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

(11) BR 112015008114-2 B1

(22) Data do Depósito: 20/09/2013

(45) Data de Concessão: 17/05/2022

(54) Título: APARELHO E MÉTODO PARA SÍNTESE EFICIENTE DE SINUSOIDES E VARREDURAS EMPREGANDO PADRÕES ESPECTRAIS

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

(30) Prioridade Unionista: 10/10/2012 US 61/712,013; 21/12/2012 EP 12199266.3.

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

(72) Inventor(es): SASCHA DISCH; BENJAMIN SCHUBERT; RALF GEIGER; BERND EDLER; MARTIN DIETZ.

(86) Pedido PCT: PCT EP2013069592 de 20/09/2013

(87) Publicação PCT: WO 2014/056705 de 17/04/2014

(85) Data do Início da Fase Nacional: 10/04/2015

(57) Resumo: APARELHO E MÉTODO PARA SÍNTESE EFICIENTE DE SINUSOIDES E VARREDURAS EMPREGANDO PADRÕES ESPECTRAIS. Um aparelho para gerar um sinal de saída de áudio com base em um espectro do sinal de áudio codificado é fornecido. O aparelho compreende uma unidade de processamento (115) para processar o espectro do sinal de áudio codificado para obter um espectro do sinal de áudio decodificado, compreendendo uma pluralidade de coeficientes espectrais, caracterizado por cada um dos coeficientes espectrais ter uma localização espectral dentro do espectro do sinal de áudio codificado e um valor espectral, em que os coeficientes espectrais são sequencialmente organizados, de acordo com sua localização espectral dentro do espectro do sinal de áudio codificado, de modo que os coeficientes espectrais formem uma sequência de coeficientes espectrais. Além disso, o aparelho compreende um determinador dos pseudocoeficientes (125) para determinar um ou mais pseudocoeficiente(s) do espectro do sinal de áudio decodificado, cada um dos pseudocoeficientes tendo um valor espectral.

"APARELHO E MÉTODO PARA SÍNTESE EFICIENTE DE SINUSOIDES E VARREDURAS EMPREGANDO PADRÕES ESPECTRAIS".

DESCRIÇÃO

[0001] A presente invenção refere-se à codificação do sinal de áudio, decodificação e processamento e, em particular, à síntese eficiente de sinusoides e varreduras empregando padrões espectrais.

[0002] O processamento do sinal de áudio torna-se cada vez mais importante. Desafios surgem conforme codecs de áudio perceptuais modernos são necessários para entregar qualidade de áudio satisfatória nas taxas de bit progressivamente baixas. Adicionalmente, muitas vezes a latência permitida também é muito baixa, por exemplo, para aplicações de comunicação bidirecionais ou jogos distribuídos, etc.

[0003] A forma de onda moderna que preserva os codificadores de áudio de transformada geralmente possui melhorias parametricamente codificadas, como substituição de ruído ou extensão da largura de banda. Além destas ferramentas paramétricas bem conhecidas, também pode ser desejável sintetizar os tons sinusoidais neste decodificador da informação adicional paramétrica. A complexidade do cálculo computacional é sempre um critério importante no desenvolvimento do codec visto que uma baixa complexidade é essencial para uma ampla aceitação e instalação de um codec. Assim, as formas eficientes para gerar estes tons são necessárias.

[0004] Por exemplo, codecs de áudio MPEG-D USAC

(MPEG-D = *Moving Picture Experts Group-D* | Grupo de Especialistas em Imagem com Movimento; USAC = *Unified Speech and Audio Coding* | Fala Unificada e Codificação de Áudio) geralmente alternam entre codificação preditiva de domínio de tempo e codificação de domínio de transformação, independentemente do conteúdo musical ser ainda predominantemente codificado no domínio de transformação. Em baixas taxas de bit, por exemplo < 14 kbit/s, componentes tonais em itens musicais geralmente soam ruim quando codificados através de codificadores de transformada, que torna a tarefa de codificar o áudio em uma qualidade satisfatória ainda mais desafiante.

[0005] Além disso, restrições de baixo atraso geralmente levam a uma resposta de frequência subaproveitada do banco de filtro do codificador de transformação (devido à forma da janela otimizada de baixo atraso e/ou comprimento de transformação) e, assim, ainda compromete a qualidade perceptual destes codecs.

[0006] De acordo com o modelo psicoacústico clássico, pré-requisitos para transparência com relação ao ruído de quantização são definidos. Em altas taxas de bit, isso refere-se a uma distribuição de tempo/frequência ideal perceptualmente adaptada de ruído de quantização que obedece os níveis de mascaramento auditivo. Em baixas taxas de bit, entretanto, a transparência não pode ser obtida. Assim, uma estratégia de redução de exigências do nível de mascaramento pode ser empregada em baixas taxas de bit.

[0007] Assim, codecs de primeira linha foram

fornecidos para o conteúdo musical, em particular, codificadores de transformada com base na Transformada de Cosseno Discreta Modificada (MDCT | *Modified Discrete Cosine Transform*) que quantiza e transmite coeficientes espectrais no domínio de frequência. Entretanto, em taxas de dados muito baixas, apenas poucas linhas espectrais de cada período de tempo podem ser codificadas pelos bits disponíveis para este período. Como uma consequência, as perturbações de modulação temporal e as chamadas perturbações melódicas são inevitavelmente introduzidas no sinal codificado.

[0008] Mais visivelmente, estes tipos de perturbações são percebidos em componentes tonais quase estacionários. Isso acontece especialmente se, devido às restrições de atraso, uma forma da janela de transformação deve ser escolhida induzindo à significante diafonia entre os coeficientes espectrais adjacentes (ampliação espectral) devido ao efeito de vazamento bem conhecido. Entretanto, geralmente apenas um ou alguns destes coeficientes espectrais adjacentes permanecem zerados após a quantização bruta pelo codificador de baixa taxa de bit.

[0009] Conforme declarado acima, na técnica anterior, de acordo com uma abordagem, codificadores de transformada são empregados. Os codecs de áudio de alto índice de compressão que são bem adequados para codificar o conteúdo musical dependem da codificação de transformação. Os exemplos mais viáveis são Codificação de Áudio Avançada MPEG2/4 (AAC | *Advanced Audio Coding*) e Fala Unificada e Codificação de Áudio MPEG-D (USAC | *Unified Speech and Audio*

Coding). A USAC tem um núcleo comutado consistente de um módulo Previsão Linear Excitada por Código Algébrico (ACELP | *Algebraic Code Excited Linear Prediction*) mais um módulo de Excitação Codificada de Transformada (TCX | *Transform Coded Excitation*) (veja [5]) direcionado principalmente para codificação de fala e, de modo alternativo, AAC principalmente direcionado para codificação de música. Como AAC, TCX também é um método de codificação com base na transformação. Em configurações de baixa taxa de bit, estes esquemas de codificação estão sujeitos a exibir perturbações melódicas, especialmente se os esquemas de codificação subjacentes baseiam-se na Transformada de Cosseno Discreta Modificada (MDCT) (veja [1]).

[0010] Para a reprodução de música, os codificadores de transformada são a técnica preferida para compressão de dados de áudio. Entretanto, em baixas taxas de bit, os codificadores de transformada tradicionais exibem perturbações melódicas fortes e brutas. A maioria das perturbações surgem dos componentes espectrais tonais codificados muito escassamente. Isso acontece especialmente se estes são espectralmente espalhados por uma função de transferência espectral subaproveitada (efeito de vazamento) que é principalmente desenhada para cumprir as restrições de atraso rigorosos.

[0011] De acordo com outra abordagem na técnica anterior, os esquemas de codificação são completamente paramétricos para transientes, sinusoides e ruído. Em particular, para taxas de bit médias e baixas, codecs de

áudio completamente paramétricos foram padronizados, o mais viável dos quais são MPEG-4 Parte 3, Subparte 7 Linhas Harmônicas e Individuais mais Ruído (HILN | *Harmonic and Individual Lines plus Noise*) (veja [2]) e MPEG-4 Parte 3, Subparte 8 Codificação Sinusoidal (SSC | *SinuSoidal Coding*) (veja [3]). Os codificadores paramétricos, entretanto, apresentam um som desagradavelmente artificial e, com o aumento da taxa de bit, não escalam bem em direção à transparência perceptual.

[0012] Outra abordagem fornece forma de onda híbrida e codificação paramétrica. Em [4], um híbrido da codificação da forma de onda com base na transformação e MPEG 4-SSC (parte sinusoidal apenas) é proposto. Em um processo iterativo, sinusoides são extraídos e subtraídos do sinal para formar um sinal residual a ser codificada pelas técnicas de codificação de transformação. Os sinusoides extraídos são codificados por um conjunto de parâmetros e transmitidos junto com o residual. Em [6], uma abordagem da codificação híbrida é fornecida codificando os sinusoides e residuais separadamente. Em [7], no chamado codec Transformada Revestida de Energia Restrita (CELT | *Constrained Energy Lapped Transform*) /página da internet fantasma, a ideia de utilizar um banco de osciladores para codificação híbrida é apresentada. Entretanto, a geração de tons artificiais por um banco de osciladores que executa em paralelo com o decodificador e a saída da qual é misturada com a saída do banco de filtro de síntese do decodificador no domínio de tempo, significa uma grande carga do cálculo computacional,

visto que muitos osciladores devem ser computados em paralelo em uma alta taxa de amostra. A complexidade computacional é sempre um critério importante no desenvolvimento e instalação do codec, assim, formas mais eficientes para gerar estes tons são necessárias.

[0013] Nas taxas de bit médias ou mais altas, os codificadores de transformada são bem adequados para codificar a música devido ao seu som natural. Assim, as exigências de transparência do modelo psicoacústico subjacente são completamente ou quase completamente cumpridas. Entretanto, nas baixas taxas de bit, os codificadores devem violar seriamente as exigências do modelo psicoacústico e nesta situação os codificadores de transformada estão sujeitos a perturbações melódicas, brutas e de ruído musical.

[0014] Embora os codecs de áudio completamente paramétricos sejam os mais adequados para taxas de bit inferior, eles são, entretanto, conhecidos pelo som desagradavelmente artificial. Além disso, estes codecs não ascendem continuamente para a transparência perceptual, visto que um refinamento gradual do modelo paramétrico ainda bruto não é viável.

[0015] A forma de onda híbrida e codificação paramétrica poderiam potencialmente superar os limites das abordagens individuais e poderiam potencialmente se beneficiar das propriedades ortogonais mútuas de ambas as técnicas. Entretanto, no estado da técnica atual, é dificultada por uma falta de interação entre a parte da

codificação de transformação e a parte paramétrica do codec híbrido. Problemas se referem à divisão de sinal entre a parte do codec paramétrica e de transformação, condução do montante de bit entre a parte de transformação e paramétrica, técnicas de sinalização do parâmetro e fusão descontínua da saída do codec paramétrico e de transformação.

[0016] Outras publicações prévias no campo se referem à síntese de tons sinusoidais diretamente no domínio de tempo, ou tons constantes em peças no domínio de frequência de DFT [13], e na otimização de SNR dos padrões truncados no domínio de DFT [12]. A incorporação dos tons constantes de frequência em peças com base nos espectros de MDCT em um ambiente de codec perceptual [10] ou um cenário da extensão da largura de banda [11] já foi descrita. Entretanto, a eficiente geração de varreduras e sua ligação às faixas descontínuas no domínio de MDCT não foi supostamente direcionada ainda, nem tem a definição de restrições sensíveis nos graus disponíveis de liberdade no espaço do parâmetro.

[0017] O objetivo da presente invenção é fornecer os conceitos melhorados para a decodificação de áudio híbrida. O objetivo da presente invenção é solucionado por um aparelho, de acordo com a reivindicação 1, por um aparelho, de acordo com a reivindicação 14, por um método, de acordo com a reivindicação 20, por um método, de acordo com a reivindicação 21 e por um programa de computador, de acordo com a reivindicação 22.

[0018] Um aparelho para gerar um sinal de saída de

áudio com base em um espectro do sinal de áudio codificado é fornecido.

[0019] O aparelho compreende uma unidade de processamento para processar o espectro do sinal de áudio codificado para obter um espectro do sinal de áudio decodificado compreendendo uma pluralidade de coeficientes espectrais, caracterizado por cada um dos coeficientes espectrais terem uma localização espectral dentro do espectro do sinal de áudio codificado e um valor espectral, em que os coeficientes espectrais são sequencialmente organizados, de acordo com sua localização espectral dentro do espectro do sinal de áudio codificado de modo que os coeficientes espectrais formam uma sequência de coeficientes espectrais.

[0020] Além disso, o aparelho compreende um determinador dos pseudocoeficientes para determinar um ou mais pseudocoeficiente(s) do espectro do sinal de áudio decodificado, cada um dos pseudocoeficientes tendo uma localização espectral e um valor espectral.

[0021] Além disso, o aparelho compreende uma unidade de substituição para substituir pelo menos um ou mais pseudocoeficiente(s) por um padrão espectral determinado para obter um espectro do sinal de áudio modificado, caracterizado pelo padrão espectral determinado compreender pelo menos dois coeficientes padrões, em que cada um de pelo menos dois coeficientes padrões tem um valor espectral.

[0022] Além disso, o aparelho compreende uma unidade de conversão de tempo do espectro para converter o espectro do sinal de áudio modificado em um domínio de tempo para

obter o sinal de saída de áudio.

