adicional compreende um computador tendo nele instalado o programa de computador para executar um dos métodos descritos aqui..

 Em algumas configurações, um dispositivo lógico programável (por exemplo, um arranjo de portas programável em campo)
25 pode ser usado para realizar algumas ou todas as funcionalidades dos métodos descritos aqui. Em algumas configurações, um *field programmable gate array* pode cooperar com um microprocessador a fim de executar um dos métodos descritos aqui.

REFERÊNCIAS

[1] L. Villemoes, "Time Warped Transform Coding of Audio Signals", PCT/EP2006/010246, Int. patent application, November 2005.

5 [2] Generic Coding of Moving Pictures and Associated Audio: Advanced Audio Coding. International Standard 13818-7, ISO/IECJTC1/SC29/WG11 Moving Pictures Expert Group, 1997.

REIVINDICAÇÕES

1. **"Decodificador de sinal de áudio",** **caracterizado** pelo decodificador de sinal de áudio (200;300;1500;1800) para fornecer uma representação de sinal de áudio multi-canal decodificado (232;312;1812) com base em uma representação de sinal de áudio multi-canal codificado (211,212;310;1810) compreender:

um decodificador de distorção de tempo (210,216,218,219,220,230,240;340;1830) configurado para seletivamente usar contornos de distorção de tempo específicos para canal de áudio, individuais (332;1824) ou um contorno de distorção de tempo multi-canal conjunto (332;1824) para uma reconstrução de uma pluralidade de canais de áudio representados pela representação de sinal de áudio multi-canal codificado.

2. Decodificador de sinal de áudio (200;300;1800), de acordo com a reivindicação 1, **caracterizado** pelo fato de que o decodificador de distorção de tempo (210,216,218,219,220,230,240;340;1830) ser configurado para seletivamente usar um contorno de distorção de tempo multi-canal conjunto (332;1824) para uma reconstrução de distorção de tempo de uma pluralidade de canais de áudio representados pela representação de sinal de áudio multi-canal codificado para a qual está disponível informação individual codificada de domínio espectral (211).

3. Decodificador de sinal de áudio (200;300;1800), de acordo com a reivindicação 2, **caracterizado** pelo fato de que o decodificador de distorção de tempo (210,216,218,219,220,230,240;340;1830) ser configurado para

receber uma primeira informação de domínio espectral associada com um primeiro dos canais de áudio, e para fornecer, com base nela, uma representação de domínio de tempo ($X[][]$) do primeiro canal de áudio usando uma transformação de domínio de frequências para domínio de tempo distorcido;

onde o decodificador de distorção de tempo é adicionalmente configurado para receber uma segunda informação de domínio espectral codificado, associada com um segundo dos canais de áudio, e para fornecer, com base nele, uma representação de domínio de tempo distorcido ($X[][]$) do segundo canal de áudio usando uma transformação de domínio de frequências para domínio de tempo;

onde a segunda informação de domínio espectral é diferente da primeira informação de domínio espectral; e

onde o decodificador de distorção de tempo é configurado para reamostrar de forma variável com o tempo, com base no contorno de distorção de tempo multi-canal conjunto, a representação de domínio de tempo distorcido do primeiro canal de áudio, uma versão processada deste, para obter uma representação regularmente amostrada do primeiro canal de áudio,

e para reamostrar de forma variável com o tempo, com base no contorno de distorção de tempo multi-canal conjunto, a representação de domínio de tempo distorcido do segundo canal de áudio, ou uma versão processada deste, para obter uma representação regularmente amostrada do segundo canal de áudio.

4. Decodificador de sinal de áudio (200;300;1800), de acordo com qualquer uma das reivindicações de 1 a 3, **caracterizado** pelo fato decodificador de distorção de tempo

ser configurado para derivar um contorno de tempo multi-canal conjunto a partir da informação de contorno de distorção de tempo multi-canal conjunto;

para derivar uma primeira forma de janela específica de canal, individual, associada com o primeiro dos canais de áudio com base em uma primeira informação de forma de janela codificada;

para derivar uma segunda forma de janela específica de canal, individual, associada com o segundo dos canais de áudio com base em uma segunda informação de forma de janela codificada;

para aplicar a primeira forma de janela à representação de domínio de tempo distorcido do primeiro canal de áudio, para obter uma versão processada da representação de domínio de tempo distorcido do primeiro canal de áudio;

para aplicar a segunda forma de janela à representação de domínio de tempo distorcido do segundo canal de áudio, para obter uma versão processada da representação de domínio de tempo distorcido do segundo canal de áudio; e

onde o decodificador de distorção de tempo é capaz de aplicar diferentes formas de janela às representações de domínio de tempo distorcido do primeiro e segundo canal de áudio de uma dada estrutura na dependência de informação de forma de janela específica de canal, individual.

5. Decodificador de sinal de áudio (200;300;1800), de acordo com a reivindicação 4, **caracterizado** pelo fato de o decodificador de distorção de tempo ser configurado para aplicar uma escalação de tempo comum, que é determinada pelo

contorno de tempo multi-canal conjunto, a diferentes formas de janela ao aplicar a função janela às representações de domínio de tempo distorcido do primeiro e segundo canais de áudio.

6. Codificador de sinal de áudio (100;1700) para fornecer uma representação codificada (150,152;1712) de um sinal de áudio multi-canal, o codificador de sinal de áudio **caracterizado** por compreender:

um provedor de representação de áudio codificado (104,106,108,108a,114;1720) configurado para seletivamente fornecer uma representação de áudio codificada (150,152;1712) compreendendo uma informação de contorno de distorção de tempo multi-canal comum, comumente associada com uma pluralidade de canais de áudio do sinal de áudio multi-canal, ou uma representação de áudio codificada compreendendo informação individual de contorno de distorção de tempo, individualmente associada com os diferentes canais de áudio da pluralidade de canais de áudio, na dependência de uma informação descrevendo uma similaridade ou diferença entre contornos de distorção de tempo associados com os canais de áudio da pluralidade de canais de áudio.

7. Codificador de sinal de áudio (100;1700) de acordo com a reivindicação 6, **caracterizado** pelo provedor de representação de áudio codificado (104,106,108,108a,114;1720) ser configurado para aplicar uma informação de contorno de distorção de tempo multi-canal comum para obter uma versão de tempo distorcido de um primeiro dos canais de áudio e para obter uma versão de tempo distorcido de um segundo dos canais de áudio, e para fornecer uma primeira informação individual codificada de domínio espectral, associada com um primeiro dos

canais de áudio, com base na versão de tempo distorcido do primeiro canal de áudio, e para fornecer uma segunda informação individual codificada de domínio espectral, associada com um segundo dos canais de áudio, com base na versão de tempo distorcido do segundo canal de áudio.

8. Codificador de sinal de áudio (100;1700), de acordo com a reivindicação 6 ou 7, **caracterizado** pelo fato de que o provedor de representação de áudio codificado (104,106,108,108a,114;1720) ser configurado para fornecer a representação codificada (150,152;1712) do sinal de áudio multi-canal tal que a representação codificada do sinal multi-canal compreenda a informação de contorno de distorção de tempo multi-canal comum, uma representação espectral codificada de uma versão de tempo distorcido do sinal de áudio de um primeiro canal, tempo distorcido de acordo com a informação de contorno de distorção de tempo multi-canal comum, e uma representação espectral codificada de uma versão de tempo distorcido do sinal de áudio de um segundo canal, tempo distorcido de acordo com a informação de contorno de distorção de tempo multi-canal comum.

9. Codificador de sinal de áudio (100;1700), de acordo com qualquer uma das reivindicações de 6 a 8, **caracterizado** pelo codificador de sinal de áudio ser configurado para obter a informação de contorno de distorção de tempo multi-canal comum tal que a informação de contorno de distorção de tempo multi-canal comum representa uma média de contornos de tempo individuais associados com o primeiro canal de sinal de áudio e o segundo canal de sinal de áudio.

10. Codificador de sinal de áudio (100;1700), de

acordo com qualquer uma das reivindicações de 6 a 9, **caracterizado** pelo provedor de representação de áudio codificado ser configurado para fornecer uma informação lateral (*tw_data_present*; *common_tw*) dentro da representação codificada (150;152;1712) do sinal de áudio multi-canal, a informação lateral indicando, conforme a estrutura de áudio, se os dados de distorção de tempo estão presentes para uma dada estrutura de áudio, e se uma informação de contorno comum de distorção de tempo está presente para a dada estrutura de áudio.

11. Representação de sinal de áudio multi-canal codificado (*usac_raw_data_block*) representando um sinal de áudio multi-canal **caracterizado** por compreender:

uma representação de domínio de frequências codificada (*fd_channel_stream*) representando uma pluralidade de canais de áudio de tempo distorcido, tempo distorcido de acordo com uma distorção de tempo comum; e

uma representação codificada (*tw_data*) de uma informação de contorno de distorção de tempo multi-canal comum, comumente associada com os canais de áudio e representando a distorção de tempo comum.

12. Representação de sinal de áudio multi-canal codificado (*usac_raw_data_block*), de acordo com a reivindicação 11, **caracterizado** pela representação de domínio de frequências codificada compreende informação de domínio de frequência codificada individual (*fd_channel_stream*) de múltiplos canais de áudio que têm diferentes conteúdos de áudio, e onde a representação codificada (*tw_data*) da informação de contorno de distorção de tempo multi-canal comum está associada aos múltiplos

canais de áudio que têm conteúdo de áudio diferente.

13. Método para fornecer uma representação de sinal de áudio multi-canal decodificado (232;300;1500;1800) com base em uma representação de sinal de áudio multi-canal codificado (211,212;310;1810), **caracterizado** por compreender:

uso seletivo de contornos individuais de distorção de tempo específicos para canal de áudio ou um contorno de distorção de tempo multi-canal conjunto para uma reconstrução de uma pluralidade de canais de áudio representados pela representação de sinal de áudio multi-canal codificado.

14. Método para fornecer uma representação codificada (150,152;1712) de um sinal de áudio multi-canal, o **caracterizado** por compreender:

fornecimento seletivo de uma representação de áudio codificada compreendendo uma informação de contorno de distorção de tempo multi-canal comum, comumente associada com uma pluralidade de canais de áudio do sinal de áudio multi-canal, ou uma representação de áudio codificada compreendendo informação individual de contorno de distorção de tempo, individualmente associada com os diferentes canais de áudio da pluralidade de canais de áudio, na dependência de uma informação descrevendo uma similaridade ou diferença entre contornos de distorção de tempo associados com os canais de áudio da pluralidade de canais de áudio.

15. Meio de armazenamento não transitório legível por máquina, caracterizado por conter conjunto de instruções que, quando executadas realizam o método de acordo com a reivindicação 13 ou reivindicação 14.