[0023] Em uma aplicação, o aparelho ainda pode compreender uma unidade de armazenamento compreendendo uma base de dados ou uma memória tendo armazenado dentro da base de dados ou dentro da memória uma pluralidade de padrões espectrais armazenados, caracterizado por cada um dos padrões espectrais armazenados terem uma determinada propriedade espectral (por exemplo, frequência constante, frequência de varredura - cada um em uma versão da localização sobre a posição ou entre posições - etc.). A unidade de substituição pode ser configurada para solicitar um dos padrões espectrais armazenados como um padrão espectral solicitado a partir da unidade de armazenamento. A unidade de armazenamento pode ser configurada para fornecer o referido padrão espectral solicitado, e a unidade de substituição pode ser configurada para substituir pelo menos um ou mais pseudocoeficiente(s) pelo padrão espectral determinado com base no padrão espectral solicitado.

[0024] De acordo com uma aplicação, a unidade de substituição pode ser configurada para solicitar um dos referidos padrões espectrais armazenados a partir da unidade de armazenamento dependendo de uma primeira localização espectral derivada de pelo menos um de um ou mais pseudocoeficiente(s) determinado(s) pelo determinador dos pseudocoeficientes.

[0025] Em uma aplicação, a primeira localização espectral derivada de pelo menos um de um ou mais pseudocoeficiente(s) pode ser a localização espectral de um

dos pseudocoeficientes.

[0026] Em outra aplicação, um ou mais pseudocoeficiente(s) são valores com sinal, cada um compreendendo um componente do sinal, e a unidade de substituição é configurada para determinar a primeira localização espectral derivada com base na localização espectral de um pseudocoeficiente de um ou mais pseudocoeficiente(s) e com base no componente do sinal do referido pseudocoeficiente, de modo que a primeira localização espectral derivada seja igual à localização espectral do referido pseudocoeficiente quando o componente do sinal tem um primeiro valor do sinal, e de modo que a primeira localização espectral derivada seja igual a uma localização modificada, a localização modificada resultante da mudança da localização espectral do referido pseudocoeficiente por um valor predefinido quando o componente do sinal tem um segundo valor diferente.

[0027] Por exemplo, uma resolução de frequência de meia posição das pseudolinhas pode ser sinalizada pelo sinal do referido pseudocoeficiente. O valor predefinido pelo qual a localização espectral do referido pseudocoeficiente é mudada pode então corresponder à metade da diferença de frequência, por exemplo, de duas posições subsequentes, por exemplo, quando um domínio de frequência de tempo é considerado, quando o componente do sinal do pseudocoeficiente tem o segundo valor do sinal.

[0028] O componente do sinal do pseudocoeficiente pode ser compreendido pelo valor espectral do

pseudocoeficiente.

[0029] Em uma aplicação, a pluralidade de padrões espectrais armazenados sendo armazenados dentro da base de dados ou da memória da unidade de armazenamento pode ser tanto padrões de tom estacionário quanto padrões de varredura de frequência. O determinador dos pseudocoeficientes pode ser configurado para determinar dois ou mais pseudocoeficientes temporalmente consecutivos do espectro do sinal de áudio decodificado. A unidade de substituição pode ser configurada para atribuir um primeiro pseudocoeficiente e um segundo pseudocoeficiente de dois ou mais pseudocoeficientes temporalmente consecutivos em uma faixa dependendo se uma diferença absoluta entre a primeira localização espectral derivada do primeiro pseudocoeficiente e uma segunda localização espectral derivada do segundo pseudocoeficiente é menor do que um valor limite. E, a unidade de substituição pode ser configurada para solicitar um dos padrões de tom estacionário a partir da unidade de armazenamento quando a primeira localização espectral derivada do primeiro pseudocoeficiente da faixa é igual à segunda localização espectral derivada do segundo pseudocoeficiente da faixa. Além disso, a unidade de substituição pode ser configurada para solicitar um dos padrões de varredura de frequência a partir da unidade de armazenamento quando a primeira localização espectral derivada do primeiro pseudocoeficiente da faixa é diferente da segunda localização espectral derivada do segundo pseudocoeficiente da faixa.

[0030] De acordo com uma aplicação, a unidade de

substituição pode ser configurada para solicitar um primeiro padrão de varredura de frequência dos padrões de varredura de frequência a partir da unidade de armazenamento quando uma diferença de frequência entre a segunda localização espectral derivada do segundo pseudocoeficiente da faixa e a primeira localização espectral derivada do primeiro pseudocoeficiente da faixa é igual à metade de um valor predefinido. Além disso, a unidade de substituição pode ser configurada para solicitar um segundo padrão de varredura de frequência, sendo diferente do primeiro padrão de varredura de frequência, dos padrões de varredura de frequência a partir da unidade de armazenamento quando a diferença de frequência entre a segunda localização espectral derivada do segundo pseudocoeficiente da faixa e a primeira localização espectral derivada do primeiro pseudocoeficiente da faixa é igual ao valor predefinido. Além disso, a unidade de substituição pode ser configurada para solicitar um terceiro padrão de varredura de frequência, sendo diferente do primeiro padrão de varredura e do segundo padrão de varredura de frequência, dos padrões de varredura de frequência a partir da unidade de armazenamento quando a diferença de frequência entre a segunda localização espectral derivada do segundo pseudocoeficiente da faixa e a primeira localização espectral derivada do primeiro pseudocoeficiente da faixa é igual a uma vez e meia o valor predefinido.

[0031] De acordo com uma aplicação, a unidade de substituição compreende uma unidade de adaptação padrão sendo configurada para modificar o padrão espectral solicitado

fornecido pela unidade de armazenamento para obter o padrão espectral determinado.

[0032] Em uma aplicação, a unidade de adaptação padrão pode ser configurada para modificar o padrão espectral solicitado fornecido pela unidade de armazenamento pelo redimensionamento dos valores espectrais dos coeficientes padrões do padrão espectral solicitado dependendo do valor espectral de um de um ou mais pseudocoeficiente(s) para obter o padrão espectral determinado.

[0033] De acordo com uma aplicação, a unidade de adaptação padrão pode ser configurada para modificar o padrão espectral solicitado fornecido pela unidade de armazenamento dependendo de uma fase inicial de modo que o valor espectral de cada um dos coeficientes padrões do padrão espectral solicitado seja modificado em uma primeira forma, quando a fase inicial tem um primeiro valor da fase inicial, e de modo que o valor espectral de cada um dos coeficientes padrões do padrão espectral solicitado seja modificado em uma segunda forma diferente, quando a fase inicial tem um segundo valor da fase inicial diferente.

[0034] De acordo com uma aplicação, o valor espectral de cada um dos coeficientes padrões do padrão espectral solicitado pode ser um coeficiente complexo compreendendo uma parte real e uma parte imaginária. Em tal aplicação, a unidade de adaptação padrão pode ser configurada para modificar o padrão espectral solicitado pela modificação da parte real e da parte imaginária de cada um dos coeficientes padrões do padrão espectral solicitado fornecido

pela unidade de armazenamento, pela aplicação de um fator de rotação complexo $e^{j\cdot\varphi}$, em que φ é um ângulo (por exemplo, valor do ângulo). Por isso, para cada um dos coeficientes complexos um vetor que representa o referido coeficiente complexo em um plano complexo é girado pelo mesmo ângulo para cada um dos coeficientes complexos.

[0035] Em uma aplicação, o valor espectral de cada um dos coeficientes padrões do padrão espectral solicitado compreende uma parte real e uma parte imaginária. A unidade de adaptação padrão pode ser configurada para modificar o padrão espectral solicitado fornecido pela unidade de armazenamento pela negação da parte real e da parte imaginária do valor espectral de cada um dos coeficientes padrões do padrão espectral solicitado, ou pela troca da parte real ou uma parte real negada e a parte imaginária ou uma parte imaginária negada do valor espectral de cada um dos coeficientes padrões do padrão espectral solicitado.

[0036] Em uma aplicação, a unidade de adaptação padrão pode ser configurada para modificar o padrão espectral solicitado fornecido pela unidade de armazenamento pela realização de um paralelismo temporal do padrão. Tipicamente, isso pode ser obtido em um domínio de frequência pelo cálculo computacional do conjugado complexo (pela multiplicação da parte imaginária por -1) do padrão e pela aplicação de um termo da fase complexa (cruzado).

[0037] De acordo com uma aplicação, o espectro do sinal de áudio decodificado é representado em um domínio de MDCT. A unidade de adaptação padrão pode ser configurada para

modificar o padrão espectral solicitado fornecido pela unidade de armazenamento pela modificação dos valores espectrais dos coeficientes padrões do padrão espectral solicitado para obter um padrão espectral modificado, caracterizado pelos valores espectrais serem representados em um domínio da Transformada Discreta de Fourier com Empilhamento Diferencial. Além disso, a unidade de adaptação padrão pode ser configurada para transformar os valores espectrais dos coeficientes padrões do padrão espectral modificado do domínio da Transformada Discreta de Fourier com Empilhamento Diferencial ao domínio de MDCT para obter o padrão espectral determinado. Além disso, a unidade de substituição pode ser configurada para substituir pelo menos um ou mais pseudocoeficiente(s) pelo padrão espectral determinado sendo representado no domínio de MDCT para obter o espectro do sinal de áudio modificado sendo representado no domínio de MDCT.

[0038] De modo alternativo, nas aplicações os valores espectrais podem ser representados em um domínio da Transformada de Cosseno Discreta Modificada Complexa (CMDCT | *Complex Modified Discrete Cosine Transform*). Além disso, nestas aplicações a unidade de adaptação padrão pode ser configurada para transformar os valores espectrais dos coeficientes padrões do padrão espectral modificado do domínio de CMDCT ao domínio de MDCT para obter o padrão espectral determinado por simplesmente extrair a parte real do padrão modificado complexo.

[0039] Além disso, um aparelho para gerar uma

pluralidade de padrões espectrais é fornecido. O aparelho compreende um gerador de sinal para gerar uma pluralidade de sinais em um primeiro domínio. Além disso, o aparelho compreende uma unidade de transformação do sinal para transformar cada sinal da pluralidade de sinais do primeiro domínio em um segundo domínio para obter uma pluralidade de padrões espectrais, cada padrão da pluralidade de padrões espectrais transformados compreendendo uma pluralidade de coeficientes. Além disso, o aparelho compreende uma unidade de pós-processamento para truncar os padrões espectrais transformados pela remoção de um ou mais dos coeficientes dos padrões espectrais transformados para obter uma pluralidade de padrões processados. Além disso, o aparelho compreende uma unidade de armazenamento compreendendo uma base de dados ou uma memória, em que a unidade de armazenamento é configurada para armazenar cada padrão processado da pluralidade de padrões processados na base de dados ou na memória. O gerador de sinal é configurado para gerar cada sinal da pluralidade de sinais com base nas fórmulas

$$x(t) = cos\,(2\pi\varphi(t))$$

e

$$\varphi(t) = \varphi(0) + \int_0^t 2\pi f(\tau)d\tau \quad ,$$

[0040] em que t e τ indicam tempo, em que φ(t) é uma fase instantânea em t e em que f(τ) é uma frequência instantânea em τ, em que cada sinal da pluralidade de sinais tem uma frequência inicial (f_0), sendo uma frequência instantânea do referido sinal em um primeiro ponto no tempo e

uma frequência alvo (f_1), sendo uma frequência instantânea do referido sinal em um diferente segundo ponto no tempo. O gerador de sinal é configurado para gerar um primeiro sinal da pluralidade de sinais de modo que a frequência alvo do primeiro sinal é igual à frequência inicial. Além disso, o gerador de sinal é configurado para gerar um segundo sinal diferente da pluralidade de sinais de modo que a frequência alvo do primeiro sinal é diferente da frequência inicial.

[0041] De acordo com uma aplicação, a unidade de transformação do sinal pode ser configurada para transformar cada sinal da pluralidade de sinais do primeiro domínio, sendo um domínio de tempo, em um segundo domínio, sendo um domínio espectral. A unidade de transformação do sinal pode ser configurada para gerar um primeiro de uma pluralidade de blocos de tempo para transformar referido sinal, caracterizado por cada bloco de tempo da pluralidade de blocos de tempo compreender uma pluralidade de amostras ponderadas, em que cada uma das referidas amostras ponderadas é uma amostra do sinal do referido sinal sendo ponderada por uma ponderação de uma pluralidade de ponderações, em que a pluralidade de ponderações é atribuída ao referido bloco de tempo, e em que cada ponderação da pluralidade de ponderações é atribuída a um ponto no tempo. A frequência inicial (f_0) de cada sinal da pluralidade de sinais pode ser uma frequência instantânea do referido sinal no primeiro ponto no tempo, onde uma primeira das ponderações do primeiro de um dos blocos de tempo é atribuída ao primeiro ponto no tempo, onde uma segunda das ponderações de um diferente segundo dos

blocos de tempo é atribuída ao primeiro ponto no tempo, em que o primeiro dos blocos de tempo e o segundo dos blocos de tempo se sobrepõem, e em que a primeira das ponderações é igual à segunda das ponderações. A frequência alvo (f_1) de cada sinal da pluralidade de sinais pode ser uma frequência instantânea do referido sinal no segundo ponto no tempo, onde uma terceira das ponderações do primeiro de um dos blocos de tempo é atribuída ao segundo ponto no tempo, onde uma quarta das ponderações de um diferente terceiro dos blocos de tempo é atribuída ao segundo ponto no tempo, em que o primeiro dos blocos de tempo e o terceiro dos blocos de tempo se sobrepõem, e em que a terceira das ponderações é igual à quarta das ponderações.