codificador de áudio de distorção de tempo

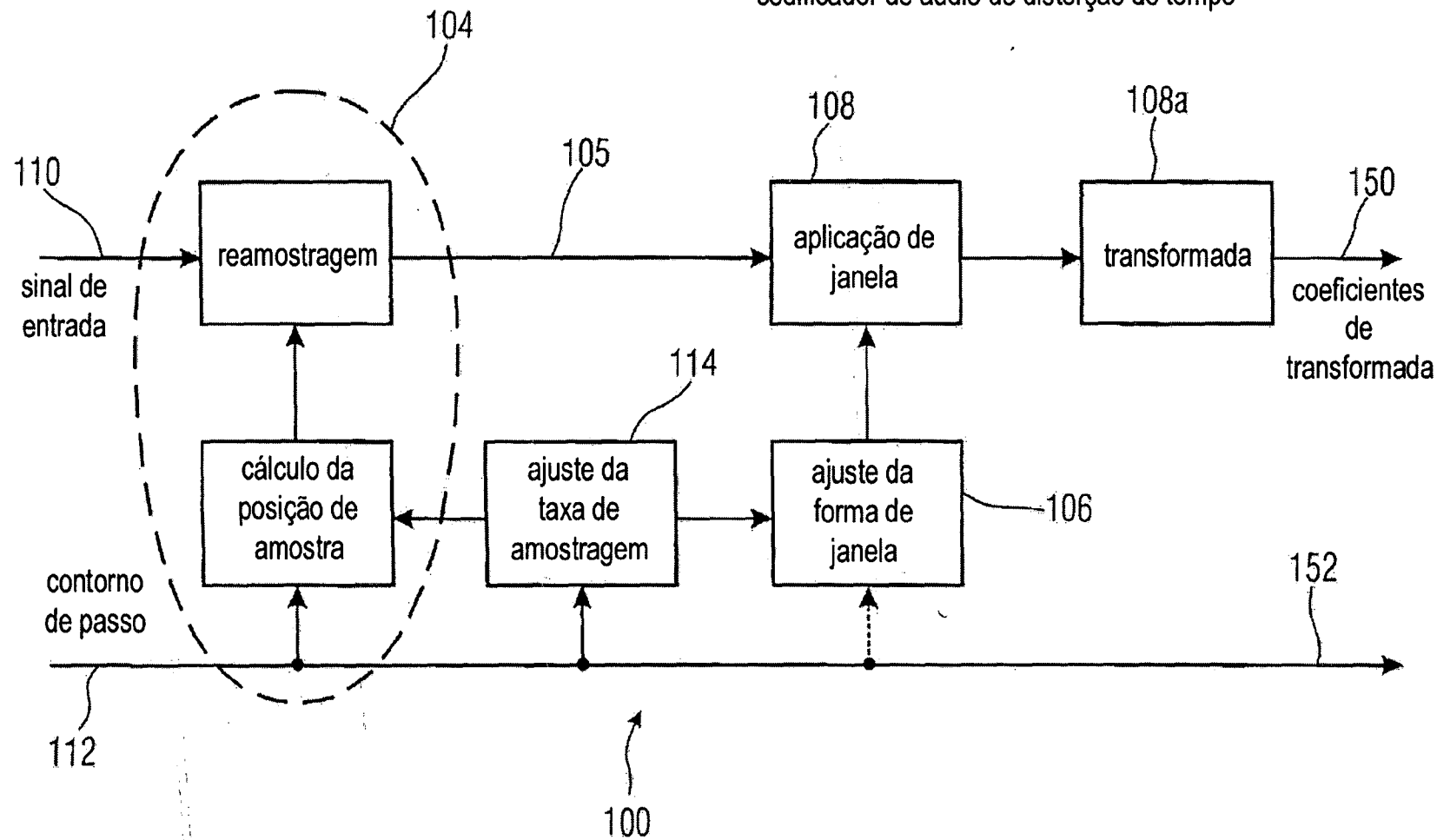


FIG 1

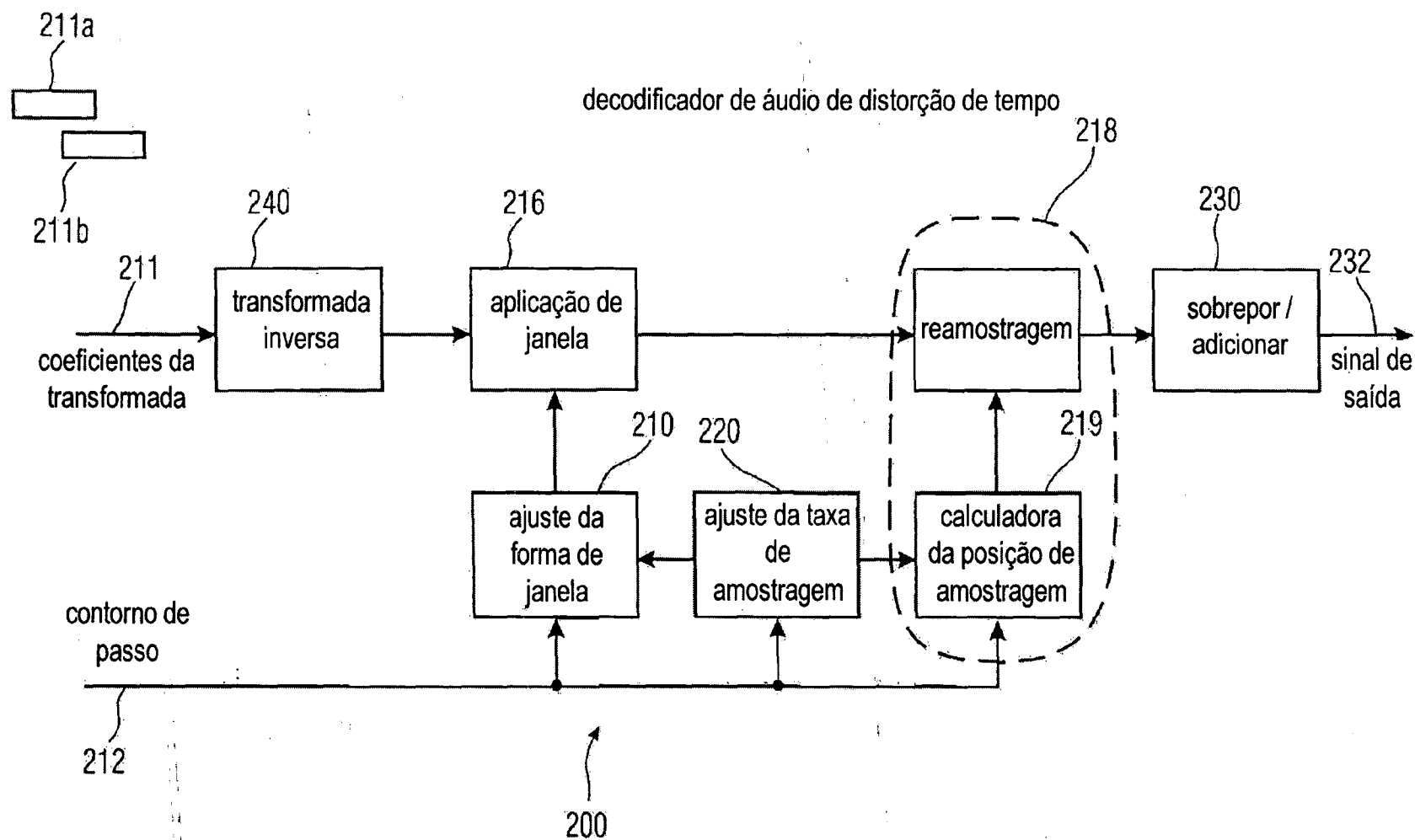


FIG 2

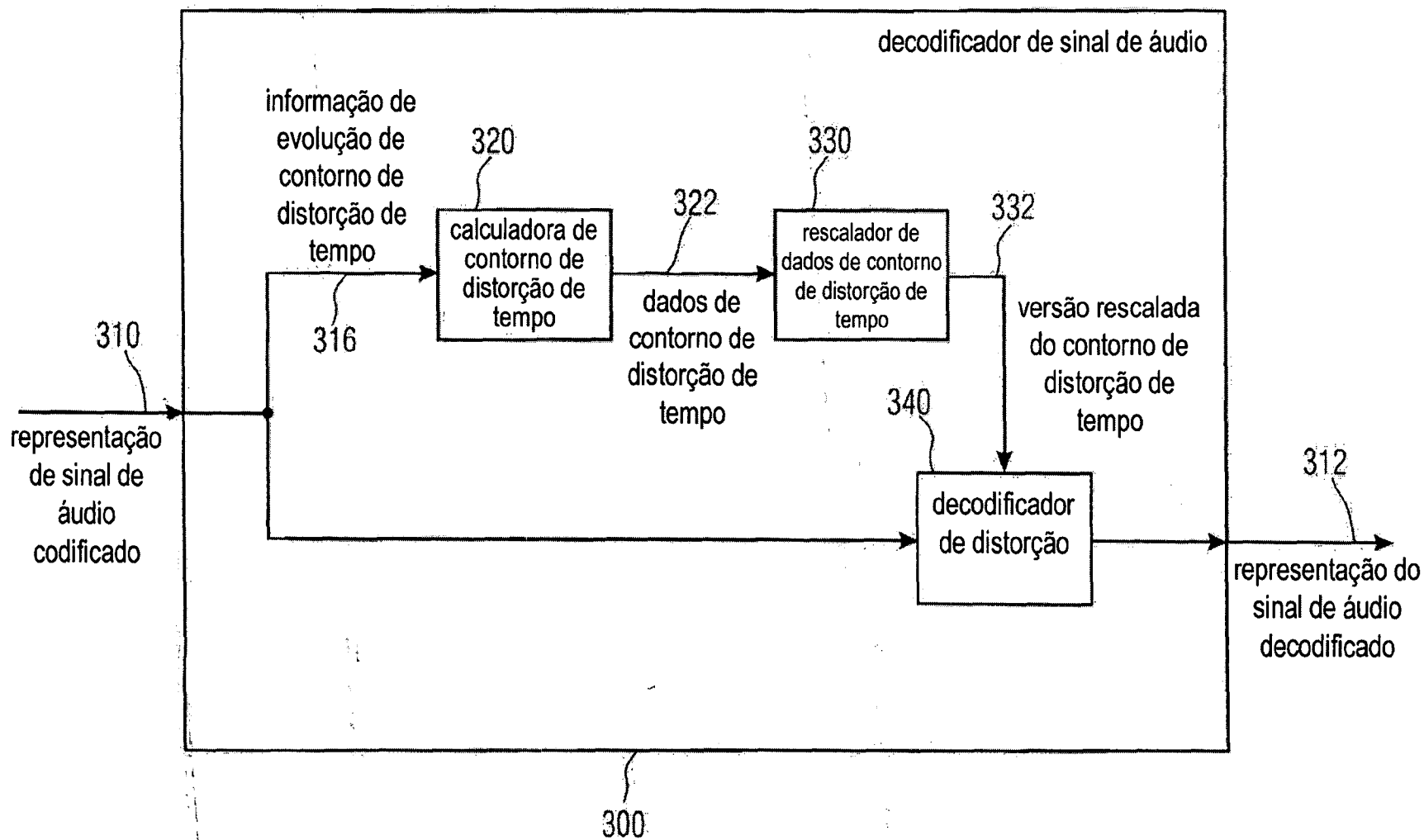


FIG 3