[0042] Deve ser observado que isso, por exemplo, pode ser suficiente para gerar apenas um bloco de tempo (por exemplo, o primeiro dos blocos de tempo) para a geração de um padrão.

[0043] De acordo com uma aplicação, cada sinal da pluralidade de sinais tem uma fase inicial (φ_0), sendo uma fase do referido sinal em um primeiro ponto no tempo, e uma fase alvo (φ_1), sendo uma fase do referido sinal em um diferente segundo ponto no tempo, caracterizado pelo gerador de sinal ser configurado para gerar a pluralidade de sinais de modo que a fase inicial (φ_0) de um primeiro da pluralidade de sinais é igual à fase inicial (φ_0) de um diferente segundo sinal da pluralidade de sinais.

[0044] A fase inicial (e, implicitamente pela escolha da frequência inicial e alvo, a fase de parada) de

cada sinal da pluralidade de sinais pode ser ajustada nos referidos pontos no tempo inicial e de parada.

[0045] Por esta escolha especial de pontos no tempo inicial e de parada, perturbações adicionadas por sobreposição são reduzidas podendo ocorrer, se padrões com diferentes propriedades espectrais forem interligados.

[0046] Em uma aplicação, a unidade de pós-processamento pode ser ainda configurada para conduzir uma rotação por $\pi/4$ nos coeficientes espectrais de cada um dos padrões espectrais transformados para obter uma pluralidade de padrões espectrais girados.

[0047] Em outra aplicação, a unidade de pós-processamento pode ser ainda configurada para conduzir uma rotação por um ângulo de fase arbitrária nos coeficientes espectrais de cada um dos padrões espectrais transformados para obter uma pluralidade de padrões espectrais arbitrariamente girados.

[0048] De acordo com uma aplicação adicional, o gerador de sinal pode ser configurado para gerar o primeiro sinal, o segundo sinal e um ou mais sinal(is) adicional(is) como a pluralidade de sinais, de modo que cada diferença da frequência alvo e da frequência inicial de cada um dos sinais adicionais seja um múltiplo inteiro de uma diferença da frequência alvo e da frequência inicial do segundo sinal.

[0049] Além disso, um método para gerar um sinal de saída de áudio com base em um espectro do sinal de áudio codificado é fornecido. O método compreende:

- Processar o espectro do sinal de áudio

codificado para obter um espectro do sinal de áudio decodificado compreendendo uma pluralidade de coeficientes espectrais, caracterizado por cada um dos coeficientes espectrais terem uma localização espectral dentro do espectro do sinal de áudio codificado e um valor espectral, em que os coeficientes espectrais são sequencialmente organizados, de acordo com sua localização espectral dentro do espectro do sinal de áudio codificado de modo que os coeficientes espectrais formam uma sequência de coeficientes espectrais.

- Determinar um ou mais pseudocoeficiente(s) do espectro do sinal de áudio decodificado, em que cada um dos pseudocoeficientes é um dos coeficientes espectrais,

- Substituir pelo menos um ou mais pseudocoeficiente(s) por um padrão espectral determinado para obter um espectro do sinal de áudio modificado, em que o padrão espectral determinado compreende pelo menos dois coeficientes padrões, em que cada um de pelo menos dois coeficientes padrões tem um valor espectral. E:

- Converter o espectro do sinal de áudio modificado em um domínio de tempo para obter o sinal de saída de áudio.

[0050] Além disso, um método para gerar uma pluralidade de padrões espectrais é fornecido. O método compreende:

- Gerar uma pluralidade de sinais em um primeiro domínio.

- Transformar cada sinal da pluralidade de sinais do primeiro domínio em um segundo domínio para obter

uma pluralidade de padrões espectrais, cada padrão da pluralidade de padrões espectrais transformados compreendendo uma pluralidade de coeficientes.

- Truncar os padrões espectrais transformados pela remoção de um ou mais dos coeficientes dos padrões espectrais transformados para obter uma pluralidade de padrões processados. E:

- Armazenar cada padrão processado da pluralidade de padrões processados em uma base de dados ou uma memória.

[0051] A geração de cada sinal da pluralidade de sinais é conduzida com base nas fórmulas

$$x(t) = cos\,(2\pi\varphi(t))$$

e

$$\varphi(t) = \varphi(0) + \int_0^t 2\pi f(\tau)d\tau \quad ,$$

[0052] caracterizado por t e τ indicam tempo, em que φ(t) é uma fase instantânea em t e em que f(τ) é uma frequência instantânea em τ e em que cada sinal da pluralidade de sinais tem uma frequência inicial (f_0), sendo uma frequência instantânea do referido sinal em um primeiro ponto no tempo e uma frequência alvo (f_1), sendo uma frequência instantânea do referido sinal em um diferente segundo ponto no tempo.

[0053] A geração da pluralidade de sinais é conduzida gerando um primeiro sinal da pluralidade de sinais de modo que a frequência alvo (f_1) do primeiro sinal é igual à frequência inicial (f_0). Além disso, a geração da

pluralidade de sinais é conduzida gerando um segundo sinal diferente da pluralidade de sinais de modo que a frequência alvo (f_1) do primeiro sinal seja diferente da frequência inicial (f_0).

[0054] Além disso, um programa de computador para implementar os métodos descritos acima quando executados em um computador ou processador de sinal é fornecido.

[0055] Visto que codecs modernos como AAC ou USAC estão com base em uma representação de áudio de domínio de MDCT, as aplicações fornecem conceitos para gerar tons sintéticos pela correção dos padrões do tom no espectro de MDCT no decodificador. É demonstrado como padrões espectrais apropriados podem ser derivados e adaptados para sua localização alvo na (e entre a) grade de tempo/frequência (t/f) de MDCT para continuamente sintetizar tons sinusoidais de alta qualidade incluindo varreduras.

[0056] Codecs modernos como Codificação de Áudio Avançada (AAC) ou Fala Unificada e Codificação de Áudio (USAC) com base em uma representação de áudio do domínio da Transformada de Cosseno Discreta Modificada (MDCT). As aplicações geram tons sintéticos pela direta correção dos padrões do tom no espectro de MDCT no decodificador. Apenas por isso, uma implementação de complexidade ultra baixa pode ser realizada.

[0057] Nas aplicações, os padrões apropriados são derivados e são adaptados para sua localização alvo na (e entre a) grade t/f de MDCT para sintetizar os tons sinusoidais de alta qualidade incluindo varreduras.

[0058] De acordo com as aplicações, a codificação de áudio de baixo atraso e baixa taxa de bit é fornecida. Algumas aplicações com base em um conceito novo e inovador referido como ToneFilling (TF). O termo ToneFilling denota uma técnica de codificação, na qual, tons naturais indevidamente codificados de outra forma são substituídos perceptualmente por tons do seno puros ainda semelhantes. Assim, as perturbações da modulação de amplitude em uma determinada taxa, dependentes da posição espectral do sinusoide com relação à localização espectral da posição mais próxima da MDCT, são evitadas (conhecidas como "melódicas").

[0059] Nas aplicações, um grau de incômodo de todas as perturbações concebíveis é ponderado. Isso se refere aos aspectos perceptuais como, por exemplo, altura, harmonização, modulação e para estacionários de perturbações. Todos os aspectos são avaliados em um Modelo de Incômodo de Percepção Sonora (SPAM | *Sound Perception Annoyance Model*). Conduzido por tal modelo, ToneFilling fornece vantagens significantes. Um erro de altura e modulação que é introduzido pela substituição de um tom natural com um tom senoidal puro, é ponderado versus um impacto de ruído aditivo e baixa estacionariedade ("melódica") causada por um tom natural insuficientemente quantizado.

[0060] ToneFilling fornece diferenças significantes em codecs sinusoides-mais-ruído. Por exemplo, TF substitui tons por sinusoides e varreduras sinusoidais lineares com declives predefinidos, ao invés de uma subtração de sinusoides. Tons perceptualmente semelhantes têm os mesmos

Centros de Gravidades (COG | *Centers of Gravity*) locais que o componente do som original a ser substituído. De acordo com as aplicações, tons originais são apagados no espectro de áudio (base da esquerda para a direita da função do COG). Tipicamente, a resolução de frequência do sinusoide utilizada para substituição é a mais bruta possível para reduzir a informação adicional, enquanto, ao mesmo tempo, responde pelas exigências perceptuais para evitar uma sensação fora de sintonização.

[0061] Em algumas aplicações, ToneFilling pode ser conduzido acima de uma frequência de corte inferior devido às referidas exigências perceptuais, mas não abaixo da frequência de corte inferior. Ao conduzir o ToneFilling, os tons são representados através das pseudolinhas espectrais dentro de um codificador de transformada. Entretanto, em um codificador equipado com ToneFilling, as pseudolinhas são submetidas ao processamento regular controlado pelo modelo psicoacústico clássico. Assim, ao conduzir o ToneFilling, não há necessidade de restrições anteriores da parte paramétrica (na taxa de bit os x, y os componentes tonais são substituídos). Esta integração rigorosa em um codec de transformação é obtida.

[0062] A funcionalidade de ToneFilling pode ser empregada no codificador, detectando os COGs locais (estimativas uniformes; medições da qualidade máxima), pela remoção de componentes tonais, gerando pseudolinhas substituídas (por exemplo, pseudocoeficientes) que carregam uma informação de nível através da amplitude das

pseudolinhas, uma informação de frequência através da posição espectral das pseudolinhas e uma informação de frequência precisa (meio deslocamento da posição) através do sinal das pseudolinhas. Pseudocoeficientes (pseudolinhas) são gerenciados por uma unidade do quantizador subsequente do codec como qualquer coeficiente espectral regular (linha espectral).

[0063] ToneFilling pode ainda ser empregado no decodificador pela detecção das linhas espectrais isoladas, caracterizado pelos verdadeiros pseudocoeficientes (pseudolinhas) poderem ser marcados pela matriz indicadora (por exemplo, um campo de bit). O decodificador pode ligar a informação da pseudolinha para criar faixas sinusoidais. Um esquema de nascimento/continuação/morte pode ser empregado para sintetizar as faixas contínuas.

[0064] Para decodificar, pseudocoeficientes (pseudolinhas) podem ser marcados por uma matriz indicadora transmitida dentro da informação adicional. Uma resolução de frequência de meia posição das pseudolinhas pode ser sinalizada pelo sinal dos pseudocoeficientes (pseudolinhas). No decodificador, as pseudolinhas podem ser apagadas do espectro antes da unidade de transformação inversa e sintetizadas separadamente por um banco de osciladores. Ao longo do tempo, pares de osciladores podem ser ligados e a interpolação do parâmetro é empregada para garantir uma saída do oscilador levemente evoluída.

[0065] As compensações/descompensações dos osciladores acionados pelo parâmetro podem ser formadas de

modo que eles correspondem proximamente às características temporais da operação de janelamento do codec de transformação, assim, garantindo a transição sem descontinuidades entre as partes do codec de transformação geradas e as partes do oscilador geradas do sinal de saída.

[0066] Os conceitos fornecidos integram bem e sem esforços nos esquemas de codificação de transformação existentes como AAC, TCX ou configurações semelhantes. A direção da precisão de quantização do parâmetro pode ser implicitamente realizada pelo controle da taxa existente do codec.

[0067] Em algumas aplicações, as pseudolinhas (pseudocoeficientes) podem ser gerenciadas pelo quantizador existente no codec apenas como qualquer linha espectral regular; como oposto para separar a sinalização dos parâmetros sinusoidais.

[0068] Em algumas aplicações, uma fase inicial opcionalmente medida de uma faixa sinusoidal obtida da extrapolação de espectros precedentes pode ser empregada.

[0069] De acordo com algumas aplicações, uma técnica opcional de Cancelamento de Pseudônimo de Domínio de Tempo (TDAC | *Time Domain Alias Cancellation*) pode ser empregada pela modelagem do pseudônimo na compensação/descompensação de uma faixa sinusoidal.

[0070] A seguir, aplicações da presente invenção são descritas em mais detalhes com referência às figuras, nas quais:

[0071] A figura 1a ilustra um aparelho para gerar um

sinal de saída de áudio com base em um espectro do sinal de áudio codificado de acordo com uma aplicação,

[0072] A figura 1b ilustra um aparelho para gerar um sinal de saída de áudio com base em um espectro do sinal de áudio codificado de acordo com outra aplicação,

[0073] A figura 1c ilustra um aparelho para gerar um sinal de saída de áudio com base em um espectro do sinal de áudio codificado de acordo com uma aplicação adicional,

[0074] A figura 1d ilustra um aparelho para gerar uma pluralidade de padrões espectrais de acordo com uma aplicação,

[0075] A figura 2 descreve o alinhamento do parâmetro de um padrão de varredura com relação a um bloco de tempo de MDCT,

[0076] A figura 3 mostra o processo de correção de um padrão do tom, em que (a-b) ilustram geração do padrão prototípico, caracterizado por (c) ilustrar a truncagem padrão, em que (d) ilustra a adaptação padrão na localização alvo e fase e em que (e-f) ilustram a correção padrão,

[0077] A figura 4 ilustra padrões do tom espectral normalizados: seno sobre a posição, seno entre a posição, varredura sobre a posição, varredura entre a posição (do painel superior ao inferior),

[0078] A figura 5 descreve um índice do sinal para ruído (SNR) do padrão do tom truncado como uma função do comprimento padrão para uma janela sinusoidal,

[0079] A figura 6a mostra uma frequência instantânea de uma varredura sinusoidal em pontos no tempo para

sobreposição dos blocos de acordo com as aplicações,

[0080] A figura 6b descreve um progresso da fase para funções base de DCT e DCT IV de acordo com as aplicações,

[0081] A figura 6c ilustra um espectro de potência, um espectro de MDCT substituído, um espectro de MDCT quantizado e um espectro de MDCT com padrões, de acordo com uma aplicação,

[0082] A figura 7 ilustra um aparelho para codificar um espectro de entrada do sinal de áudio de acordo com uma aplicação,

[0083] A figura 8 descreve um espectro de entrada do sinal de áudio, um espectro de potência correspondente e um espectro do sinal de áudio modificado (substituído),

[0084] A figura 9 ilustra outro espectro de potência, outro espectro do sinal de áudio modificado (substituído) e um espectro do sinal de áudio quantizado, em que o espectro do sinal de áudio quantizado gerado em um lado do codificador, pode, em algumas aplicações, corresponder ao espectro do sinal de áudio decodificado em um lado de decodificação,

[0085] A figura 10 ilustra um aparelho para gerar um sinal de saída de áudio com base em um espectro do sinal de áudio codificado de acordo com uma aplicação,

[0086] A figura 11 descreve um aparelho para gerar um sinal de saída de áudio com base em um espectro do sinal de áudio codificado, de acordo com outra aplicação, e

[0087] A figura 12 mostra dois diagramas comparando

sinusoides originais e sinusoides após serem processados por uma cadeia de MDCT / de MDCT inversa.

[0088] A figura 7 ilustra um aparelho para codificar um espectro de entrada do sinal de áudio de acordo com uma aplicação. O aparelho para codificar compreende um determinador de extremos 410, um modificador de espectro 420, uma unidade de processamento 430 e um gerador de informação adicional 440.

[0089] Antes de considerar o aparelho da figura 7 em mais detalhes, o espectro de entrada do sinal de áudio que é codificado pelo aparelho da figura 7 é considerado em mais detalhes.

[0090] A princípio, qualquer tipo de espectro de sinal de áudio pode ser codificado pelo aparelho da figura 7. O espectro de entrada do sinal de áudio pode, por exemplo, ser um espectro de MDCT (Transformada de Cosseno Discreta Modificada), um espectro de magnitude de DFT (Transformada Discreta de Fourier) ou um espectro de MDST (Transformada de Seno Discreta Modificada).

[0091] A figura 8 ilustra um exemplo de um espectro de entrada do sinal de áudio 510. Na figura 8, o espectro de entrada do sinal de áudio 510 é um espectro de MDCT.

[0092] O espectro de entrada do sinal de áudio compreende uma pluralidade de coeficientes espectrais. Cada um dos coeficientes espectrais tem uma localização espectral dentro do espectro de entrada do sinal de áudio e um valor espectral.

[0093] Considerando o exemplo da figura 8, onde o

espectro de entrada do sinal de áudio resulta de uma transformação de MDCT do sinal de áudio, por exemplo, um banco de filtro que transformou o sinal de áudio para obter o espectro de entrada do sinal de áudio, pode, por exemplo, utilizar 1024 canais. Então, cada um dos coeficientes espectrais está associado com um dos 1024 canais e o número do canal (por exemplo, um número entre 0 e 1023) pode ser considerado como a localização espectral dos referidos coeficientes espectrais. Na figura 8, a abscissa 511 se refere à localização espectral dos coeficientes espectrais. Para melhor ilustrar, apenas os coeficientes com localizações espectrais entre 52 e 148 são ilustrados pela figura 8.

[0094] Na figura 8, a ordenada 512 ajuda a determinar o valor espectral dos coeficientes espectrais. No exemplo da figura 8 que descreve um espectro de MDCT, os valores espectrais dos coeficientes espectrais do espectro de entrada do sinal de áudio, a abscissa 512 se refere aos valores espectrais dos coeficientes espectrais. Deve ser observado que os coeficientes espectrais de um espectro de entrada do sinal de áudio de MDCT podem ter números reais positivos bem como negativos como valores espectrais.

[0095] Outros espectros de entrada do sinal de áudio, entretanto, podem ter apenas coeficientes espectrais com valores espectrais que são positivos ou zero. Por exemplo, o espectro de entrada do sinal de áudio pode ser um espectro de magnitude de DFT, com coeficientes espectrais tendo valores espectrais que representam as magnitudes dos coeficientes resultantes da Transformada Discreta de Fourier.

Estes valores espectrais podem ser apenas positivos ou zero.

[0096] Em outras aplicações, o espectro de entrada do sinal de áudio compreende coeficientes espectrais com valores espectrais que são números complexos. Por exemplo, um espectro de DFT indicando a informação de magnitude e fase pode compreender coeficientes espectrais tendo valores espectrais que são números complexos.

[0097] Conforme mostrado de forma exemplar na figura 8, os coeficientes espectrais são sequencialmente organizados, de acordo com sua localização espectral dentro do espectro de entrada do sinal de áudio de modo que os coeficientes espectrais formam uma sequência de coeficientes espectrais. Cada um dos coeficientes espectrais tem pelo menos um de um ou mais precursores e um ou mais sucessores, em que cada precursor do referido coeficiente espectral é um dos coeficientes espectrais que precede o referido coeficiente espectral dentro da sequência. Cada sucessor do referido coeficiente espectral é um dos coeficientes espectrais que sucede o referido coeficiente espectral dentro da sequência. Por exemplo, na figura 8, um coeficiente espectral tendo a localização espectral 81, 82 ou 83 (e assim por diante) é um sucessor para o coeficiente espectral com a localização espectral 80. Um coeficiente espectral tendo a localização espectral 79, 78 ou 77 (e assim por diante) é um precursor para o coeficiente espectral com a localização espectral 80. Para o exemplo de um espectro de MDCT, a localização espectral de um coeficiente espectral pode ser o canal da transformação de MDCT, o coeficiente espectral se

refere a (por exemplo, um número do canal entre, por exemplo, 0 e 1023). Novamente, deve ser observado que, para fins ilustrativos, o espectro de MDCT 510 da figura 8 apenas ilustra coeficientes espectrais com localizações espectrais entre 52 e 148.

[0098] Com referência à figura 7, o determinador de extremos 410 é agora descrito em mais detalhes. O determinador de extremos 410 é configurado para determinar um ou mais coeficientes extremos.

[0099] No geral, o determinador de extremos 410 avalia os espectros de entrada do sinal de áudio ou um espectro que está relacionado ao espectro de entrada do sinal de áudio para coeficientes extremos. A finalidade de determinação de coeficientes extremos é que, posteriormente, uma ou mais regiões tonais locais devem ser substituídas no espectro de sinal de áudio pelos pseudocoeficientes, por exemplo, por um único pseudocoeficiente para cada região tonal.

[0100] Geralmente, áreas mais elevadas em um espectro de potência do sinal de áudio, cujo espectro de entrada do sinal de áudio se refere, indicam regiões tonais. Então, pode ser preferido identificar áreas mais elevadas em um espectro de potência do sinal de áudio ao qual o espectro de entrada do sinal de áudio se refere. O determinador de extremos 410 pode, por exemplo, avaliar um espectro de potência, compreendendo coeficientes, que pode ser referido como coeficientes de comparação (pois seus valores espectrais são em pares comparados pelo determinador de extremos), de

modo que cada um dos coeficientes espectrais do espectro de entrada do sinal de áudio tenha um valor de comparação associado a ele.

[0101] Na figura 8, um espectro de potência 520 é ilustrado. O espectro de potência 520 e o espectro de entrada do sinal de áudio de MDCT 510 referem-se ao mesmo sinal de áudio. O espectro de potência 520 compreende coeficientes referidos como coeficientes de comparação. Cada coeficiente espectral compreende uma localização espectral referente à abscissa 521 e um valor de comparação. Cada coeficiente espectral do espectro de entrada do sinal de áudio tem um coeficiente de comparação associado a ele e, assim, ainda tem o valor de comparação de seu coeficiente de comparação associado a ele. Por exemplo, o valor de comparação associado a um valor espectral do espectro de entrada do sinal de áudio pode ser o valor de comparação do coeficiente de comparação com a mesma posição espectral que o coeficiente espectral do espectro de entrada do sinal de áudio considerou. A associação entre três dos coeficientes espectrais do espectro de entrada do sinal de áudio 510 e três dos coeficientes de comparação (e, assim, a associação com os valores de comparação destes coeficientes de comparação) do espectro de potência 520 é indicada pelas linhas tracejadas 513, 514, 515 indicando uma associação dos respectivos coeficientes de comparação (ou seus valores de comparação) e os respectivos coeficientes espectrais do espectro de entrada do sinal de áudio 510.

[0102] O determinador de extremos 410 pode ser

configurado para determinar um ou mais coeficiente(s) extremo(s), de modo que cada um dos coeficientes extremos seja um dos coeficientes espectrais cujo valor de comparação é superior ao valor de comparação de um de seus precursores e cujo valor de comparação é superior ao valor de comparação de um de seus sucessores.

[0103] Por exemplo, o determinador de extremos 410 pode determinar os valores máximos locais do espectro de potência. Em outras palavras, o determinador de extremos 410 pode ser configurado para determinar um ou mais coeficiente(s) extremo(s), de modo que cada um dos coeficientes extremos seja um dos coeficientes espectrais cujo valor de comparação é superior ao valor de comparação de seu precursor imediato e cujo valor de comparação é superior ao valor de comparação de seu sucessor imediato. Aqui, o precursor imediato de um coeficiente espectral é um dos coeficientes espectrais que imediatamente precede o referido coeficiente espectral no espectro de potência. O sucessor imediato do referido coeficiente espectral é um dos coeficientes espectrais que imediatamente sucede o referido coeficiente espectral no espectro de potência.

[0104] Entretanto, outras aplicações não exigem que o determinador de extremos 410 determine todos os máximos locais. Por exemplo, em algumas aplicações, o determinador de extremos pode apenas avaliar as partes determinadas do espectro de potência, por exemplo, referente a uma faixa de frequência determinada, apenas.

[0105] Em outras aplicações, o determinador de

extremos 410 é configurado para apenas estes coeficientes como coeficientes extremos, onde uma diferença entre o valor de comparação do máximo local considerado e o valor de comparação do mínimo local subsequente e/ou mínimo local precedente é superior a um valor limite.

[0106] O determinador de extremos 410 pode determinar o extremo ou os extremos em um espectro de comparação, caracterizado por um valor de comparação de um coeficiente do espectro de comparação ser atribuído a cada um dos coeficientes de MDCT do espectro de MDCT. Entretanto, o espectro de comparação pode ter uma resolução espectral mais alta do que o espectro de entrada do sinal de áudio. Por exemplo, o espectro de comparação pode ser um espectro de DFT tendo duas vezes a resolução espectral do que o espectro de entrada do sinal de áudio de MDCT. Por isso, apenas cada segundo valor espectral do espectro de DFT é então atribuído a um valor espectral do espectro de MDCT. Entretanto, os outros coeficientes do espectro de comparação podem ser considerados quando o extremo ou os extremos do espectro de comparação são determinados. Por isso, um coeficiente do espectro de comparação pode ser determinado como um extremo que não é atribuído a um coeficiente espectral do espectro de entrada do sinal de áudio, mas que tem um precursor imediato e um sucessor imediato, que são atribuídos a um coeficiente espectral do espectro de entrada do sinal de áudio e ao sucessor imediato deste coeficiente espectral do espectro de entrada do sinal de áudio, respectivamente. Assim, pode ser considerado que o referido extremo do espectro de comparação

(por exemplo, do espectro de DFT de alta resolução) é atribuído a uma localização espectral dentro do espectro de entrada do sinal de áudio (MDCT) que está localizado entre o referido coeficiente espectral do espectro de entrada do sinal de áudio (MDCT) e o referido sucessor imediato do referido coeficiente espectral do espectro de entrada do sinal de áudio (MDCT). Tal situação pode ser codificada pela escolha de um valor do sinal apropriado do pseudocoeficiente conforme explicado posteriormente. Por isso, a resolução subposição é obtida.

[0107] Deve ser observado que em algumas aplicações, um coeficiente extremo não tem que realizar a exigência que seu valor de comparação é superior ao valor de comparação de seu precursor imediato e o valor de comparação de seu sucessor imediato. Ao invés disso, nestas aplicações, pode ser suficiente que o valor de comparação do coeficiente extremo é superior a um de seus precursores e um de seus sucessores. Considere, por exemplo, a situação, onde:

Localização espectral	212	213	214	215	216
Valor de comparação	0,02	0,84	0,83	0,85	0,01

[0108] Tabela 1

[0109] Na situação descrita pela Tabela 1, o determinador de extremos 410 pode razoavelmente considerar o coeficiente espectral na localização espectral 214 como um coeficiente extremo. O valor de comparação do coeficiente espectral 214 não é superior ao seu precursor imediato 213 (0,83 < 0,84) e não superior ao seu sucessor imediato 215 (0,83 < 0,85), mas é (significantemente) superior ao valor de

comparação de outro de seus precursores, precursor 212 (0,83 > 0,02), e é (significantemente) superior ao valor de comparação de outro de seus sucessores, o sucessor 216 (0,83 > 0,01). Ainda parece razoável considerar o coeficiente espectral 214 como o extremo de sua "área mais elevada", pois o coeficiente espectral está localizado no meio dos três coeficientes 213, 214, 215 que têm valores de comparação relativamente grandes comparados aos valores de comparação de coeficientes 212 e 216.