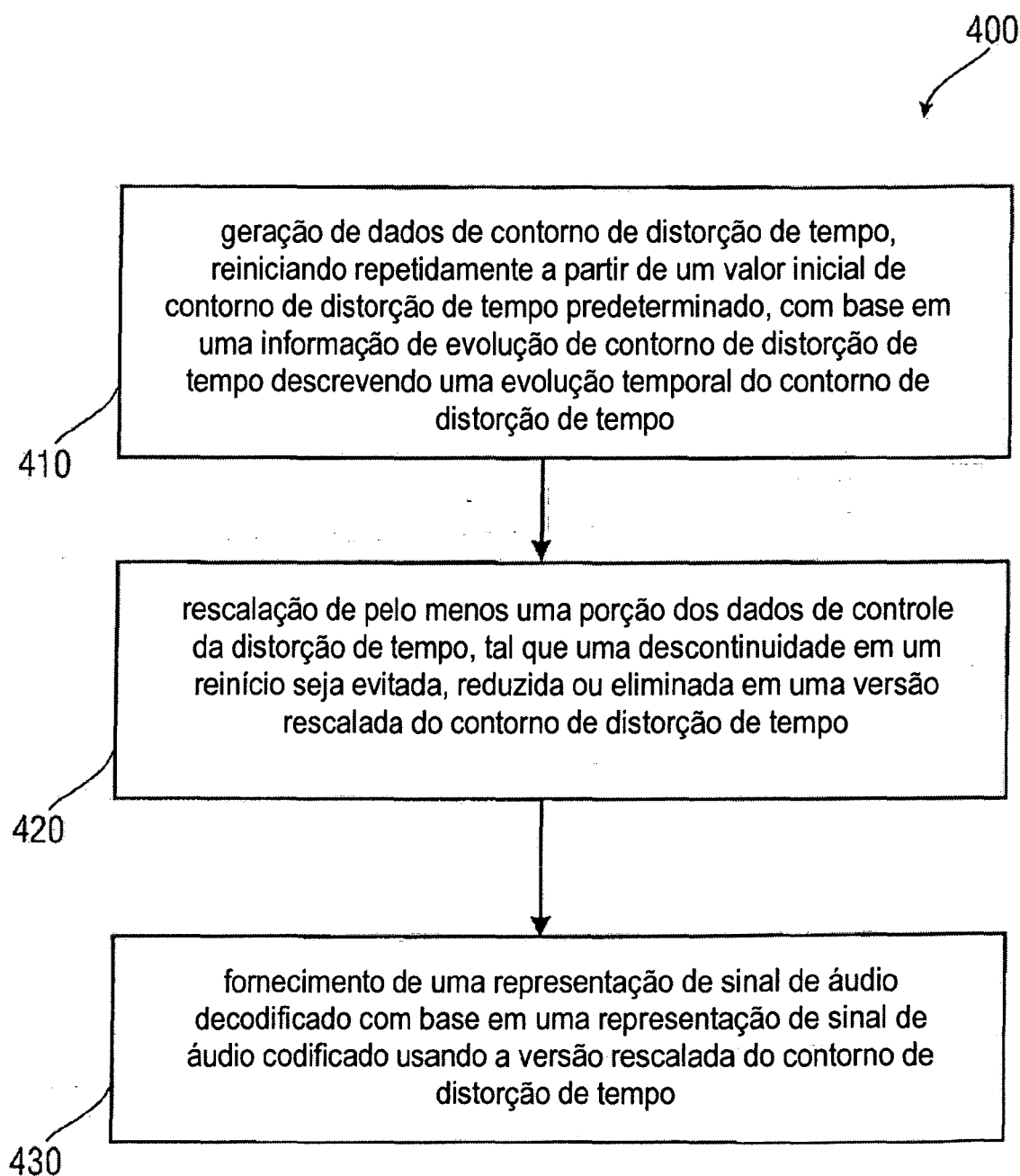


FIG 4

FIG 5A

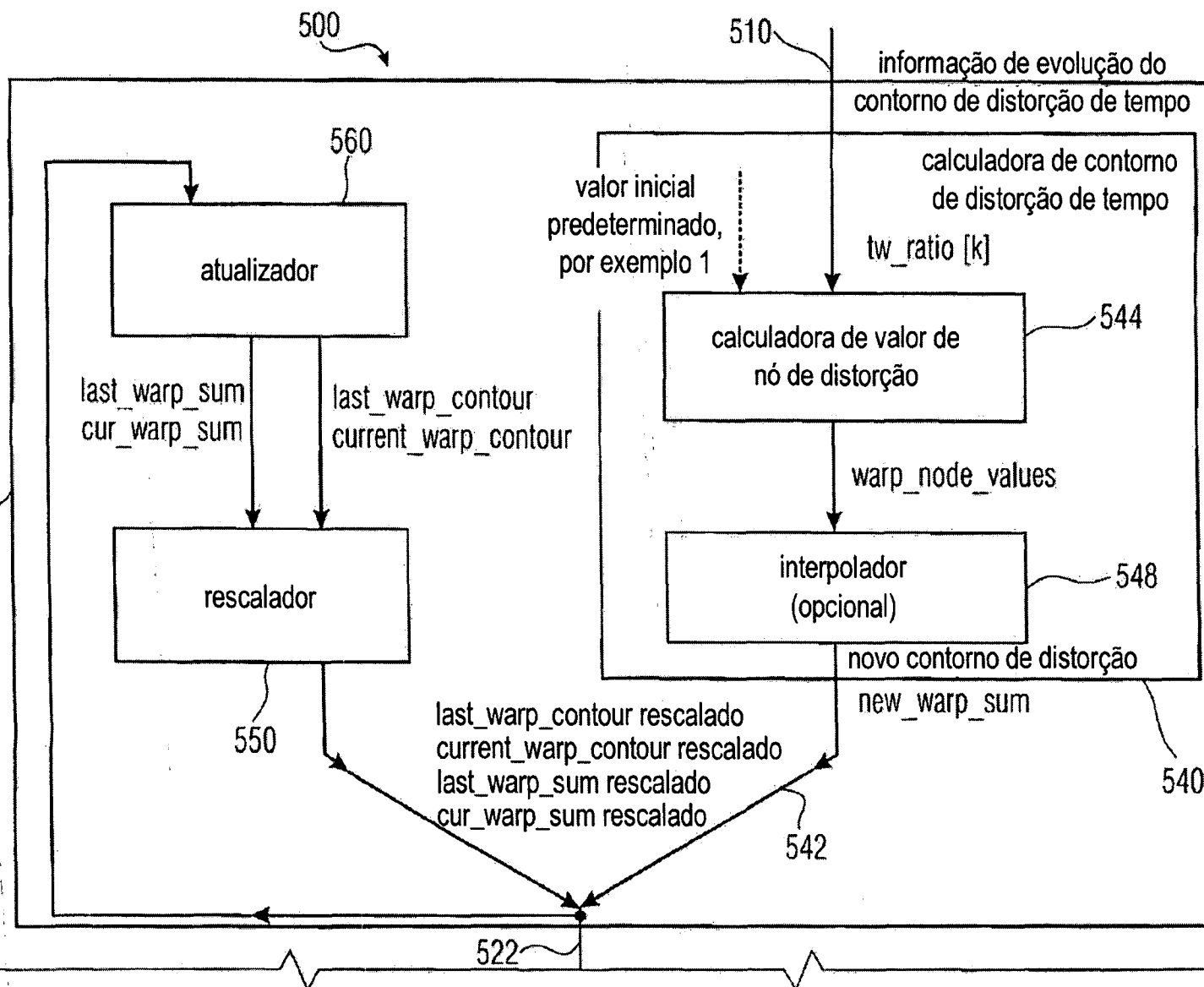


FIG 5A

FIG 5B

FIG 5B

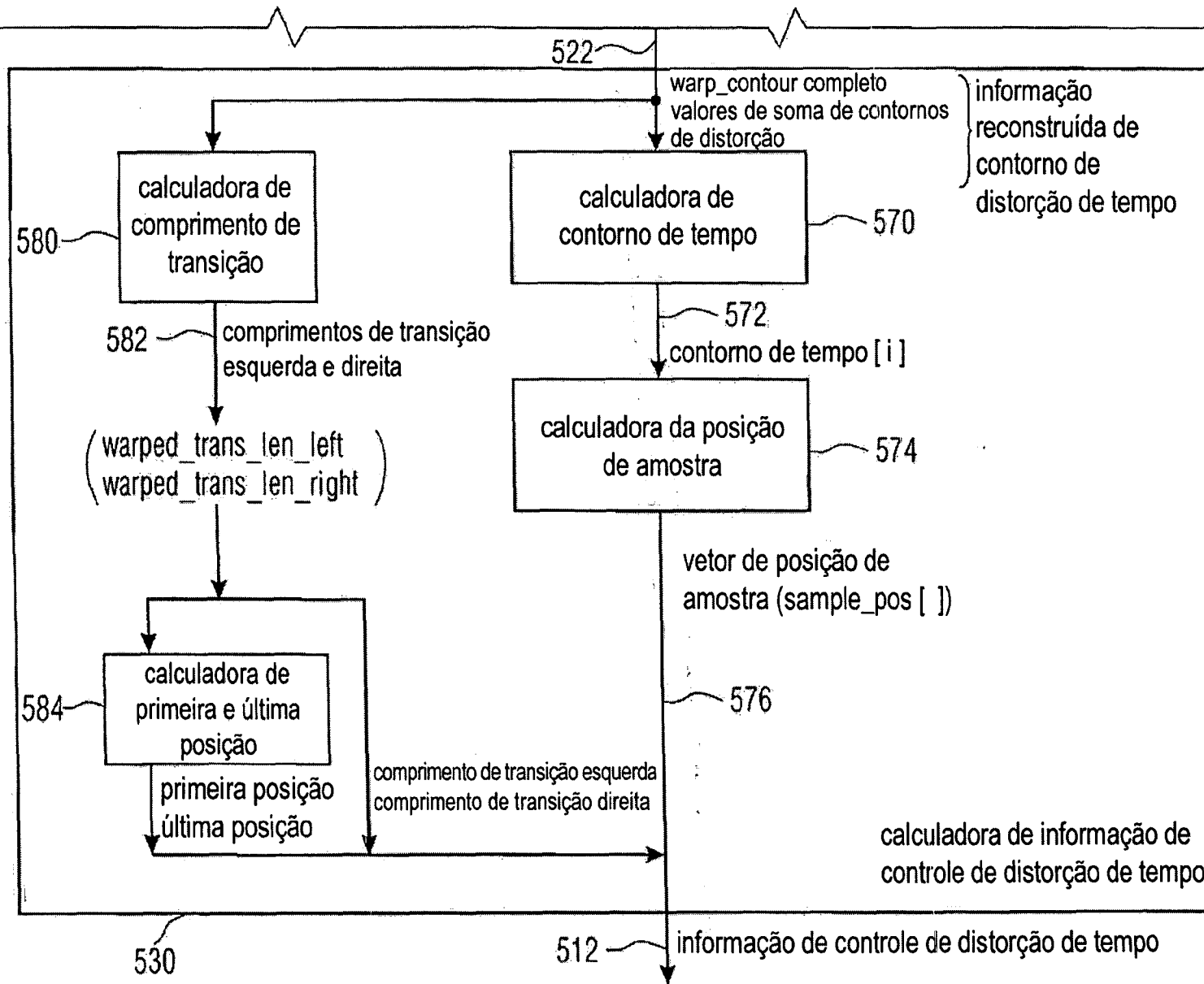
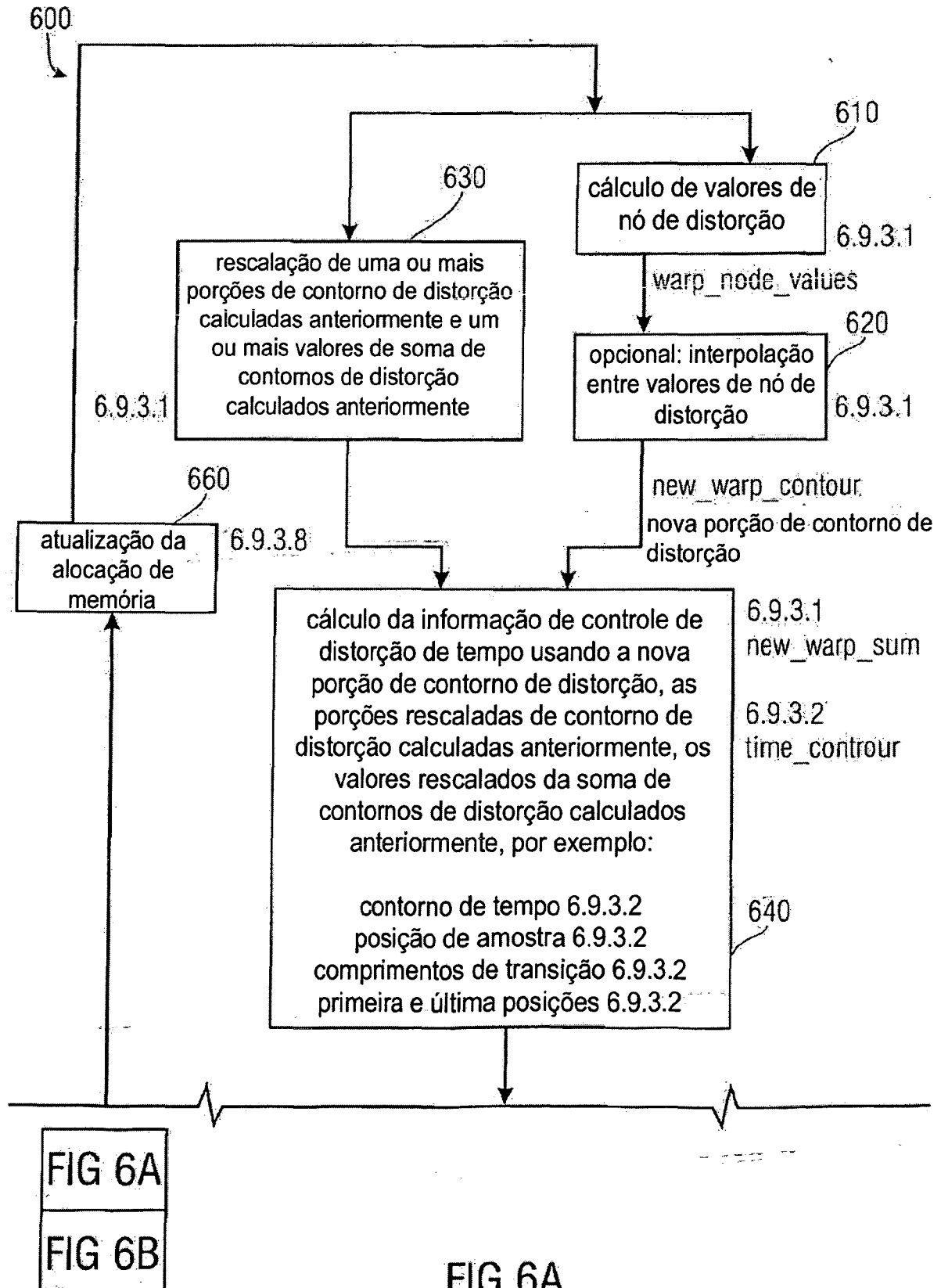


FIG 5A

FIG 5B



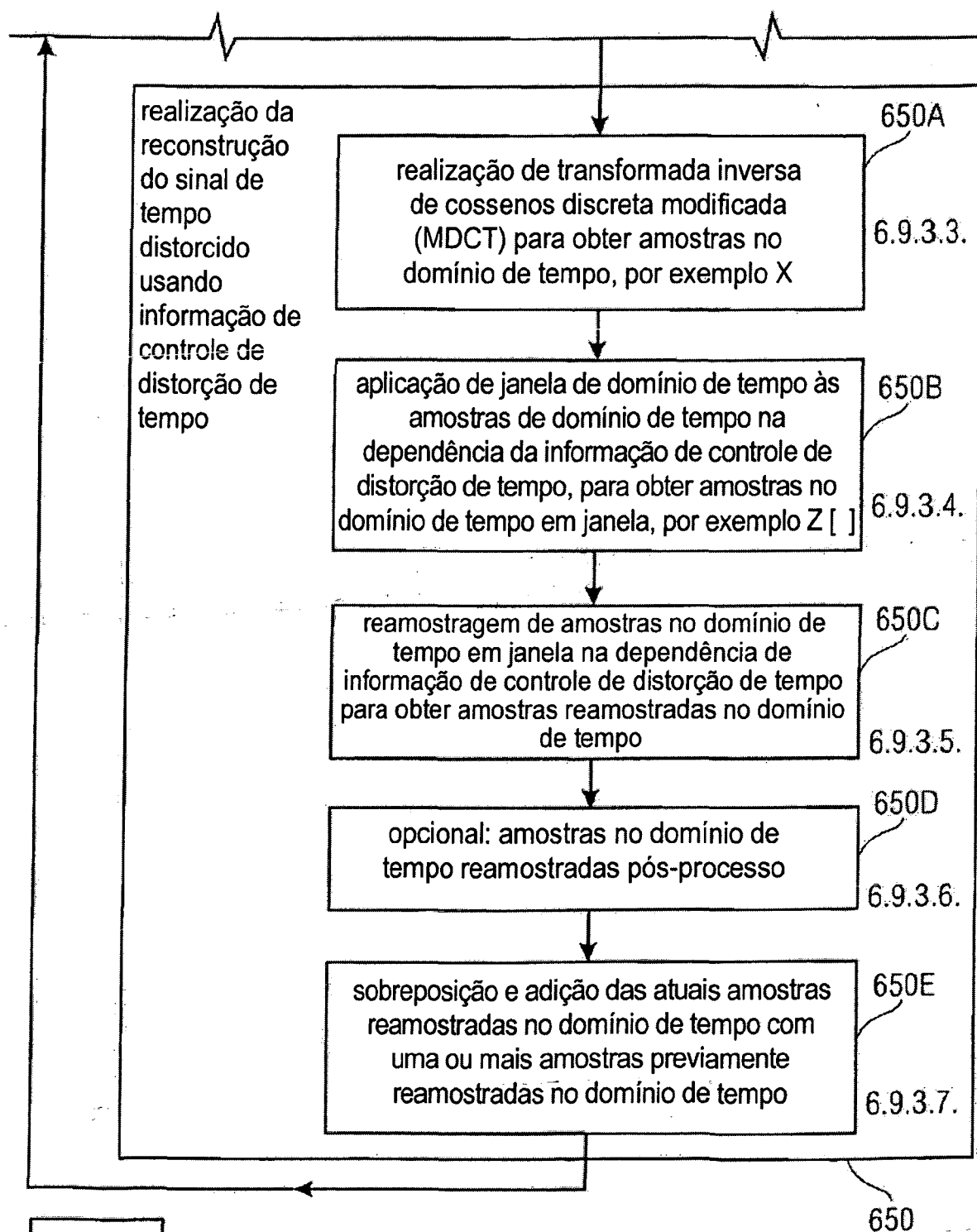


FIG 6A

FIG 6B

FIG 6B

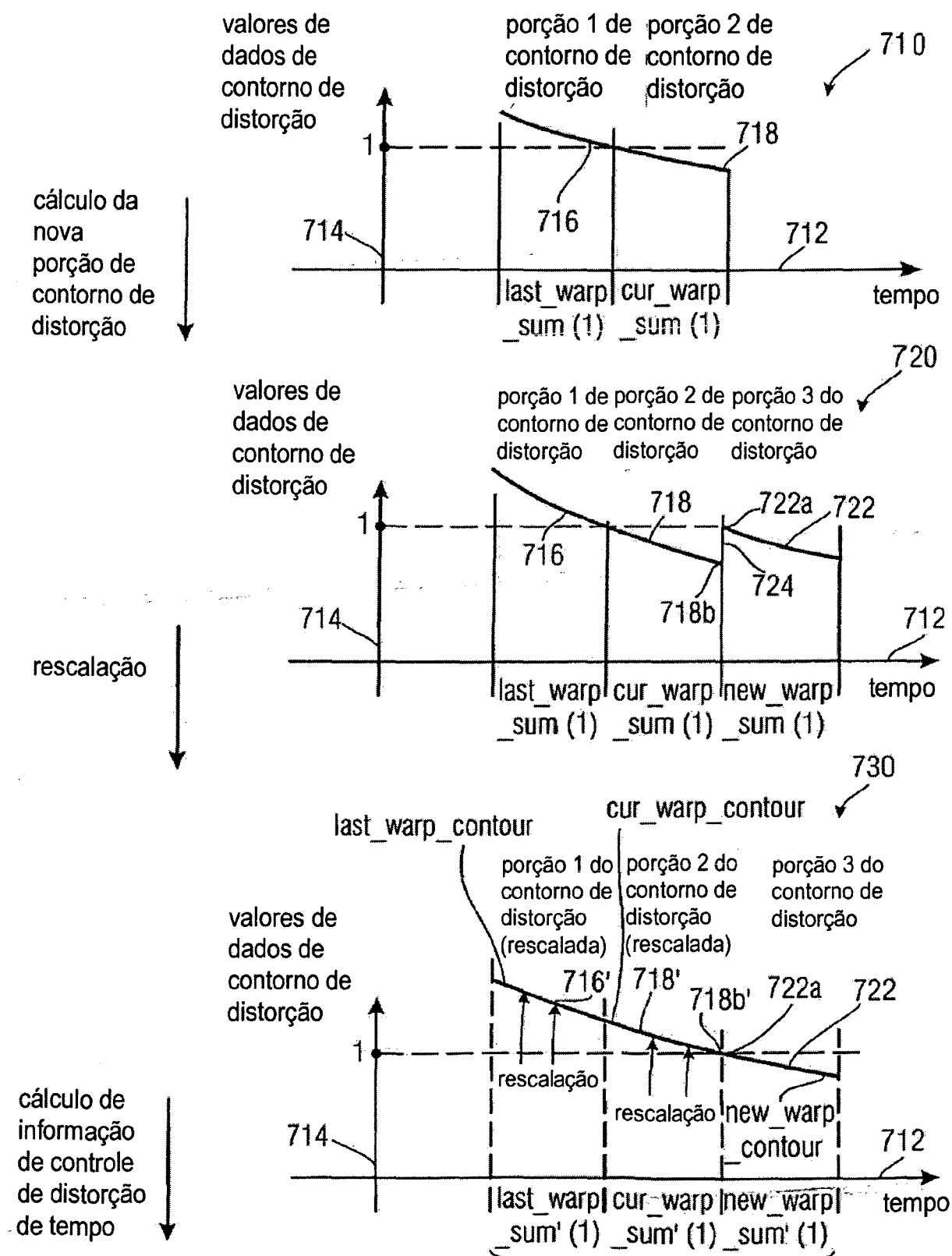


FIG 7A

seção de contorno de distorção de tempo usada para o cálculo de informação de controle de distorção de tempo

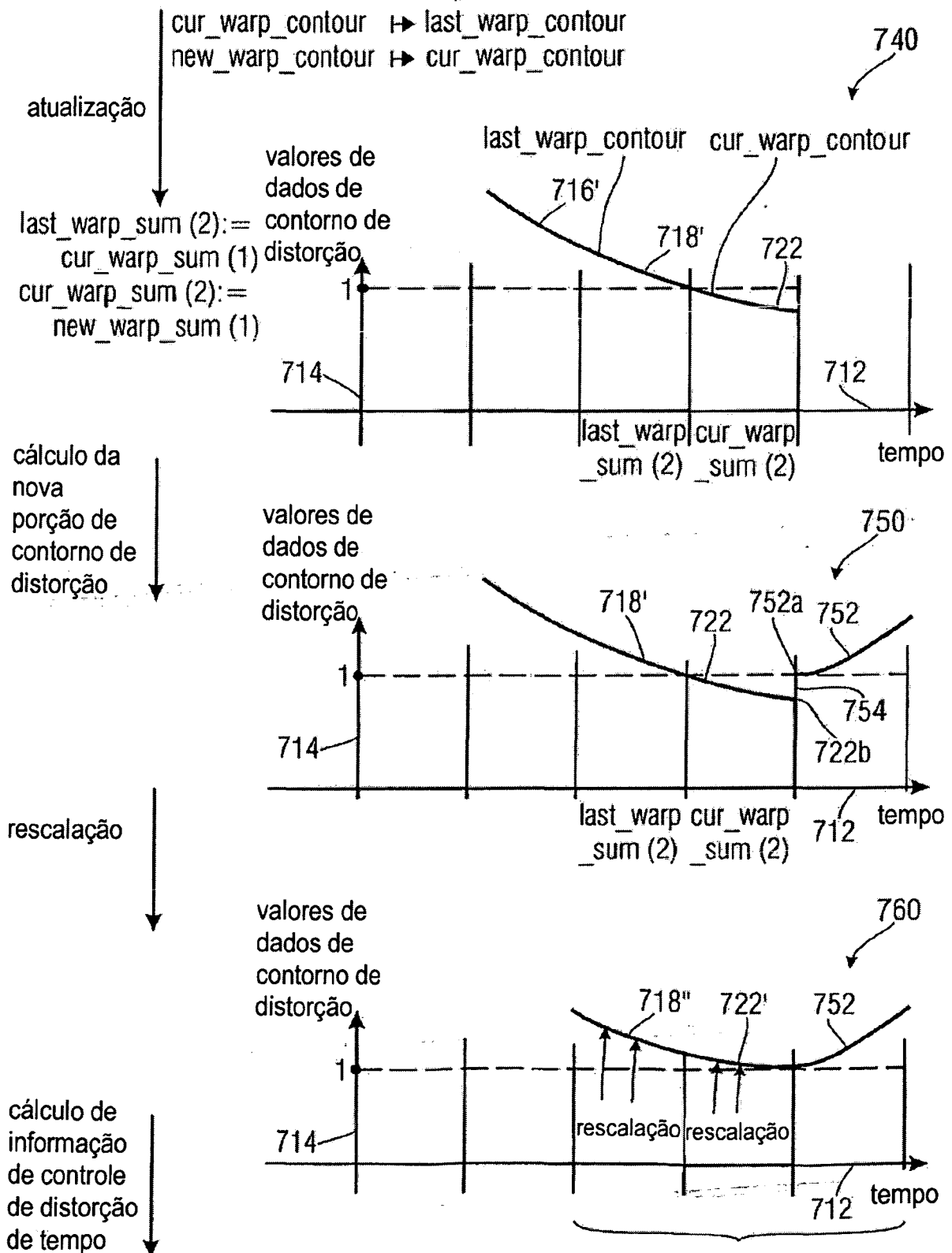


FIG 7B

seção de contorno de distorção de tempo
usada para o cálculo de informação de
controle de distorção de tempo

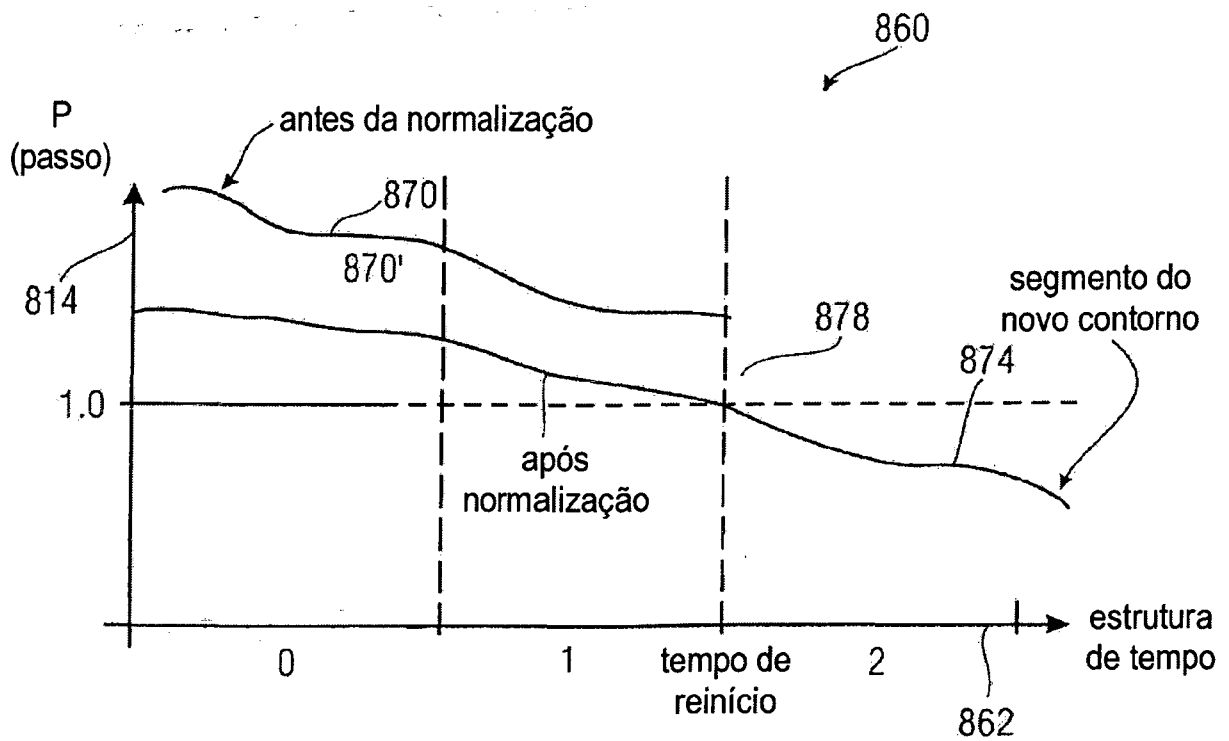
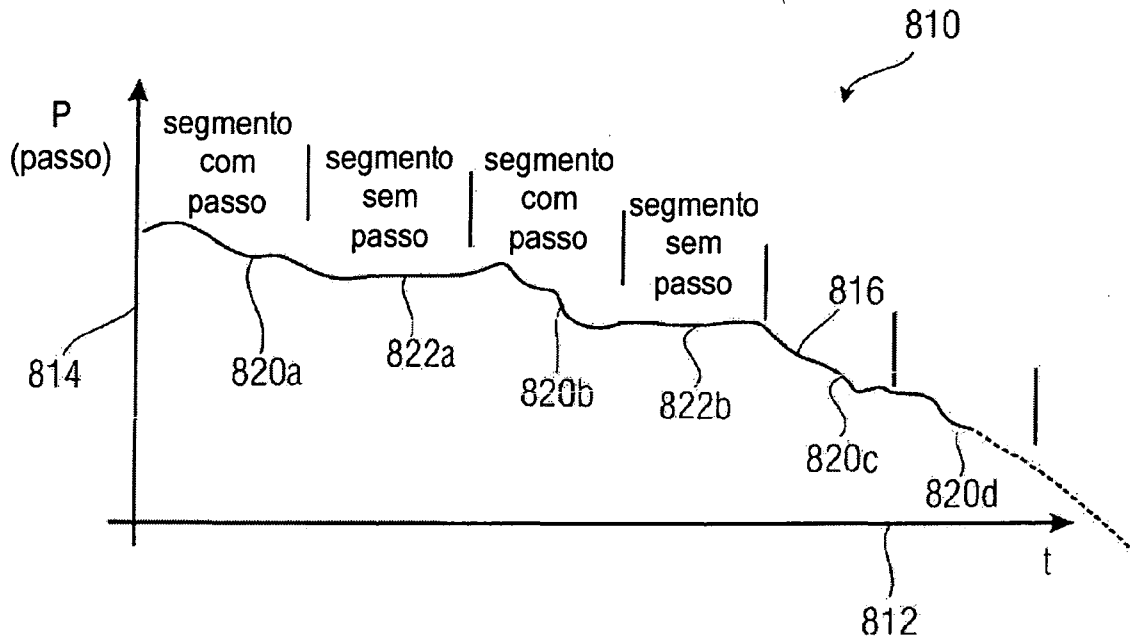