[0110] Por exemplo, o determinador de extremos 410 pode ser configurado para determinar o formato de alguns ou de todos os coeficientes de comparação, se o valor de comparação do referido coeficiente de comparação for superior a, pelo menos, um dos valores de comparação dos três precursores sendo mais próximos à localização espectral do referido coeficiente de comparação. E/ou, o determinador de extremos 410 pode ser configurado para determinar o formato de alguns ou todos os coeficientes de comparação, se o valor de comparação do referido coeficiente de comparação for superior a, pelo menos, um dos valores de comparação dos três sucessores sendo mais próximos à localização espectral do referido coeficiente de comparação. O determinador de extremos 410 pode então decidir selecionar o referido coeficiente de comparação dependendo do resultado das referidas determinações.

[0111] Em algumas aplicações, o valor de comparação de cada coeficiente espectral é um valor quadrático de outro coeficiente de outro espectro (um espectro de comparação)

resultante a partir de uma transformação que preserva a energia do sinal de áudio.

[0112] Em outras aplicações, o valor de comparação de cada coeficiente espectral é um valor de amplitude de outro coeficiente de outro espectro resultante a partir de uma transformação que preserva uma energia do sinal de áudio.

[0113] De acordo com uma aplicação, outro espectro é um espectro da Transformada Discreta de Fourier e caracterizado pela transformação que preserva a energia ser uma Transformada Discreta de Fourier.

[0114] De acordo com uma aplicação adicional, outro espectro é um espectro da Transformada de Cosseno Discreta Modificada Complexa (CMDCT), e caracterizado pela transformação que preserva a energia ser uma CMDCT.

[0115] Em outra aplicação, o determinador de extremos 410 pode não examinar um espectro de comparação, mas ao invés disso, pode examinar o próprio espectro de entrada do sinal de áudio. Isso pode, por exemplo, ser razoável, quando o próprio espectro de entrada do sinal de áudio resulta a partir de uma transformação que preserva a energia, por exemplo, quando o espectro de entrada do sinal de áudio é um espectro de magnitude da Transformada Discreta de Fourier.

[0116] Por exemplo, o determinador de extremos 410 pode ser configurado para determinar um ou mais coeficiente(s) extremo(s), de modo que cada um dos coeficientes extremos seja um dos coeficientes espectrais cujo valor espectral que é superior ao valor espectral de um de seus precursores e cujo valor espectral que é superior ao

valor espectral de um de seus sucessores.

[0117] Em uma aplicação, o determinador de extremos 410 pode ser configurado para determinar um ou mais coeficiente(s) extremo(s), de modo que cada um dos coeficientes extremos seja um dos coeficientes espectrais cujo valor espectral que é superior ao valor espectral de seu precursor imediato e cujo valor espectral que é superior ao valor espectral de seu sucessor imediato.

[0118] Além disso, o aparelho compreende um modificador de espectro 420 para modificar o espectro de entrada do sinal de áudio para obter um espectro do sinal de áudio modificado pelo ajuste do valor espectral do precursor ou do sucessor de, pelo menos, um dos coeficientes extremos em um valor predefinido. O modificador de espectro 420 é configurado para não definir os valores espectrais de um ou mais coeficiente(s) extremo(s) ao valor predefinido, ou é configurado para substituir, pelo menos, um de um ou mais coeficiente(s) extremo(s) por um pseudocoeficiente, caracterizado pelo valor espectral do pseudocoeficiente ser diferente do valor predefinido.

[0119] Preferivelmente, o valor predefinido pode ser zero. Por exemplo, nenhum espectro do sinal de áudio modificado (substituído) 530 da figura 8, os valores espectrais de muitos coeficientes espectrais foram definidos como zero pelo modificador de espectro 420.

[0120] Em outras palavras, para obter o espectro do sinal de áudio modificado, o modificador de espectro 420 definirá, pelo menos, o valor espectral de um precursor ou um

sucessor de um dos coeficientes extremos em um valor predefinido. O valor predefinido pode, por exemplo, ser zero. O valor de comparação deste precursor ou sucessor é inferior ao valor de comparação do referido valor extremo.

[0121] Além disso, referente aos próprios coeficientes extremos, o modificador de espectro 420 procederá como segue:

- O modificador de espectro 420 não definirá os coeficientes extremos ao valor predefinido, ou:

- O modificador de espectro 420 substituirá, pelo menos, um dos coeficientes extremos por um pseudocoeficiente, caracterizado pelo valor espectral do pseudocoeficiente ser diferente do valor predefinido. Isso significa que o valor espectral de, pelo menos, um dos coeficientes extremos é definido ao valor predefinido, e o valor espectral de outro de um dos coeficientes espectrais é definido em um valor que é diferente do valor predefinido. Tal valor pode, por exemplo, ser derivado do valor espectral do referido coeficiente extremo, de um dos precursores do referido coeficiente extremo ou de um dos sucessores do referido coeficiente extremo. Ou, tal valor pode, por exemplo, ser derivado a partir do valor de comparação do referido coeficiente extremo, de um dos precursores do referido coeficiente extremo ou de um dos sucessores do referido coeficiente extremo

[0122] O modificador de espectro 420 pode, por exemplo, ser configurado para substituir um dos coeficientes extremos por um pseudocoeficiente tendo um valor espectral

derivado a partir do valor espectral ou do valor de comparação do referido coeficiente extremo, do valor espectral ou do valor de comparação de um dos precursores do referido coeficiente extremo ou do valor espectral ou do valor de comparação de um dos sucessores do referido coeficiente extremo.

[0123] Além disso, o aparelho compreende uma unidade de processamento 430 para processar o espectro do sinal de áudio modificado para obter um espectro do sinal de áudio codificado.

[0124] Por exemplo, a unidade de processamento 430 pode ser qualquer tipo de codificador de áudio, por exemplo, um codificador de áudio MP3 (MPEG-1 Camada de Áudio III ou MPEG-2 Camada de Áudio III; MPEG = Grupo de Especialistas em Imagens com Movimento), um codificador de áudio para WMA (*Windows Media Audio* | Áudio de Mídia Windows), um codificador de áudio para arquivos WAVE ou um codificador de áudio MPEG-2/4 AAC (Codificação de Áudio Avançada) ou um codificador MPEG-D USAC (Fala Unificada e Codificação de áudio).

[0125] A unidade de processamento 430 pode, por exemplo, ser um codificador de áudio conforme descrito em [8] (ISO/IEC 14496-3:2005 - tecnologia da informação - Codificação de objetos visuais de áudio - Parte 3: Subparte, Áudio 4) ou conforme descrito em [9] (ISO/IEC 14496-3:2005 - Tecnologia de Informação - codificação de objetos audiovisuais - Parte 3: Áudio, Subparte 4). Por exemplo, a unidade de processamento 430 pode compreender um quantizador,

e/ou uma ferramenta de modelagem de ruído temporal como, por exemplo, descrito em [8] e/ou a unidade de processamento 430 pode compreender uma ferramenta da substituição de ruído perceptual como, por exemplo, descrito em [8].

[0126] Além disso, o aparelho compreende um gerador de informação adicional 440 para gerar e transmitir a informação adicional. O gerador de informação adicional 440 é configurado para localizar um ou mais candidato(s) pseudocoeficiente(s) dentro do espectro de entrada do sinal de áudio modificado gerado pelo modificador de espectro 420. Além disso, o gerador de informação adicional 440 é configurado para selecionar, pelo menos, um dos candidatos pseudocoeficientes como candidatos selecionados. Além disso, o gerador de informação adicional 440 é configurado para gerar a informação adicional de modo que a informação adicional indica os candidatos selecionados como os pseudocoeficientes.

[0127] Na aplicação ilustrada pela figura 7, o gerador de informação adicional 440 é configurado para receber as posições dos pseudocoeficientes (por exemplo, a posição de cada um dos pseudocoeficientes) pelo modificador de espectro 420. Além disso, na aplicação da figura 7, o gerador de informação adicional 440 é configurado para receber as posições dos candidatos pseudocoeficientes (por exemplo, a posição de cada um dos candidatos pseudocoeficientes).

[0128] Por exemplo, em algumas aplicações, a unidade de processamento 430 pode ser configurada para determinar os

candidatos pseudocoeficientes com base em um espectro do sinal de áudio quantizado. Em uma aplicação, a unidade de processamento 430 pode ter gerado o espectro do sinal de áudio quantizado pela quantização do espectro do sinal de áudio modificado. Por exemplo, a unidade de processamento 430 pode determinar, pelo menos, um coeficiente espectral do espectro do sinal de áudio quantizado como um candidato pseudocoeficiente, que tem um precursor imediato, o valor espectral que é igual ao valor predefinido (por exemplo, igual a 0), e que tem um sucessor imediato, o valor espectral que é igual ao valor predefinido.

[0129] De modo alternativo, em outras aplicações, a unidade de processamento 430 pode passar o espectro do sinal de áudio quantizado ao gerador de informação adicional 440 e ao gerador de informação adicional 440 pode determinar os candidatos pseudocoeficientes com base no espectro do sinal de áudio quantizado. De acordo com outras aplicações, os candidatos pseudocoeficientes são determinados em uma forma alternativa com base no espectro do sinal de áudio modificado.

[0130] A informação adicional gerada pelo gerador de informação adicional pode ser de um tamanho estático, predefinido ou seu tamanho pode ser estimado de forma iterativa em uma forma adaptativa por sinal. Neste caso, o tamanho real da informação adicional também é transmitido ao decodificador. De acordo com uma aplicação, o gerador de informação adicional 440 é configurado para transmitir o tamanho da informação adicional.

[0131] De acordo com uma aplicação, o determinador de extremos 410 é configurado para avaliar os coeficientes de comparação, por exemplo, os coeficientes do espectro de potência 520 na figura 8, e é configurado para determinar um ou mais coeficiente(s) mínimo(s), de modo que cada um dos coeficientes mínimos seja um dos coeficientes espectrais cujo valor de comparação é inferior ao valor de comparação de um de seu precursores e cujo valor de comparação é inferior ao valor de comparação de um de seu sucessores. Em tal aplicação, o modificador de espectro 420 pode ser configurado para determinar um valor de representação com base nos valores de comparação de um ou mais dos coeficientes extremos e de um ou mais dos coeficientes mínimos, de modo que o valor de representação seja diferente do valor predefinido. Além disso, o modificador de espectro 420 pode ser configurado para mudar o valor espectral de um dos coeficientes do espectro de entrada do sinal de áudio pela definição do referido valor espectral ao valor de representação.

[0132] Em uma aplicação específica, o determinador de extremos é configurado para avaliar os coeficientes de comparação, por exemplo, os coeficientes do espectro de potência 520 na figura 8, e é configurado para determinar um ou mais coeficiente(s) mínimo(s), de modo que cada um dos coeficientes mínimos seja um dos coeficientes espectrais cujo valor de comparação é inferior ao valor de comparação de seu precursor imediato e cujo valor de comparação é inferior ao valor de comparação de seu sucessor imediato.

[0133] De modo alternativo, o determinador de

extremos 410 é configurado para avaliar o próprio espectro de entrada do sinal de áudio 510 e é configurado para determinar um ou mais coeficiente(s) mínimo(s), de modo que cada um ou mais coeficiente(s) mínimo(s) seja um dos coeficientes espectrais cujo valor espectral é inferior ao valor espectral de um de seus precursores e cujo valor espectral é inferior ao valor espectral de um de seus sucessores. Em tal aplicação, o modificador de espectro 420 pode ser configurado para determinar um valor de representação com base nos valores espectrais de um ou mais dos coeficientes extremos e de um ou mais dos coeficientes mínimos, de modo que o valor de representação seja diferente do valor predefinido. Além disso, o modificador de espectro 420 pode ser configurado para mudar o valor espectral de um dos coeficientes do espectro de entrada do sinal de áudio pela definição do referido valor espectral ao valor de representação.

[0134] Em uma aplicação específica, o determinador de extremos 410 é configurado para avaliar o próprio espectro de entrada do sinal de áudio 510 e é configurado para determinar um ou mais coeficiente(s) mínimo(s), de modo que cada um ou mais coeficiente(s) mínimo(s) seja um dos coeficientes espectrais cujo valor espectral é inferior ao valor espectral de seu precursor imediato e cujo valor espectral é inferior ao valor espectral de seu sucessor imediato.

[0135] Em ambas as aplicações, o modificador de espectro 420 considera o coeficiente extremo e um ou mais dos coeficientes mínimos, em particular seus valores de

comparação associados ou seus valores espectrais, para determinar o valor de representação. Então, o valor espectral de um dos coeficientes espectrais do espectro de entrada do sinal de áudio é definido ao valor de representação. Para o coeficiente espectral, o valor espectral que é definido no valor de representação pode, por exemplo, ser o próprio coeficiente extremo, ou o coeficiente espectral, o valor espectral que é definido no valor de representação pode ser o pseudocoeficiente que substitui o coeficiente extremo.

[0136] Em uma aplicação, o determinador de extremos 410 pode ser configurado para determinar uma ou mais subsequência(s) da sequência de valores espectrais, de modo que cada uma das subsequências compreenda uma pluralidade de coeficientes espectrais subsequentes do espectro de entrada do sinal de áudio. Os coeficientes espectrais subsequentes são sequencialmente organizados dentro da subsequência, de acordo com sua posição espectral. Cada uma das subsequências tem um primeiro elemento sendo primeiro na referida subsequência sequencialmente-organizada e um último elemento sendo o último na referida subsequência sequencialmente-organizada.