FIG 8

$$\text{warp_node_values}[i] = \begin{cases} 1 & \text{for tw_data_present} = 0, 0 \leq i \leq \text{num_tw_nodes} \\ 1 & \text{for tw_data_present} = 1, i = 0 \\ \prod_{k=0}^{i-1} \text{warp_values_tbl}[\text{tw_ratio}[k]] & \text{for tw_data_present} = 1, 0 \leq i \leq \text{num_tw_nodes} \end{cases}$$

```

for ( i = 0 ; i < num_tw_nodes ; i++ ) {
    d = (warp_node_values[i+1] - warp_node_values[i]) / interp_dist;
    for ( j = 0 ; j < interp_dist ; j++ ) {
        new_warp_contour [i*interp_dist + j] = warp_node_values[i-1] +
        (j+1)*d;
    }
}

```

FIG 9A-1

$$\text{norm_fac} = \frac{1}{\text{past_warp_contour}[2 \cdot n_long - 1]}$$

$$\text{past_warp_contour}[i] = \text{past_warp_contour}[i] \cdot \text{norm_fac} \text{ for } 0 \leq i < 2 \cdot n_long$$

$$\text{last_warp_sum} = \text{last_warp_sum} \cdot \text{norm_fac}$$

$$\text{cur_warp_sum} = \text{cur_warp_sum} \cdot \text{norm_fac}$$

930

← 932

← 934

← 936

$$\text{new_warp_sum} = \sum_{i=0}^{n_long-1} \text{new_warp_contour}[i]$$

← 940

13/38

FIG 9A-2

past_warp_contour[n] = warp_contour[n+n_long], for $0 \leq n < 2 \cdot n_long$
cur_warp_sum = new_warp_sum
last_warp_sum = cur_warp_sum

← 950
← 952
← 954

past_warp_contour[n] = 1, for $0 \leq n < 2 \cdot n_long$
cur_warp_sum = n_long
last_warp_sum = n_long

← 960
← 962
← 964

14/38

FIG 9B

warp_value_tbl

990


Índice (e.g. tw_ratio)	Valor (e.g. warp_value_tbl [tw_ratio])
0	0.982857168
1	0.988571405
2	0.994285703
3	1
4	1.0057143
5	1.01142859
6	1.01714289
7	1.02285719

FIG 9C

cálculo do contorno de tempo

$$\text{time_contour}[i] = \begin{cases} -w_{\text{res}} \cdot \text{last_warp_sum} & \text{for } i = 0 \\ w_{\text{res}} \left(-\text{last_warp_sum} + \sum_{k=0}^{i-1} \text{warp_contour}[k] \right) & \text{for } 0 < i \leq 3 \cdot n_{\text{long}} \end{cases} \quad \leftarrow 1010$$

where $w_{\text{res}} = \frac{n_{\text{long}}}{\text{cur_warp_sum}} \quad \leftarrow 1012$

funções auxiliares

```
warp_time_inv (time_contour[], t_warp) {
    i = 0;
    if (t_warp < contour [0] ) {
        return NOTIME;
    }
    while ( t_warp > time_contour [i+1] ) {
        i++;
    }
    return (i + (t_warp - time_contour [i]) / (time_contour [i+1] -
time_contour[i]));
}
```

1020

```
warp_inv_vec (time_contour[], t_start, n_samples, sample_pos[]) {
    t_warp = t_star;
    j = 0;
    while ((i = floor (warp_time_inv (time_contour, t_warp-0.5))) == NOTIME) {
        t_warp += 1;
        j++;
    }
    while (j < n_samples && (t_warp + 0.5) < time_contour [3*n_long] ) {
        while (t_warp < time_contour [i+1]) {
            i++;
        }
        sample_pos[j] =
            i + (t_warp - time_contour[i]) / (time_contour[i+1] -
time_contour[i]);
        j++;
        t_warp += 1;
    }
}
```

1022

FIG 10A

cálculo de posição de amostra
 cálculo de comprimento de transição
 cálculo de primeira posição e de última posição

```

t_start = n_long - 3*N_f/4 - ip_len_2s + 0.5

warp_inv_vec (time_contour,
               t_start,
               N_f + 2*ip_len_2s,
               sample_pos [] ) ;
  }
  }
  }

if ( last_warp_sum > cur_warp_sum ) {
    warped_trans_len_left = n_long/2;
}
else {
    warped_trans_len_left = n_long/2*last_warp_sum/cur_warp_sum;
}
if ( new_warpSum > cur_warp_sum ) {
    warped_trans_len_right = n_long/2;
}
else {
    warped_trans_len_right = n_long/2*new_warp_sum/cur_warp_sum;
}

switch ( window_sequence ) {
    case LONG_START_SEQUENCE:
        warped_trans_len_right /= 8;
        break;
    case LONG_STOP_SEQUENCE:
    case STOP_1152_SEQUENCE:
        warped_trans_len_right /= 8;
        break;
    case EIGHT_SHORT_SEQUENCE:
    case STOP_START_SEQUENCE:
    case STOP_START_1152_SEQUENCE:
        warped_trans_len_right /= 8;
        warped_trans_len_left /= 8;

        break;
}

first_pos = ceil (N_f/4 - 0.5 - warped_trans_len_left);
last_pos = floor (3*N_f/4 - 0.5 + warped_trans_len_right);
  
```

1030

1032

1034

1036

cálculo de forma da janela

$$N_{os} = 2 \cdot n_long \cdot os_factor_win$$

← 1040

$$W_{KBD}\left(n - \frac{N_{os}}{2}\right) = \sqrt{\frac{\sum_{p=0}^{N_{os}-n-1} [W(p,a)]}{\sum_{p=0}^{N_{os}/2} [W(p,a)]}} \quad \text{para } \frac{N_{os}}{2} \leq n < N_{os}$$

← 1042

onde:

$$W(n,a) = \frac{I_0\left[\pi\alpha\sqrt{1.0 - \left(\frac{n - N_{os}/4}{N_{os}/4}\right)^2}\right]}{I_0[\pi\alpha]} \quad \text{para } 0 \leq n \leq \frac{N_{os}}{2}$$

$$I_0[x] = \sum_{k=0}^{\infty} \left[\frac{\left(\frac{x}{2}\right)^k}{k!} \right]^2$$

α = fator alfa da janela central, $\alpha = 4$

← 1044

$$W_{SIN}\left(n - \frac{N_{os}}{2}\right) = \sin\left(\frac{\pi}{N_{os}}\left(n + \frac{1}{2}\right)\right) \quad \text{para } \frac{N_{os}}{2} \leq n < N_{os}$$

← 1046

$$\text{left_window_shape}[n] = \begin{cases} W_{KBD}[n], & \text{se window_shape_previous_block} = 1 \\ W_{SIN}[n], & \text{se window_shape_previous_block} = 0 \end{cases}$$

← 1048

$$\text{right_window_shape}[n] = \begin{cases} W_{KBD}[n], & \text{se window_shape} = 1 \\ W_{SIN}[n], & \text{se window_shape} = 0 \end{cases}$$

← 1050

FIG 10C

aplicação de janela ("SEQUÊNCIA DE OITO TIROS")

```
tw_windowing_short (X [], z [], first_pos, last_pos, warped_trans_len_left, warped_trans_len_right, left_window_shape [], right_window_shape []) {
```

```
    offset = n_long - 4*n_short - n_short/2;
```

```
    tr_scale_l = 0.5*n_long/warped_trans_len_left*os_factor_win;
```

```
    tr_pos_l = warped_trans_len_left + (first_pos - n_long/2) + 0.5*tr_scale_l;
```

```
    tr_scale_r = 8*os_factor_win;
```

```
    tr_pos_r = tr_scale_r/2;
```

```
    for (i = 0; i < n_shot; i++) {
```

```
        z[i] = X[0][i];
```

```
    }
```

```
    for (i=0; i<first_pos; i++)
```

```
        z[i] = 0.;
```

```
    for (i = n_long-1-first_pos; i >= first_pos; i--) {
```

```
        z[i] *= left_window_shape[floor(tr_pos_l)];
```

```
        tr_pos_l += tr_scale_l;
```

```
    }
```

```
    for (i = 0; i < n_short; i++) {
```

```
        z[offset+i+n_short] =
```

```
            X[0][i+n_short]*right_window_shape[floor(tr_pos_r)];
```

```
        tr_pos_r += tr_scale_r;
```

```
    }
```

```
    offset += n_short;
```

```
    for (k = 1; k < 7; k++) {
```

```
        tr_scale_l = n_short*os_factor_win;
```

```
        tr_pos_l = tr_scale_l/2;
```

```
        tr_pos_r = os_factor_win*n_long-tr_pos_l;
```

```
        for (i = 0; i < n_short; i++) {
```

```
            z[i + offset] += X[k][i]*right_window_shape[floor(tr_pos_r)];
```

```
            z[offset + n_short + i] =
```

```
                X[k][n_short + i]*right_window_shape[floor(tr_pos_l)];
```

```
            tr_pos_l += tr_scale_l;
```

```
            tr_pos_r -= tr_scale_l;
```

```
        }
```

```
        offset += n_short;
```

```
    }
```

1060 →

FIG 10D-1

```

tr_scale_l = n_short*os_factor_win;
tr_pos_l = tr_scale_l/2;

for (i = n_short - 1; i >= 0; i--) {
    z[i + offset] += X[7][i]*right_window_shape[(int) floor(tr_pos_l)];
    tr_pos_l += tr_scale_l;
}

for (i = 0; i < n_short; i++) {
    z[offset + n_short + i] = X[7][n_short + i];
}

tr_scale_r = 0.5*n_long/warpedTransLenRight*os_factor_win;
tr_pos_r = 0.5*tr_scale_r+.5;

tr_pos_r = (1.5*n_long-(float) wEnd - 0.5 + warpedTransLenRight)*tr_scale_r;

for (i = 3*n_long-1-last_pos; i <= wEnd; i++) {
    z[i] *= right_window_shape[floor(tr_pos_r)];
    tr_pos_r += tr_scale_r;
}

for (i=lsat_pos+1; i<2*n_long; i++)
    z[i] = 0.;

```

FIG 10D-2

aplicação de janela (TODOS OS OUTROS)

```
tw_windowing_long (X [] [], z [], first_pos, last_pos, warped_trans_len_left, warped_trans_len_right, left_window_shape [], right_window_shape[]) {
```

```
    for (i = 0 ; i < first_pos; i++)
        z [i] = 0.;
    for (i = last_pos+1 ; i < N_f; i++)
        z [i] = 0.;
```

```
    tr_scale = 0.5*n_long/warped_trans_len_left*os_factor_win;
    tr_pos = (warped_trans_len_left+first_pos-N_f/4)+0.5)*tr_scale;
```

```
1064 →    for (i = N_f/2-1-first_pos ; i >= first_pos; i--) {
        z [i] = X[0] [i]*left_window_shape [floor (tr_pos)];
        tr_pos += tr_scale;
    }
```

```
    tr_scale = 0.5*n_long/warped_trans_len_right*os_factor_win;
    tr_pos = (3*N_f/4 - last_pos-0.5+warped_trans_len_right)*tr_scale;
```

```
    for (i = 3*N_f/2-1- last_pos; i <= last_pos; i++) {
        z [i] = X[0] [i]*right_window_shape [floor (tr_pos)];
        tr_pos += tr_scale;
    }
```

```
}
```