[0137] Em uma aplicação específica, cada uma das subsequências pode, por exemplo, compreender exatamente dois dos coeficientes mínimos e exatamente um dos coeficientes extremos, um dos coeficientes mínimos sendo o primeiro elemento da subsequência, o outro coeficiente dos coeficientes mínimos sendo o último elemento da subsequência.

[0138] Em uma aplicação, o modificador de espectro

420 pode ser configurado para determinar o valor de representação com base nos valores espectrais ou nos valores de comparação dos coeficientes de uma das subsequências. Por exemplo, se o determinador de extremos 410 examinou os coeficientes de comparação do espectro de comparação, por exemplo, do espectro de potência 520, o modificador de espectro 420 pode ser configurado para determinar o valor de representação com base nos valores de comparação dos coeficientes de uma das subsequências. Se, entretanto, o determinador de extremos 410 examinou os coeficientes espectrais do espectro de entrada do sinal de áudio 510, o modificador de espectro 420 pode ser configurado para determinar o valor de representação com base nos valores espectrais dos coeficientes de uma das subsequências.

[0139] O modificador de espectro 420 é configurado para mudar o valor espectral de um dos coeficientes da referida subsequência pela definição do referido valor espectral no valor de representação.

[0140] A tabela 2 fornece um exemplo com cinco coeficientes espectrais nas localizações espectrais de 252 a 258.

Localização espectral	252	253	254	255	256	257	258
Valor de comparação	0,12	0,05	0,48	0,73	0,45	0,03	0,18

[0141] Tabela 2

[0142] O determinador de extremos 410 pode determinar que o coeficiente espectral 255 (o coeficiente espectral com a localização espectral 255) é um coeficiente

extremo, pois seu valor de comparação (0,73) é superior ao valor de comparação (0,48) de seu precursor (aqui: imediato) 254, e pois seu valor de comparação (0,73) é superior ao valor de comparação (0,45) de seu sucessor (aqui: imediato) 256.

[0143] Além disso, o determinador de extremos 410 pode determinar que o coeficiente espectral 253 (é o coeficiente mínimo, pois seu valor de comparação (0,05) é inferior ao valor de comparação (0,12) de seu precursor (aqui: imediato) 252, e pois seu valor de comparação (0,05) é inferior ao valor de comparação (0,48) de seu sucessor (aqui: imediato) 254.

[0144] Além disso, o determinador de extremos 410 pode determinar que o coeficiente espectral 257 é um coeficiente mínimo, pois seu valor de comparação (0,03) é inferior ao valor de comparação (0,45) de seu precursor (aqui: imediato) 256 e seu valor de comparação (0,03) é inferior ao valor de comparação (0,18) de seu sucessor (aqui: imediato) 258.

[0145] O determinador de extremos 410 pode então determinar uma subsequência, compreendendo os coeficientes espectrais 253 a 257, pela determinação que o coeficiente espectral 255 é um coeficiente extremo, pela determinação do coeficiente espectral 253 como o coeficiente mínimo sendo o coeficiente mínimo precedente mais próximo ao coeficiente extremo 255, e pela determinação do coeficiente espectral 257 como o coeficiente mínimo sendo o coeficiente mínimo sucessor mais próximo ao coeficiente extremo 255.

[0146] O modificador de espectro 420 pode agora determinar um valor de representação para subsequência 253 - 257 com base nos valores de comparação de todos os coeficientes espectrais 253 - 257.

[0147] Por exemplo, o modificador de espectro 420 pode ser configurado para somar os valores de comparação de todos os coeficientes espectrais da subsequência. (Por exemplo, para a Tabela 2, o valor de representação para a subsequência 253 - 257 então soma até: 0,05 + 0,48 + 0,73 + 0,45 + 0,03 = 1,74).

[0148] Ou, por exemplo, o modificador de espectro 420 pode ser configurado para somar os quadrados dos valores de comparação de todos os coeficientes espectrais da subsequência. (Por exemplo, para a Tabela 2, o valor de representação para a subsequência 253 - 257 então soma até: $(0,05)^2 + (0,48)^2 + (0,73)^2 + (0,45)^2 + (0,03)^2 = 0,9692$).

[0149] Ou, por exemplo, o modificador de espectro 420 pode ser configurado para colocar em raiz quadrada a soma dos quadrados dos valores de comparação de todos os coeficientes espectrais da subsequência 253 - 257. (Por exemplo, para a Tabela 2, o valor de representação é então 0,98448).

[0150] De acordo com algumas aplicações, o modificador de espectro 420 definirá o valor espectral do coeficiente extremo (na Tabela o valor espectral do coeficiente espectral 253) ao valor predefinido.

[0151] Outras aplicações, entretanto, utilizam uma abordagem do centro da gravidade. A Tabela 3 ilustra uma

subsequência, compreendendo os coeficientes espectrais 282 - 288:

Localização espectral	281	282	283	284	285	286	287	288	289
Valor de comparação	0,12	0,04	0,10	0,20	0,93	0,92	0,90	0,05	0,15

[0152] Tabela 3

[0153] Embora o coeficiente extremo esteja localizado na localização espectral 285, de acordo com a abordagem do centro de gravidade, o centro de gravidade está localizado em uma diferente localização espectral.

[0154] Para determinar a localização espectral do centro de gravidade, o determinador de extremos 410 soma até as localizações espectrais ponderadas de todos os coeficientes espectrais da subsequência e divide o resultado pela soma dos valores de comparação dos coeficientes espectrais da subsequência. O arredondamento comercial pode então ser empregado no resultado da divisão para determinar o centro de gravidade. A localização espectral ponderada de um coeficiente espectral é o produto de sua localização espectral e seus valores de comparação.

[0155] Em resumo: o determinador de extremos pode obter o centro de gravidade por:

1) Determinar o produto do valor de comparação e localização espectral para cada coeficiente espectral da subsequência.

2) Somar os produtos determinados em 1) para obter uma primeira soma

3) Somar os valores de comparação de todos os coeficientes espectrais da subsequência para obter uma segunda soma

4) Dividir a primeira soma pela segunda soma para gerar um resultado intermediário; e

5) Aplicar o arredondamento ao próximo número inteiro no resultado intermediário para obter o centro de gravidade (arredondamento ao próximo número inteiro: 8,49 é arredondado para 8; 8,5 é arredondado para 9)

Assim, para o exemplo da Tabela 3, o centro de gravidade é obtido por:

(0,04 · 282 + 0,10 · 283 + 0,20 · 284 + 0,93 · 285 + 0,92 · 286 + 0,90 · 287 + 0,05 · 288) /

/ (0,04 + 0,10 + 0,20 + 0,93 + 0,92 + 0,90 + 0,05) = 897.25 / 3.14 = 285.75 = 286.

[0156] Assim, no exemplo da Tabela 3, o determinador de extremos 410 seria configurado para determinar a localização espectral 286 como o centro de gravidade.

[0157] Em algumas aplicações, o determinador de extremos 410 não avalia o completo espectro de comparação (por exemplo, o espectro de potência 520) ou não avalia o completo espectro de entrada do sinal de áudio. Ao invés disso, o determinador de extremos 410 pode apenas avaliar parcialmente o espectro de comparação ou o espectro de entrada do sinal de áudio.

[0158] A figura 9 ilustra tal exemplo. O espectro de potência 620 (como um espectro de comparação) foi avaliado por um determinador de extremos 410 começando no coeficiente 55. Os coeficientes nas localizações espectrais inferiores a 55 não foram avaliados. Assim, os coeficientes espectrais nas localizações espectrais inferiores a 55 permanecem não modificados no espectro de MDCT substituído 630. Ao contrário, a figura 8 ilustra um espectro de MDCT substituído 530 onde todas as linhas espectrais de MDCT foram modificadas

pelo modificador de espectro 420.

[0159] Assim, o modificador de espectro 420 pode ser configurado para modificar o espectro de entrada do sinal de áudio, de modo que os valores espectrais de, pelo menos, alguns dos coeficientes espectrais do espectro de entrada do sinal de áudio sejam deixados sem mudança.

[0160] Em algumas aplicações, o modificador de espectro 420 é configurado para determinar se uma diferença do valor entre um valor de comparação ou o valor espectral de um dos coeficientes extremos é inferior a um valor limite. Nestas aplicações, o modificador de espectro 420 é configurado para modificar o espectro de entrada do sinal de áudio, de modo que os valores espectrais de, pelo menos, alguns dos coeficientes espectrais do espectro de entrada do sinal de áudio sejam deixados sem modificação no espectro do sinal de áudio modificado dependendo se a diferença do valor é inferior ao valor limite.

[0161] Por exemplo, em uma aplicação, o modificador de espectro 420 pode ser configurado para modificar ou substituir todos, mas ao invés de modificar ou substituir apenas alguns dos coeficientes extremos. Por exemplo, quando a diferença entre o valor de comparação do coeficiente extremo (por exemplo, um máximo local) e o valor de comparação do valor mínimo subsequente e/ou precedente for inferior a um valor limite, o modificador de espectro pode ser determinado não para modificar estes valores espectrais (e, por exemplo, os valores espectrais dos coeficientes espectrais entre eles), mas para deixar estes valores

espectrais não modificados no espectro de MDCT modificado (substituído) 630. No espectro de MDCT modificado 630 da figura 9, os valores espectrais dos coeficientes espectrais de 100 a 112 e os valores espectrais dos coeficientes espectrais de 124 a 136 foram deixados sem modificação pelo modificador espectral no espectro não modificado (substituído) espectro 630.

[0162] A unidade de processamento pode ainda ser configurada para quantizar os coeficientes do espectro de MDCT modificado (substituído) 630 para obter um espectro de MDCT quantizado 635.

[0163] De acordo com uma aplicação, o modificador de espectro 420 pode ser configurado para receber a informação de sintonização precisa. Os valores espectrais dos coeficientes espectrais do espectro de entrada do sinal de áudio podem ser valores com sinal, cada um compreendendo um componente do sinal. O modificador de espectro pode ser configurado para definir o componente do sinal de um ou mais coeficiente(s) extremo(s) ou do pseudocoeficiente em um primeiro valor do sinal, quando a informação de sintonização precisa estiver em um primeiro estado de sintonização precisa. E o modificador de espectro pode ser configurado para definir o componente do sinal do valor espectral de um ou mais coeficiente(s) extremo(s) ou do pseudocoeficiente em um diferente segundo valor do sinal, quando a informação de sintonização precisa estiver em um segundo estado de sintonização precisa diferente.

[0164] Por exemplo, na Tabela 4,

Localização espectral	291	301	321	329	342	362	388	397	405
Valor espectral	+0,88	-0,91	+0,79	-0,82	+0,93	-0,92	-0,90	+0,95	-0,92
Estado de sintonização precisa	1°	2°	1°	2°	1°	2°	2°	1°	2°

[0165] Tabela 4

[0166] os valores espectrais dos coeficientes espectrais indicam que o coeficiente espectral 291 está em um primeiro estado de sintonização precisa, o coeficiente espectral 301 está em um segundo estado de sintonização precisa, o coeficiente espectral 321 está no primeiro estado de sintonização precisa, etc.

[0167] Por exemplo, retornando à determinação do centro de gravidade explicada acima, se o centro de gravidade está (por exemplo, aproximadamente no meio) entre duas localizações espectrais, o modificador espectral pode definir o sinal, de modo que o segundo estado de sintonização precisa seja indicado.

[0168] De acordo com uma aplicação, a unidade de processamento 430 pode ser configurada para quantizar o espectro do sinal de áudio modificado para obter um espectro do sinal de áudio quantizado. A unidade de processamento 430 pode, ainda, ser configurada para processar o espectro do sinal de áudio quantizado para obter um espectro do sinal de áudio codificado.

[0169] Além disso, a unidade de processamento 430 pode, ainda, ser configurada para gerar a informação adicional indicando apenas para estes coeficientes espectrais

do espectro do sinal de áudio quantizado que têm um precursor imediato cujo valor espectral é igual ao valor predefinido e um sucessor imediato, cujo valor espectral que é igual ao valor predefinido, se o referido coeficiente é um dos coeficientes extremos.

[0170] Esta informação pode ser fornecida pelo determinador de extremos 410 à unidade de processamento 430.

[0171] Por exemplo, tal informação pode ser armazenada pela unidade de processamento 430 em um campo de bit, indicando para cada um dos coeficientes espectrais do espectro do sinal de áudio quantizado que tem um precursor imediato cujo valor espectral é igual ao valor predefinido e um sucessor imediato, cujo valor espectral é igual ao valor predefinido, se o referido coeficiente for um dos coeficientes extremos (por exemplo, por um valor de bit 1) ou se o referido coeficiente não for um dos coeficientes extremos (por exemplo, por um valor de bit 0). Em uma aplicação, um decodificador pode posteriormente utilizar esta informação para recuperar o espectro de entrada do sinal de áudio. O campo de bit pode ter um comprimento fixo ou um comprimento escolhido adaptavelmente do sinal. No último caso, o comprimento do campo de bit pode ser adicionalmente conduzido ao decodificador.

[0172] Por exemplo, um campo de bit [000111111] gerado pela unidade de processamento 430 pode indicar que os três primeiros coeficientes "independentes" (seu valor espectral não é igual ao valor predefinido, mas os valores espectrais de seu precursor e de seu sucessor são iguais ao

valor predefinido) que aparecem no espectro de sinal de áudio (sequencialmente organizado) (quantizado) não são coeficientes extremos, mas os próximos seis coeficientes "independentes" são coeficientes extremos. Este campo de bit descreve a situação que pode ser vista no espectro de MDCT quantizado 635 na figura 9, onde os três primeiros coeficientes "independentes" 5, 8, 25 não são coeficientes extremos, mas onde os próximos seis coeficientes "independentes" 59, 71, 83, 94, 116, 141 são coeficientes extremos.

[0173] Novamente, o precursor imediato do referido coeficiente espectral é outro coeficiente espectral que imediatamente precede o referido coeficiente espectral dentro do espectro do sinal de áudio quantizado, e o sucessor imediato do referido coeficiente espectral é outro coeficiente espectral que imediatamente sucede o referido coeficiente espectral dentro do espectro do sinal de áudio quantizado.

[0174] Os conceitos propostos melhoram a qualidade perceptual dos codecs de transformação com base no bloco convencional em baixas taxas de bit. É proposto substituir as regiões tonais locais nos espectros do sinal de áudio, abrangendo os mínimos locais próximos, abrangendo um máximo local, por pseudolinhas (também referidas como pseudocoeficientes) tendo, em algumas aplicações, uma energia ou nível semelhante como as referidas regiões a serem substituídas.

[0175] Em baixas taxas de bit, as aplicações

fornecem conceitos em como integrar de forma rigorosa a codificação da forma de onda e a codificação paramétrica para obter uma qualidade perceptual melhorada e uma escala melhorada da qualidade perceptual versus taxa de bit sobre as técnicas simples.

[0176] Em algumas aplicações, áreas mais elevadas (abrangendo os mínimos locais próximos, abrangendo um máximo local) de espectros podem ser completamente substituídas por um único sinusoide cada; em oposição aos codificadores sinusoidais que de forma iterativa subtraem os sinusoides sintetizados do residual. As áreas mais elevadas adequadas são extraídas em uma representação espectral uniforme e levemente clareada e são selecionadas com relação às determinadas características (altura máxima, forma máxima).

[0177] De acordo com algumas aplicações, estes sinusoides de substituição podem ser representadas como pseudolinhas (pseudocoeficientes) dentro do espectro a ser codificado e refletem a amplitude ou energia total do sinusoide (em oposição a, por exemplo, linhas de MDCT regulares correspondem à projeção real do valor verdadeiro).

[0178] De acordo com algumas aplicações, as pseudolinhas (pseudocoeficientes) podem ser marcadas como tais pela matriz indicadora da informação adicional.

[0179] Em algumas aplicações, a escolha de sinal das pseudolinhas pode denotar resolução de frequência da semissub-banda.

[0180] De acordo com algumas aplicações, uma frequência de corte inferior para substituição sinusoidal

pode ser aconselhável devido à resolução de frequência limitada (por exemplo, semissub-banda).

[0181] A seguir, os conceitos são fornecidos para gerar um sinal de saída de áudio com base em um sinal de áudio codificado. Estes conceitos implementam uma síntese eficiente de sinusoides e varreduras no domínio de MDCT.

[0182] A figura 1a ilustra um aparelho para gerar um sinal de saída de áudio com base em um espectro do sinal de áudio codificado de acordo com uma aplicação.

[0183] O aparelho compreende uma unidade de processamento 115 para processar o espectro do sinal de áudio codificado para obter um espectro do sinal de áudio decodificado compreendendo uma pluralidade de coeficientes espectrais, caracterizado por cada um dos coeficientes espectrais terem uma localização espectral dentro do espectro do sinal de áudio codificado e um valor espectral, em que os coeficientes espectrais são sequencialmente organizados, de acordo com sua localização espectral dentro do espectro do sinal de áudio codificado de modo que os coeficientes espectrais formam uma sequência de coeficientes espectrais.

[0184] Além disso, o aparelho compreende um determinador dos pseudocoeficientes 125 para determinar um ou mais pseudocoeficiente(s) do espectro do sinal de áudio decodificado, caracterizado por cada um dos pseudocoeficientes ser um dos coeficientes espectrais (pois cada um dos pseudocoeficientes é um dos coeficientes espectrais, cada um dos pseudocoeficientes tem uma localização espectral e um valor espectral).

[0185] Além disso, o aparelho compreende uma unidade de substituição 135 para substituir pelo menos um ou mais pseudocoeficiente(s) por um padrão espectral determinado para obter um espectro do sinal de áudio modificado, caracterizado pelo padrão espectral determinado compreender pelo menos dois coeficientes padrões, em que cada um de pelo menos dois coeficientes padrões tem um valor espectral.

[0186] Por exemplo, em algumas aplicações, a unidade de substituição 135 pode obter um padrão espectral como um padrão espectral obtido de uma unidade de armazenamento, caracterizado pela unidade de armazenamento ser compreendida pelo aparelho, e em que a unidade de armazenamento compreende uma base de dados ou uma memória. Em outras aplicações, a unidade de substituição 135 pode obter um padrão espectral de uma unidade remota, por exemplo, uma base de dados remota, por exemplo, localizada longe do aparelho. Em outras aplicações, o padrão será gerado analiticamente em tempo real (no tempo de execução, quando necessário). O padrão espectral obtido pode então ser empregado como o padrão espectral determinado. Ou, o padrão espectral determinado pode ser derivado do padrão espectral obtido, por exemplo, pela modificação do padrão espectral obtido.

[0187] Além disso, o aparelho compreende uma unidade de conversão de tempo do espectro 145 para converter o espectro do sinal de áudio modificado em um domínio de tempo para obter o sinal de saída de áudio.

[0188] A figura 1b ilustra um aparelho para gerar um sinal de saída de áudio com base em um espectro do sinal de

áudio codificado, de acordo com outra aplicação. O aparelho da figura 1b difere do aparelho da aplicação da figura 1a em que ainda compreende uma unidade de armazenamento 155 que compreende uma base de dados ou uma memória.

[0189] Em particular, o aparelho da aplicação da figura 1b ainda compreende uma unidade de armazenamento 155 compreendendo uma base de dados ou uma memória tendo armazenado dentro da base de dados ou dentro da memória uma pluralidade de padrões espectrais armazenados. Cada um dos padrões espectrais armazenados tem uma propriedade espectral (por exemplo, frequência constante, frequência de varredura - cada um em uma versão da localização sobre a posição ou entre posições - etc.). A unidade de substituição 135 é configurada para solicitar um dos padrões espectrais armazenados como um padrão espectral solicitado a partir da unidade de armazenamento 155. A unidade de armazenamento 155 é configurada para fornecer o referido padrão espectral solicitado. Além disso, a unidade de substituição 135 é configurada para substituir pelo menos um ou mais pseudocoeficiente(s) pelo padrão espectral determinado com base no padrão espectral solicitado.

[0190] Nas aplicações preferidas, os padrões espectrais armazenados não foram armazenados para frequências específicas. Isso exigiria quantidades enormes de memória. Assim cada padrão (por exemplo, um padrão sobre a posição constante, um padrão entre posição constante e alguns padrões para várias varreduras) é armazenado apenas uma vez. Este padrão geral é então solicitado de, por exemplo, uma base de

dados, adaptado à frequência alvo, por exemplo, em uma frequência alvo de 8200 Hz, adaptado à fase necessária (por exemplo, 0 rad), e então corrigido na localização alvo espectral.

[0191] Em uma aplicação, a unidade de substituição 135 é configurada para solicitar um dos padrões espectrais armazenados a partir da unidade de armazenamento 155 dependendo de uma primeira localização espectral derivada de pelo menos um de um ou mais pseudocoeficiente(s) determinado(s) pelo determinador dos pseudocoeficientes 125. Por exemplo, a solicitação depende da natureza do padrão (constante, varredura, etc.) e a adaptação padrão depende da localização espectral e do predecessor dentro de uma faixa sinusoidal ou uma fase inicial determinada adaptavelmente do sinal de uma faixa sinusoidal.

[0192] Em uma aplicação, a primeira localização espectral derivada de pelo menos um de um ou mais pseudocoeficiente(s) pode ser a localização espectral de um dos pseudocoeficientes.

[0193] Em outra aplicação, um ou mais pseudocoeficiente(s) são valores com sinal, cada um compreendendo um componente do sinal, e a unidade de substituição 135 é configurada para determinar a primeira localização espectral derivada com base na localização espectral de um pseudocoeficiente de um ou mais pseudocoeficiente(s) e com base no componente do sinal do referido pseudocoeficiente, de modo que a primeira localização espectral derivada seja igual à localização

espectral do referido pseudocoeficiente quando o componente do sinal tem um primeiro valor do sinal, e de modo que a primeira localização espectral derivada seja igual a uma localização modificada, a localização modificada resultante da mudança de localização espectral do referido pseudocoeficiente por um valor predefinido quando o componente do sinal tem um segundo valor diferente.

[0194] Por exemplo, uma resolução de frequência de meia posição das pseudolinhas pode ser sinalizada pelo sinal do referido pseudocoeficiente. O valor predefinido pelo qual a localização espectral do referido pseudocoeficiente é mudada pode então corresponder à metade da diferença de frequência, por exemplo, de duas posições subsequentes, por exemplo, quando um domínio de frequência de tempo é considerado, quando o componente do sinal do pseudocoeficiente tem o segundo valor do sinal.

[0195] Em uma aplicação específica, o determinador de pseudocoeficientes 125 é configurado para determinar dois ou mais pseudocoeficientes temporalmente consecutivos do espectro do sinal de áudio decodificado. A unidade de substituição 135 é configurada para atribuir um primeiro pseudocoeficiente e um segundo pseudocoeficiente de dois ou mais pseudocoeficientes temporalmente consecutivos a uma faixa dependendo se uma diferença absoluta entre a primeira localização espectral derivada do primeiro pseudocoeficiente e uma segunda localização espectral derivada do segundo pseudocoeficiente é menor do que um valor limite. A pluralidade de padrões espectrais armazenados sendo

armazenados dentro da base de dados ou da memória da unidade de armazenamento pode ser tanto padrões de tom estacionário quanto padrões de varredura de frequência. A unidade de substituição 135 pode então ser configurada para solicitar um dos padrões de tom estacionário a partir da unidade de armazenamento 155 quando a primeira localização espectral derivada do primeiro pseudocoeficiente da faixa é igual à segunda localização espectral derivada do segundo pseudocoeficiente da faixa. Além disso, a unidade de substituição 135 pode ser configurada para solicitar um dos padrões de varredura de frequência a partir da unidade de armazenamento 155 quando a primeira localização espectral derivada do primeiro pseudocoeficiente da faixa é diferente da segunda localização espectral derivada do segundo pseudocoeficiente da faixa.

[0196] Por exemplo, a primeira localização espectral derivada do primeiro pseudocoeficiente da faixa pode ser a localização espectral do primeiro pseudocoeficiente. Por exemplo, a segunda localização espectral derivada do segundo pseudocoeficiente da faixa pode ser a localização espectral do segundo pseudocoeficiente.

[0197] Por exemplo, um pseudocoeficiente pode ser atribuído a uma de uma pluralidade de posições de tempo-frequência ou a uma localização de frequência intermediária entre duas posições de tempo-frequência, por exemplo, à posição de tempo-frequência (n, k), caracterizado por n denotar tempo, e em que k denota frequência. A frequência da posição de tempo-frequência dos pseudocoeficiente ou a

localização de frequência entre as duas posições de tempo-frequência pode então indicar a localização espectral do pseudocoeficiente. Ao receber a posição de tempo-frequência (n, k) a unidade de substituição 135 verificará se já recebeu um pseudocoeficiente sendo atribuído a uma posição de tempo-frequência que imediatamente precede a posição de tempo-frequência do pseudocoeficiente atual no tempo (n-1) e que é igual ou próximo à frequência da posição de tempo-frequência do pseudocoeficiente atual (igual ou próximo a k). A unidade de substituição 135 então atribuirá ambos os pseudocoeficientes a uma faixa.

[0198] Por exemplo, o pseudocoeficiente tendo uma posição de tempo-frequência que imediatamente precede a posição de tempo-frequência atual no tempo pode ser considerado próximo à frequência da posição de tempo-frequência atual, se a diferença absoluta das frequências de ambas as frequências for menor do que um valor limite. (Por exemplo, se os índices de frequência forem considerados como frequências, se a diferença absoluta for menor do que 2).

[0199] Se ambos os pseudocoeficientes da faixa têm a mesma localização espectral, a unidade de substituição 135 considera isso como uma indicação que um tom estacionário está presente e solicita um padrão estacionário de tom tendo a frequência correspondente.

[0200] Entretanto, se as localizações espectrais dos coeficientes espectrais de uma faixa diferem, a unidade de substituição 135 considera isso como uma indicação que uma varredura está presente e solicita um padrão de varredura de

frequência a partir da unidade de armazenamento 155. A frequência indicada pela localização de frequência do pseudocoeficiente precedente dentro da faixa pode então indicar uma frequência inicial do padrão de varredura e da frequência indicada pela localização de frequência do pseudocoeficiente atual dentro da faixa pode então indicar uma frequência alvo do padrão de varredura.

[0201] De acordo com uma aplicação, a unidade de substituição 135 pode ser configurada para solicitar um primeiro padrão de varredura de frequência dos padrões de varredura de frequência a partir da unidade de armazenamento quando uma diferença de frequência entre o segundo pseudocoeficiente da faixa e o primeiro pseudocoeficiente da faixa é igual à metade de um valor predefinido.