FIG 10E

reamostragem variável com o tempo

$$b[n] = I_0[\alpha]^{-1} \cdot I_0 \left[\alpha \sqrt{1 - \frac{n^2}{ip_len_2^2}} \right] \cdot \frac{\sin \left(\frac{\pi n}{os_factor_resamp} \right)}{\frac{\pi n}{os_factor_resamp}} \cdot \text{for } 0 \leq n < ip_size - 1 \quad \left. \vphantom{\frac{\sin \left(\frac{\pi n}{os_factor_resamp} \right)}{\frac{\pi n}{os_factor_resamp}}} \right\} \leftarrow 1070$$

$$\alpha = 8$$

$$zp[n] = \begin{cases} 0, & \text{for } 0 \leq n < ip_len_2s \\ z[n-ip_len_2s], & \text{for } ip_len_2s \leq n < N_f + ip_len_2s \\ 0, & \text{for } 2 \cdot N_f + ip_len_2s \leq n < N_f + 2 \cdot ip_len_2s \end{cases} \quad \left. \vphantom{\begin{cases} 0, \\ z[n-ip_len_2s], \\ 0, \end{cases}} \right\} \leftarrow 1072$$

```

offset_pos = 0.5;
num_samples_in = N_f + 2 * ip_len_2s;
num_samples_out = 3 * n_long;
j_center = 0;
for (i = 0; i < numSamplesOut; i++) {
    while (j_center < num_samples_in && sample_pos[j_center] - offset_pos <= i)
        j_center++;
    j_center--;
    y[i] = 0;
    if (j_center < num_samples_in - 1 && j_center > 0) {
        frac_time = floor((i - (sample_pos[j_center] - offset_pos))
                        / (sample_pos[j_center + 1] - sample_pos[j_center])
                        * os_factor);
        j = ip_len_2s * os_factor + frac_time;

        for(k=j_center-ip_len_2s; k<=j_center+ip_len_2s; k++) {
            if (k >= 0 && k < num_samples_in)

                y[i] += b[abs(j)] * zp[k];
            j -= os_factor;
        }
    }
    if (j_center < 0)
        j_center++;
}

```

← 1074

FIG 10F

a) EIGHT_SHORT_SEQUENCE, LONG_START_SEQUENCE,
STOP_START_SEQUENCE, STOP_START_1152_SEQUENCE seguido de uma sequência LPD

$$W_{\text{corr}}(n) = \frac{W_{\text{short}}(n)}{W_{\text{FD_LPD}}(n)}, \text{ para } 0 \leq n < \frac{n_{\text{short}}}{2} \quad \leftarrow 1080a$$

$$W_{\text{short}}(n) = \begin{cases} W_{\text{SIN_RIGHT}, n_{\text{short}}}(n), & \text{se } \text{window_shape} = 0 \\ W_{\text{KBD_RIGHT}, n_{\text{short}}}(n), & \text{se } \text{window_shape} = 1 \end{cases} \quad \leftarrow 1080b$$

$$W_{\text{FD_LPD}}(n) = \begin{cases} W_{\text{SIN_RIGHT}, n_{\text{short}}/2}(n), & \text{se } \text{window_shape} = 0 \\ W_{\text{KBD_RIGHT}, n_{\text{short}}/2}(n), & \text{se } \text{window_shape} = 1 \end{cases}$$

é aplicado:

$$y'_{i,n} = \begin{cases} y[n] & \text{para } 0 \leq n < \frac{5n_{\text{long}} - n_{\text{short}}}{2} \\ y[n] \cdot W_{\text{corr}}\left(n - \left(\frac{5n_{\text{long}} - n_{\text{short}}}{2}\right)\right) & \text{para } \frac{5n_{\text{long}} - n_{\text{short}}}{2} \leq n < \frac{5n_{\text{long}}}{2} \\ 0 & \text{para } \frac{5n_{\text{long}}}{2} \leq n < 3n_{\text{long}} \end{cases} \quad \leftarrow 1082$$

b) todos os outros casos

nada tem que ser feito

$$y'_{i,n} = y[n] \text{ for } 0 \leq n < 3 \cdot n_{\text{long}} \quad \leftarrow 1084$$

sobreposição e adição

$$\text{out}_{i,n} = \begin{cases} y'_{i,n} + y'_{i-1, n+n_{\text{long}}} + y'_{i-2, n+2 \cdot n_{\text{long}}} & \text{para } 0 \leq n < n_{\text{long}}/2 \\ y'_{i,n} + y'_{i-1, n+n_{\text{long}}} & \text{para } n_{\text{long}}/2 \leq n < n_{\text{long}} \end{cases} \quad \leftarrow 1086$$

Elementos de dados

<code>tw_data()</code>	contém a informação lateral necessária para decodificar e aplicar a MDCT de tempo distorcido em um <code>fd_channel_stream()</code> para elementos SCE e CPE. O <code>fd_channel_streams</code> de um <code>channel_pair_element()</code> pode compartilhar um <code>tw_data()</code> comum
<code>tw_data_present</code>	1 bit indicando que um contorno de distorção não-plano é transmitido nesta estrutura
<code>tw_ratio[]</code>	índice do livro de códigos da proporção de distorção para o nó i
<code>window_sequence</code>	2 bit indicando qual sequência de janela (isto é, tamanho de bloco) é usada
<code>window_shape</code>	1 bit indicando qual função de janela está selecionada

Elementos de ajuda

<code>warp_node_values[]</code>	valores de nó de contorno de distorção decodificados
<code>warp_value_tbl</code>	vide tabela "warp_value_tbl"
<code>warp_value_tbl[]</code>	tabela de quantização para os valores de proporção de nó de distorção, favor ver Tabela da Figura 9C
<code>new_warp_contour[]</code>	contorno de distorção decodificado e interpolado para esta estrutura (amostras n_long)
<code>past_warp_contour[]</code>	contorno de distorção passada (amostras $2*n_long$), compreendendo <code>last_warp_contour</code> e <code>cur_warp_contour</code>
<code>norm_fac</code>	fator de normalização para o <code>past_warp_contour</code>
<code>warp_contour[]</code>	contorno de distorção completo (amostras $3*n_long$)
<code>last_warp_sum</code>	soma da primeira parte do contorno de distorção
<code>cur_warp_sum</code>	soma da parte média do contorno de distorção
<code>next_warp_sum</code>	soma da última parte do contorno de distorção
<code>time_contour[]</code>	contorno de tempo completo ($3*n_long+1_samples$)
<code>sample_pos[]</code>	posições das amostras distorcidas em uma escala de tempo linear ($2*n_long\ samples+2*ip_len_2s$)
<code>X[w][]</code>	saída da MDCT para a janela w

FIG 11A

Elementos de ajuda (continuação)

$z[]$	vetor de tempo em janela e (opcionalmente) internamente sobreposto para uma estrutura no domínio distorcido de tempo
$zp[]$	$z[]$ com preenchimento com zeros
$y[]$	vetor de tempo para uma estrutura no domínio de tempo linear após reamostragem
$y'_{i,n}$	vetor de tempo para a estrutura i depois de pós-processamento
$out[]$	vetor de saída para uma estrutura
$b[]$	resposta de impulso do filtro de reamostragem
N	comprimento de janela de síntese, vide abaixo
N_f	comprimento de estrutura, seja 2304 no caso de STOP_1152_SEQUENCE, STOP_START_1152_SEQUENCE ou 2048 para todas as outras sequências de janela

Constantes

num_tw_nodes	16
os_factor_win	16
os_factor_resamp	128
ip_len_2s	64
ip_len_2	$os_factor_resamp * ip_len_2s + 1$
ip_size	$ip_len_2 + os_factor_resamp$
n_long	1024 (960)
n_short	128 (120)
$Interp_dist$	n_long / num_tw_nodes
NOTIME	-100000

Comprimento de janela 2304:

$$N = \begin{cases} 2304, & \text{if STOP_1152_SEQUENCE} \\ 2304, & \text{if STOP_START_1152_SEQUENCE} \end{cases}$$

Comprimento de janela 2048:

$$N = \begin{cases} 2048, & \text{if ONLY_LONG_SEQUENCE} \\ 2048, & \text{if LONG_START_SEQUENCE} \\ 256, & \text{if EIGHT_SHORT_SEQUENCE} \\ 2048, & \text{if LONG_STOP_SEQUENCE} \\ 2048, & \text{if STOP_START_SEQUENCE} \end{cases}$$

} ← 1190

FIG 11B-2

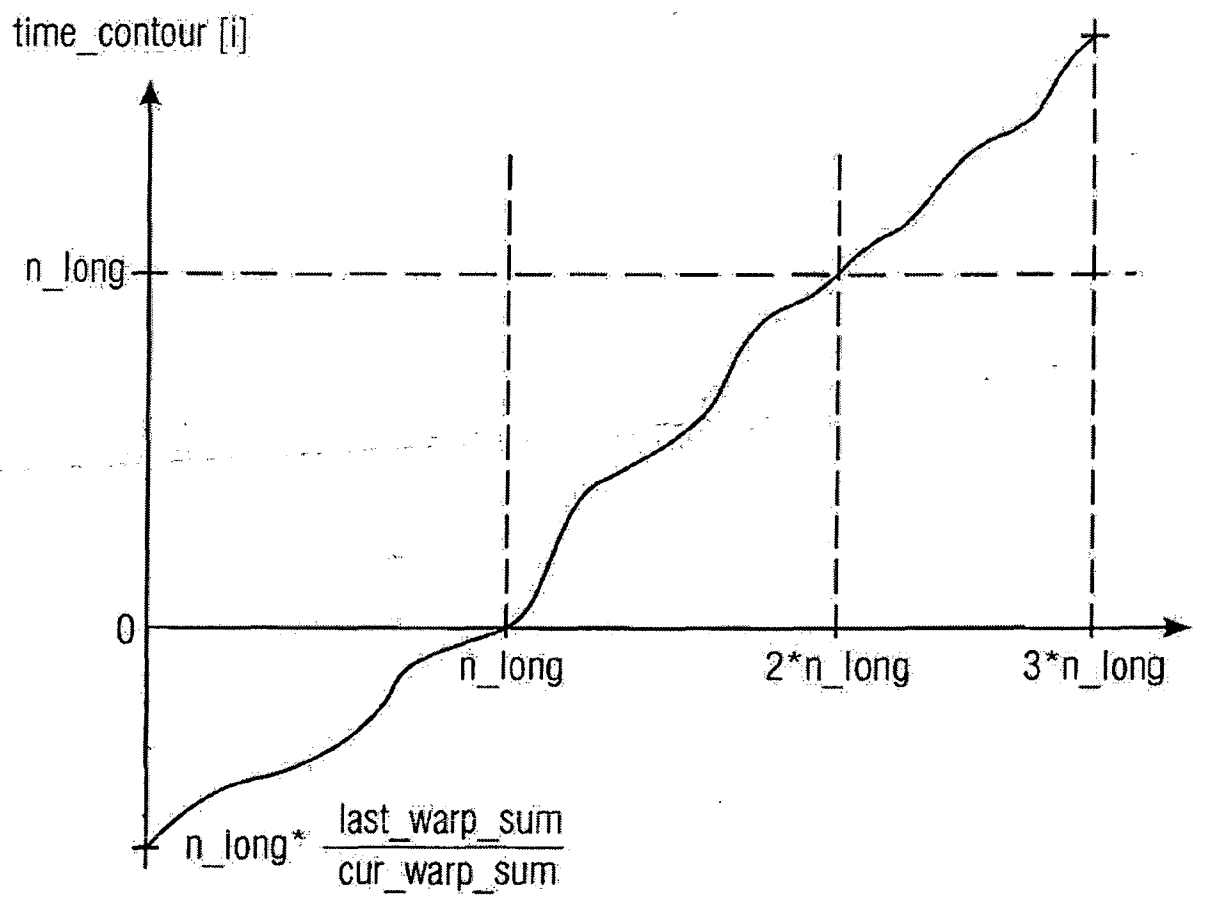


FIG 12

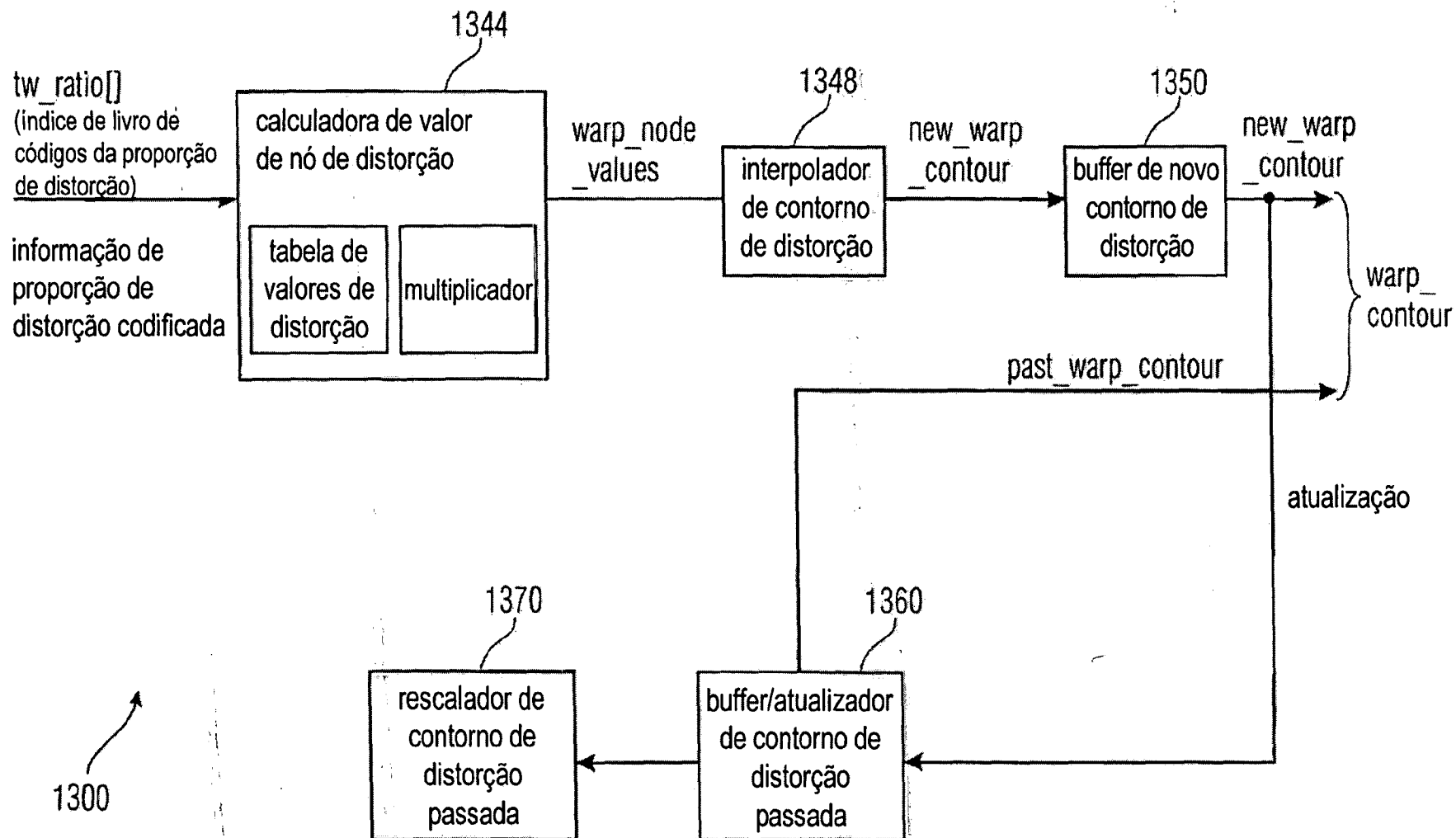


FIG 13

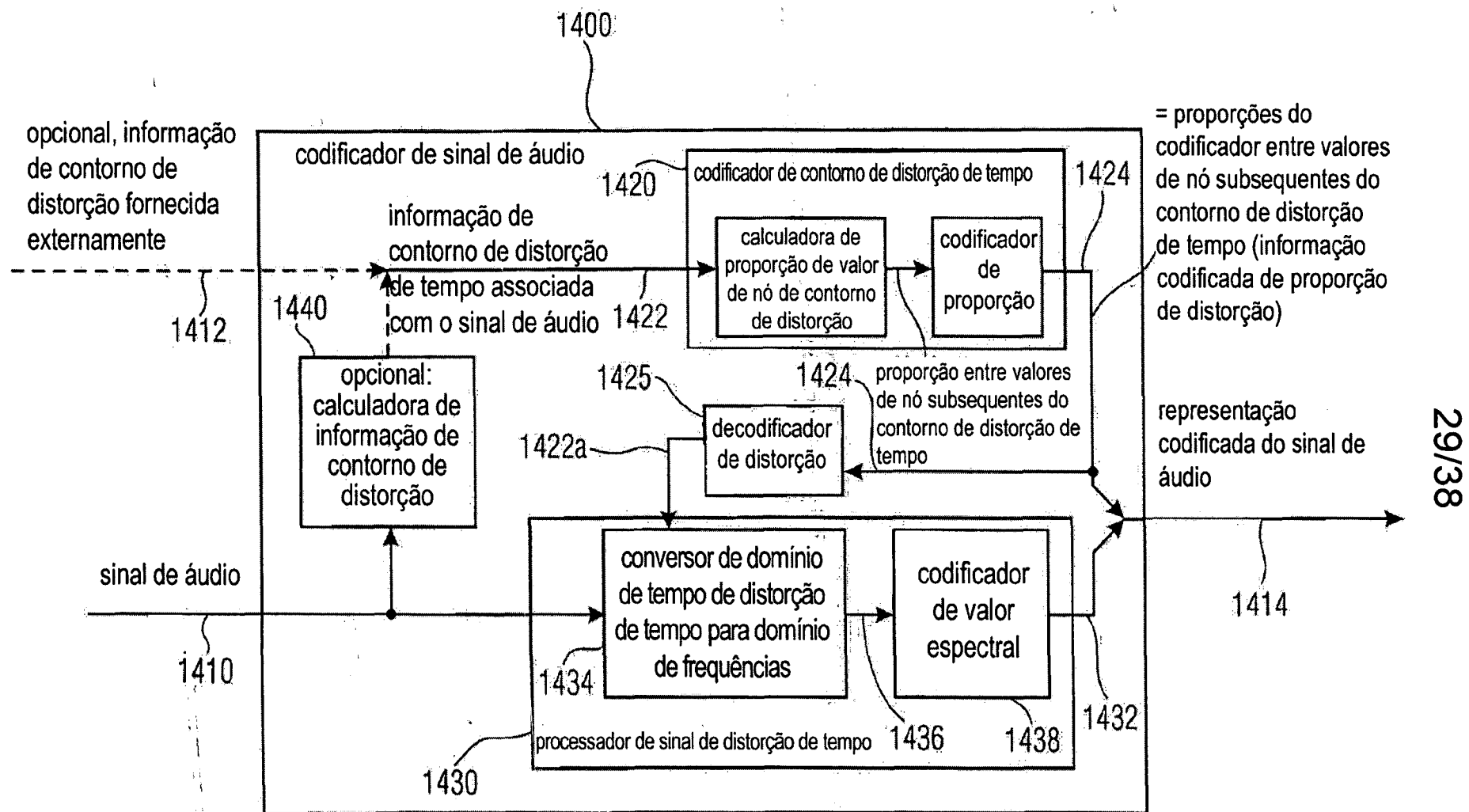


FIG 14

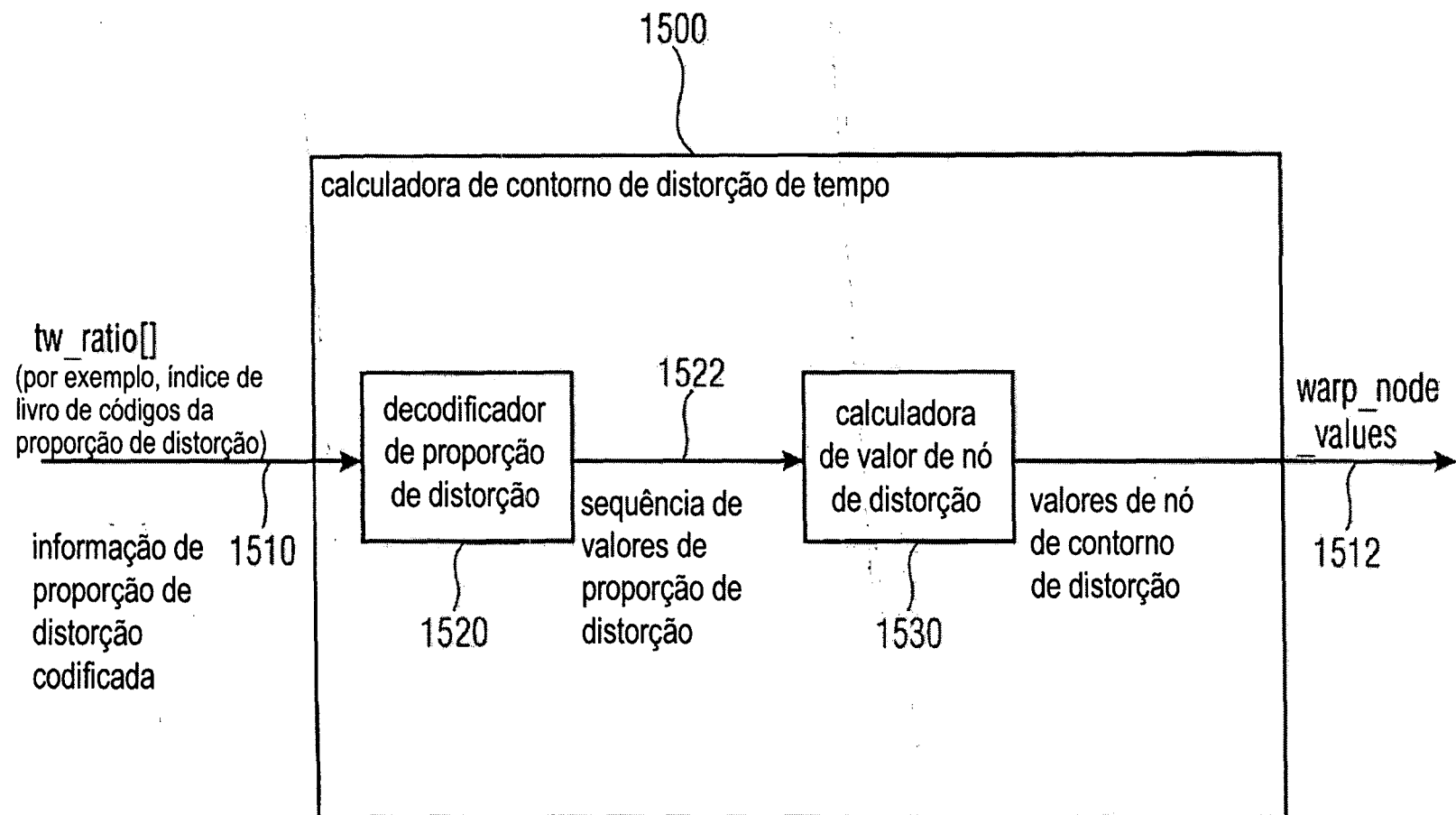
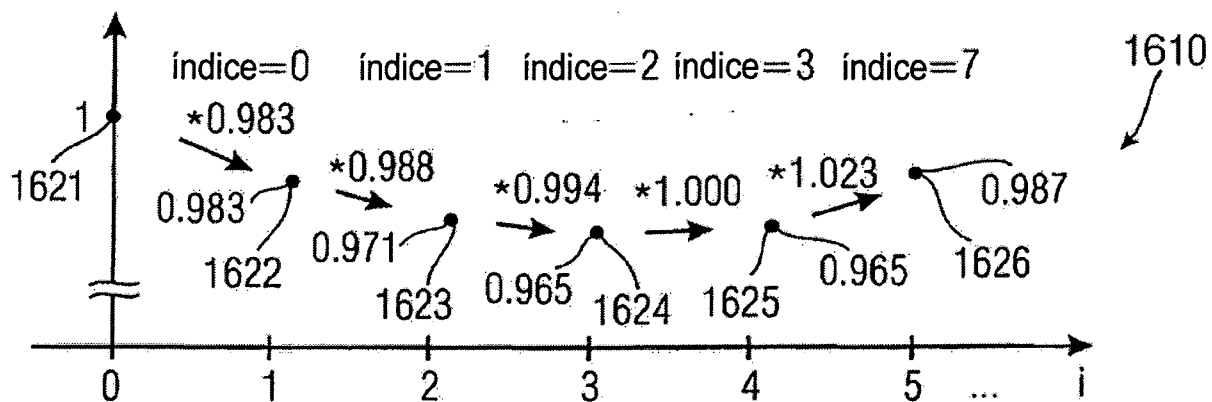


FIG 15

warp_node_values



interpolação linear

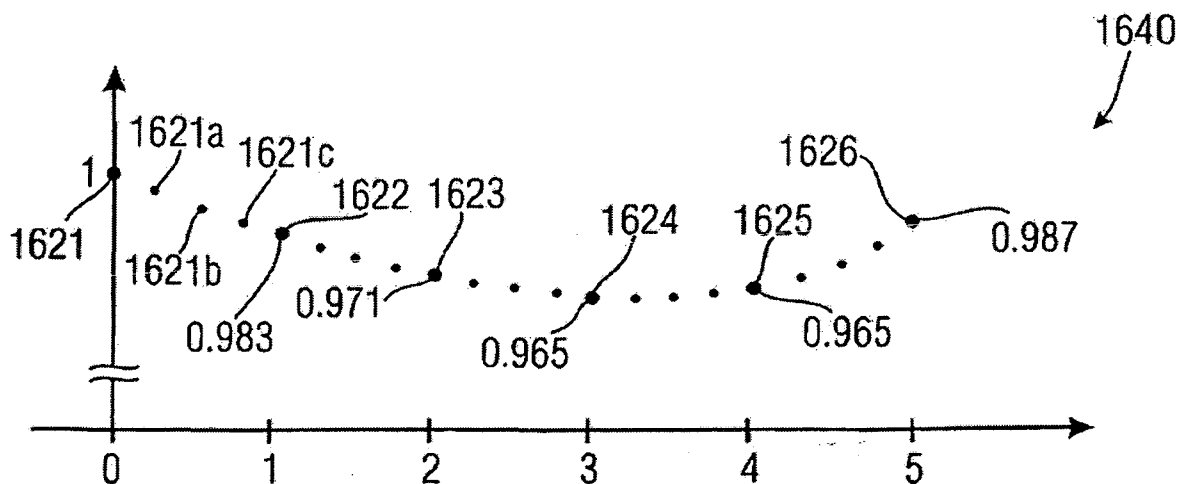


FIG 16A

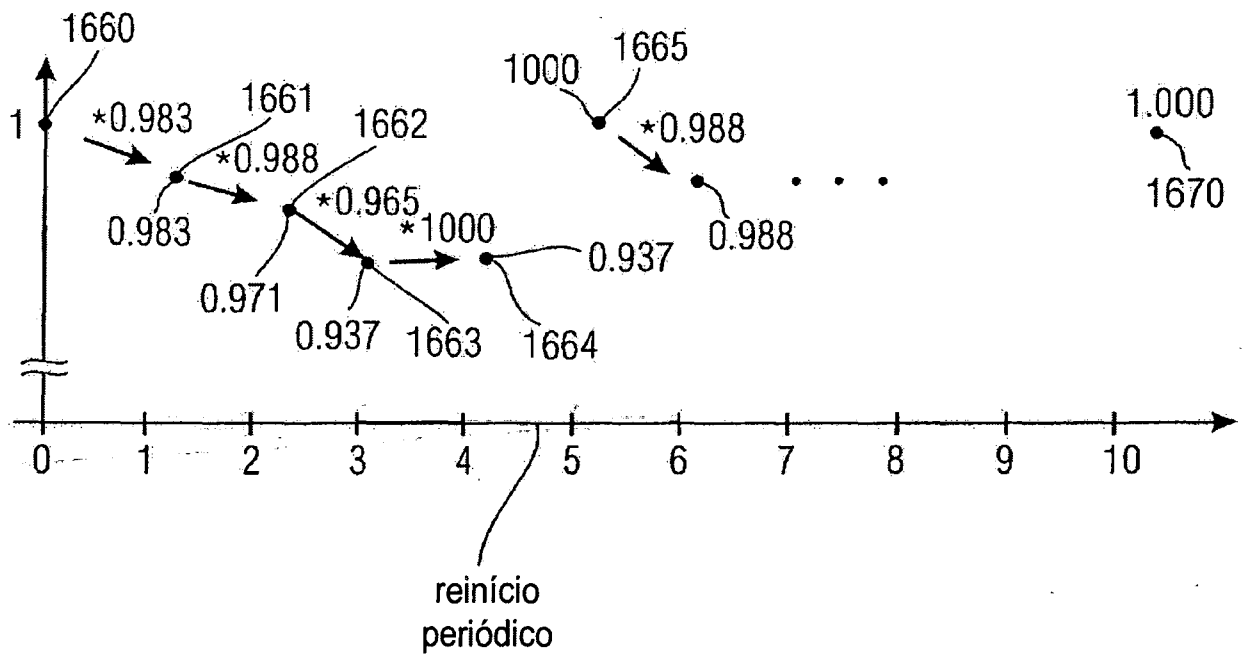


FIG 16B

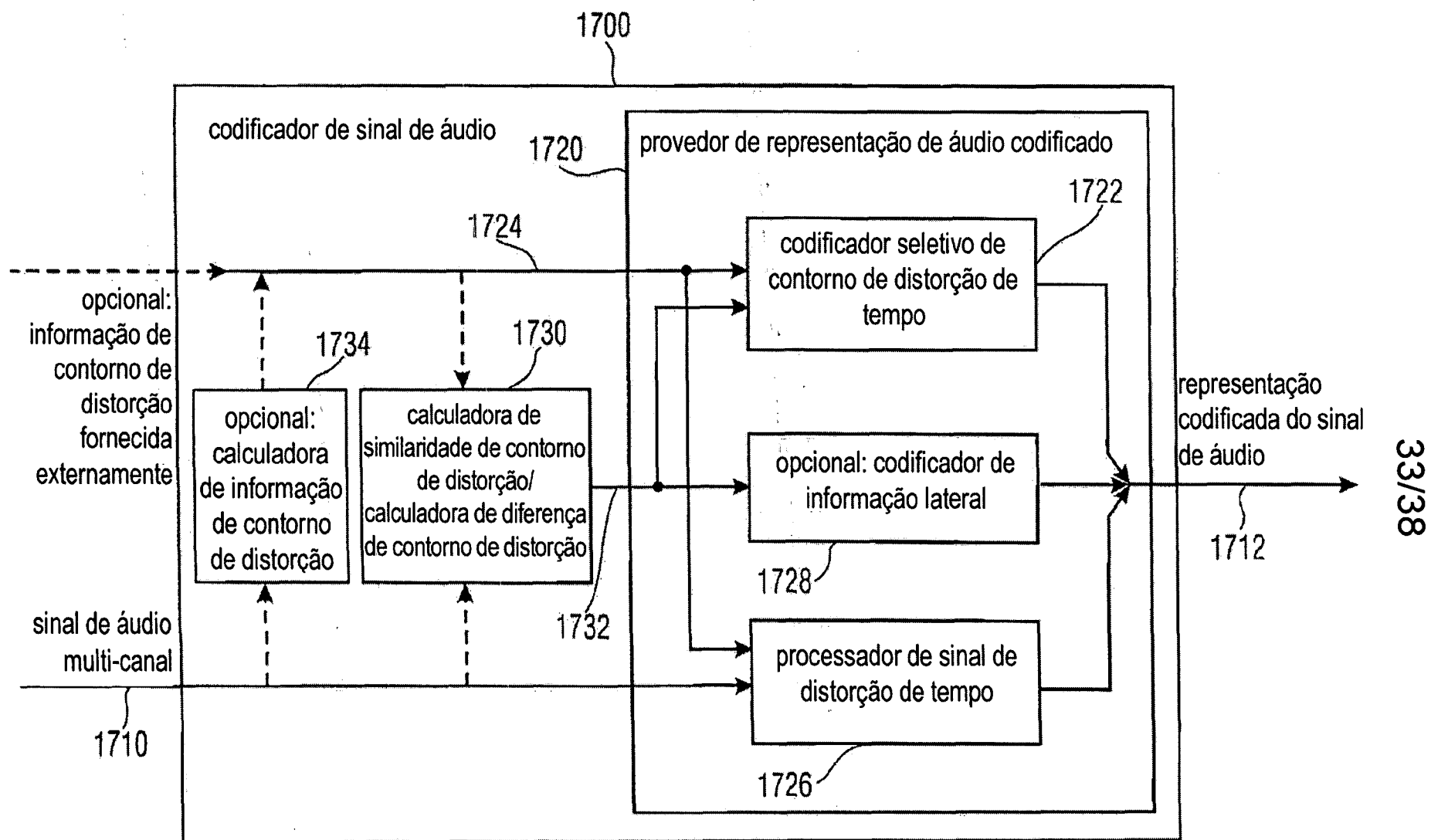


FIG 17

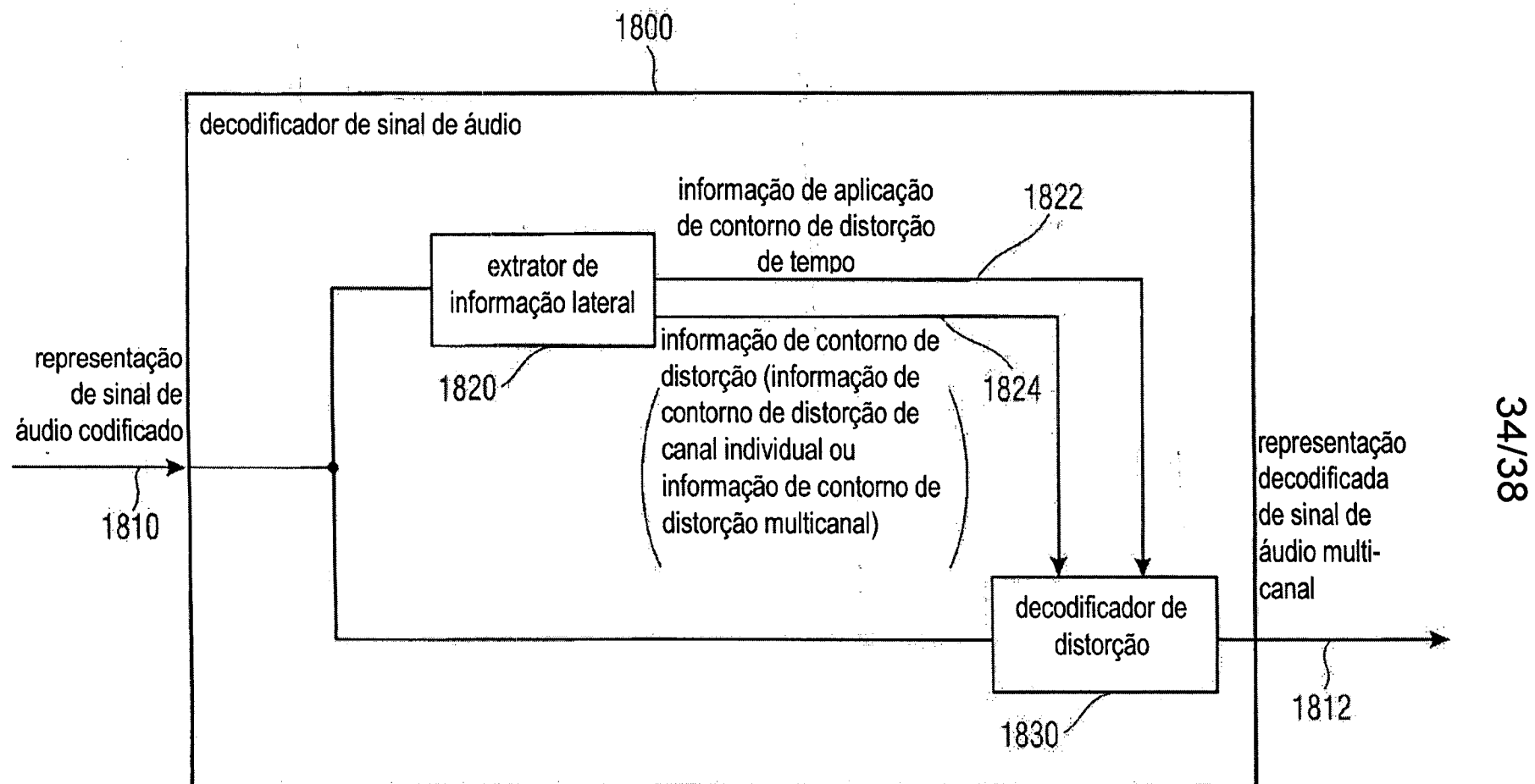


FIG 18

```
usac_raw_data_block ()  
{  
    single_channel_element ();  
    or  
    channel_pair_element ();  
    or  
    single_channel_element ();  
    and  
    channel_pair_element ();  
}
```

FIG 19A

```
single_channel_element ()  
{  
    fd_channel_stream (*, *, *);  
}
```

FIG 19B

```
channel_pair_element
{
    if (tw_mdct) {
        common_tw;
        if (common_tw) {
            tw_data();
        }
    }
    fd_channel_stream (*, *, *);
    fd_channel_stream (*, *, *);
}
```

FIG 19C

```

fd_channel_stream (*,*,*);
{
    global gain;
    if (tw_mdct){
        if (not common_tw) {
            tw_data ();
        }
    }
    scale_factor_data ();
    ac_spectral_data ();
}

```

FIG 19D

Sintaxe	Número de bits Mnemônico	
tw_data()		
{		
tw_data_present	1	ulmsfb
if (tw_data_present == 1) {		
for (i= 1 ; i < num_tw_nodes ; i++) {		
tw_ratio[i];	3	ulmsbf
}		
}		
}		

FIG 19E

Syntaxe de pitch_data()

Sintaxe	Número de bits	Mnemônico
<pre>pitch_data() { activePitchData if (activePitchData == 1) { for (i= 1 ; i < numPitches ; i++) { pitchidx[i]; } } }</pre>	1 1 numPitchBits	uimssf uimssf uimssf

FIG 19F