[0202] Além disso, a unidade de substituição 135 pode ser configurada para solicitar um segundo padrão de varredura de frequência, sendo diferente do primeiro padrão de varredura de frequência, dos padrões de varredura de frequência a partir da unidade de armazenamento quando a diferença de frequência entre o segundo pseudocoeficiente da faixa e o primeiro pseudocoeficiente da faixa for igual ao valor predefinido.

[0203] Além disso, a unidade de substituição 135 pode ser configurada para solicitar um terceiro padrão de varredura de frequência, sendo diferente do primeiro padrão de varredura e do segundo padrão de varredura de frequência, dos padrões de varredura de frequência a partir da unidade de armazenamento quando a diferença de frequência entre o

segundo pseudocoeficiente da faixa e o primeiro pseudocoeficiente da faixa for igual a uma vez e meia do valor predefinido.

[0204] Por exemplo, o valor predefinido pode ser uma diferença de frequência entre duas posições de tempo-frequência temporalmente subsequentes. Assim, nesta aplicação, os padrões para varreduras são fornecidos onde a diferença de frequência entre uma frequência inicial e uma frequência alvo difere pela diferença da posição de frequência de 1/2, pela diferença da posição de frequência de 1,0 e pela diferença da posição de frequência de 3/2.

[0205] A figura 1c ilustra um aparelho de acordo com uma aplicação, onde a unidade de substituição 135 compreende uma unidade de adaptação padrão 138 sendo configurada para modificar o padrão espectral solicitado fornecido pela unidade de armazenamento 155 para obter o padrão espectral determinado.

[0206] Em uma aplicação, a unidade de adaptação padrão 138 pode ser configurada para modificar o padrão espectral solicitado fornecido pela unidade de armazenamento 155 pelo redimensionamento dos valores espectrais dos coeficientes padrões do padrão espectral solicitado dependendo do valor espectral de um ou mais pseudocoeficiente(s) para obter um padrão espectral determinado. A unidade de substituição espectral 135 é então configurada para substituir pelo menos um ou mais pseudocoeficiente(s) pelo padrão espectral determinado para obter o espectro do sinal de áudio modificado. Assim, de

acordo com esta aplicação, o tamanho dos valores espectrais dos coeficientes padrões do padrão espectral solicitado pode ser ajustado dependendo do valor espectral dos pseudocoeficiente.

[0207] De acordo com uma aplicação, a unidade de adaptação padrão 138 pode ser configurada para modificar o padrão espectral solicitado fornecido pela unidade de armazenamento dependendo de uma fase inicial de modo que o valor espectral de cada um dos coeficientes padrões do padrão espectral solicitado seja modificado em uma primeira forma, quando a fase inicial tem um primeiro valor da fase inicial, e de modo que o valor espectral de cada um dos coeficientes padrões do padrão espectral solicitado seja modificado em uma segunda forma diferente, quando a fase inicial tem um segundo valor da fase inicial diferente. Pelo ajuste da fase dos padrões de uma faixa, a transição sem descontinuidades de um padrão de uma faixa ao seguinte padrão pode ser obtida.

[0208] De acordo com uma aplicação, o valor espectral de cada um dos coeficientes padrões do padrão espectral solicitado é um coeficiente complexo compreendendo uma parte real e uma parte imaginária. A unidade de adaptação padrão 138 pode ser configurada para modificar o padrão espectral solicitado pela modificação da parte real e da parte imaginária de cada um dos coeficientes padrões do padrão espectral solicitado fornecido pela unidade de armazenamento 155, de modo que para cada um dos coeficientes complexos um vetor que representa o referido coeficiente complexo em um plano complexo seja girado pelo mesmo ângulo

para cada um dos coeficientes complexos. De modo alternativo, a fase de um padrão armazenado pode ser girada pela aplicação de um fator de rotação complexo $e^{j\cdot\varphi}$, com φ sendo um ângulo de fase arbitrária.

[0209] Em uma aplicação particular, o valor espectral de cada um dos coeficientes padrões do padrão espectral solicitado compreende uma parte real e uma parte imaginária. Em tal aplicação, a unidade de adaptação padrão 138 pode ser configurada para modificar o padrão espectral solicitado fornecido pela unidade de armazenamento 155 pela negação da parte real e da parte imaginária do valor espectral de cada um dos coeficientes padrões do padrão espectral solicitado, ou pela troca da parte real ou de uma parte real negada e da parte imaginária ou de uma parte imaginária negada do valor espectral de cada um dos coeficientes padrões do padrão espectral solicitado.

[0210] Em uma aplicação, a unidade de adaptação padrão 138 pode ser configurada para modificar o padrão espectral solicitado fornecido pela unidade de armazenamento 155 pela realização de um espelhamento temporal do padrão. Tipicamente, isso pode ser obtido em um domínio de frequência pelo cálculo computacional do conjugado complexo (pela multiplicação da parte imaginária por -1) do padrão e aplicando um termo da fase complexa (cruzado).

[0211] De acordo com uma aplicação, o espectro do sinal de áudio decodificado é representado em um domínio de MDCT. Em tal aplicação, a unidade de adaptação padrão 138 é então configurada para modificar o padrão espectral

solicitado fornecido pela unidade de armazenamento 155 pela modificação dos valores espectrais dos coeficientes padrões do padrão espectral solicitado para obter um padrão espectral modificado, caracterizado pelos valores espectrais serem representados em um domínio da Transformada Discreta de Fourier com Empilhamento Diferencial. Além disso, a unidade de adaptação padrão 138 está nesta aplicação configurada para transformar os valores espectrais dos coeficientes padrões do padrão espectral modificado do domínio da Transformada Discreta de Fourier com Empilhamento Diferencial ao domínio de MDCT para obter o padrão espectral determinado. Além disso, a unidade de substituição 135 está nesta aplicação configurada para substituir pelo menos um ou mais pseudocoeficiente(s) pelo padrão espectral determinado sendo representado no domínio de MDCT para obter o espectro do sinal de áudio modificado sendo representado no domínio de MDCT.

[0212] De modo alternativo, nas aplicações, os valores espectrais podem ser representados em um domínio da Transformada de Cosseno Discreta Modificada Complexa (CMDCT). Além disso, nestas aplicações a unidade de adaptação padrão 138 pode ser configurada para transformar os valores espectrais dos coeficientes padrões do padrão espectral modificado do domínio de CMDCT ao domínio de MDCT para obter o padrão espectral determinado simplesmente pela extração da parte real do padrão modificado complexo.

[0213] A figura 1d ilustra um aparelho para gerar uma pluralidade de padrões espectrais de acordo com uma

aplicação.

[0214] O aparelho compreende um gerador de sinal 165 para gerar uma pluralidade de sinais em um primeiro domínio.

[0215] Além disso, o aparelho compreende uma unidade de transformação do sinal 175 para transformar cada sinal da pluralidade de sinais do primeiro domínio em um segundo domínio para obter uma pluralidade de padrões espectrais, cada padrão da pluralidade de padrões espectrais transformados compreendendo uma pluralidade de coeficientes.

[0216] Além disso, o aparelho compreende uma unidade de pós-processamento 185 para truncar os padrões espectrais transformados pela remoção de um ou mais dos coeficientes dos padrões espectrais transformados para obter uma pluralidade de padrões processados.

[0217] Além disso, o aparelho compreende uma unidade de armazenamento 195 compreendendo uma base de dados ou uma memória, caracterizado pela unidade de armazenamento 195 ser configurada para armazenar cada padrão processado da pluralidade de padrões processados na base de dados ou na memória.

[0218] O gerador de sinal 165 é configurado para gerar cada sinal da pluralidade de sinais com base nas fórmulas

[0219] $$x(t) = cos\,(2\pi\varphi(t))$$

[0220] e

[0221] $$\varphi(t) = \varphi(0) + \int_0^t 2\pi f(\tau)d\tau \quad ,$$

[0222] caracterizado por t e τ indicarem tempo, em

que $\varphi(t)$ é uma fase instantânea em t, e em que $f(\tau)$ é uma frequência instantânea em τ, em que cada sinal da pluralidade de sinais tem uma frequência inicial (f_0), sendo uma frequência instantânea do referido sinal em um primeiro ponto no tempo, e uma frequência alvo (f_1), sendo uma frequência instantânea do referido sinal em um diferente segundo ponto no tempo.

[0223] O gerador de sinal 165 é configurado para gerar um primeiro sinal da pluralidade de sinais de modo que a frequência alvo (f_1) do primeiro sinal seja igual à frequência inicial (f_0). Além disso, o gerador de sinal 165 é configurado para gerar um segundo sinal diferente da pluralidade de sinais de modo que a frequência alvo (f_1) do primeiro sinal seja diferente da frequência inicial (f_0).

[0224] De acordo com uma aplicação, a unidade de transformação do sinal 175 é configurada para transformar cada sinal da pluralidade de sinais do primeiro domínio, sendo um domínio de tempo, em um segundo domínio, sendo um domínio espectral. A unidade de transformação do sinal 175 é configurada para gerar um primeiro de uma pluralidade de blocos de tempo para transformar o referido sinal, caracterizado por cada bloco de tempo da pluralidade de blocos de tempo compreender uma pluralidade de amostras ponderadas, em que cada uma das referidas amostras ponderadas é uma amostra do sinal do referido sinal sendo ponderado por uma ponderação de uma pluralidade de ponderações, em que a pluralidade de ponderações é atribuída ao referido bloco de tempo, e em que cada ponderação da pluralidade de ponderações

é atribuída a um ponto no tempo. A frequência inicial (f_0) de cada sinal da pluralidade de sinais é uma frequência instantânea do referido sinal no primeiro ponto no tempo, onde uma primeira das ponderações do primeiro de um dos blocos de tempo é atribuída ao primeiro ponto no tempo, onde uma segunda das ponderações de um diferente segundo dos blocos de tempo é atribuída ao primeiro ponto no tempo, em que o primeiro dos blocos de tempo e o segundo dos blocos de tempo se sobrepõem, e em que a primeira das ponderações é igual à segunda das ponderações. A frequência alvo (f_1) de cada sinal da pluralidade de sinais é uma frequência instantânea do referido sinal no segundo ponto no tempo, onde uma terceira das ponderações do primeiro de um dos blocos de tempo é atribuída ao segundo ponto no tempo, onde uma quarta das ponderações de um terceiro diferente dos blocos de tempo é atribuída ao segundo ponto no tempo, em que o primeiro dos blocos de tempo e o terceiro dos blocos de tempo se sobrepõem, e em que a terceira das ponderações é igual à quarta das ponderações.

[0225] Por exemplo, a figura 6a ilustra um exemplo, caracterizado pelo primeiro ponto no tempo ser indicado por n_0 e o segundo ponto no tempo é indicado por n_1. Os blocos de sobreposição são ilustrados por blocos L e L+1. As ponderações são descritas pela curva no bloco L e a curva no bloco L+1, respectivamente.

[0226] Deve ser observado que, por exemplo, é suficiente para gerar apenas um bloco de tempo (por exemplo, o primeiro dos blocos de tempo) para a geração de um padrão.

[0227] De acordo com uma aplicação, cada sinal da pluralidade de sinais tem uma fase inicial (φ_0), estando uma fase do referido sinal em um primeiro ponto no tempo, e uma fase alvo (φ_1), estando uma fase do referido sinal em um diferente segundo ponto no tempo, em que o gerador de sinal (165) é configurado para gerar a pluralidade de sinais de modo que a fase inicial (φ_0) de um primeiro da pluralidade de sinais seja igual à fase inicial (φ_0) de um segundo sinal diferente da pluralidade de sinais.

[0228] A fase inicial (e, implicitamente pela escolha da frequência inicial e de parada, a fase alvo (de parada)) de cada sinal da pluralidade de sinais é ajustada nos referidos pontos no tempo inicial e de parada.

[0229] Por esta escolha especial do primeiro (inicial) e do segundo (de parada) pontos no tempo, as perturbações adicionadas por sobreposição são reduzidas podendo ocorrer, se os padrões com diferentes propriedades espectrais são interligados.

[0230] Em uma aplicação, a unidade de pós-processamento 185 pode ser ainda configurada para conduzir uma rotação por $\pi/4$ nos coeficientes espectrais de cada um dos padrões espectrais transformados para obter uma pluralidade de padrões espectrais rodados.

[0231] De acordo com uma aplicação adicional, o gerador de sinal 165 pode ser configurado para gerar o primeiro sinal, o segundo sinal e um ou mais sinal(is) adicional(is) como a pluralidade de sinais, de modo que cada diferença da frequência alvo e da frequência inicial de cada

um dos sinais adicionais seja um múltiplo inteiro de uma diferença da frequência alvo e da frequência inicial do segundo sinal.

[0232] Por exemplo, a diferença de frequência da frequência alvo e da frequência inicial do segundo sinal pode corresponder a uma diferença de frequência de meia posição, por exemplo, uma diferença de frequência da metade da diferença de frequência de duas posições subsequentes quando posições de tempo-frequência são consideradas. A diferença de frequência da frequência alvo e da frequência inicial de outro terceiro sinal pode corresponder a uma diferença de frequência da posição, por exemplo, uma diferença de frequência correspondente à diferença de frequência de duas posições subsequentes quando as posições de tempo-frequência são consideradas. A diferença de frequência da frequência alvo e da frequência inicial de um quarto sinal adicional pode corresponder a diferença de frequência de posição de um e meio, por exemplo, uma diferença de frequência correspondente a um e meio da diferença de frequência de duas posições subsequentes quando as posições de tempo-frequência são consideradas.

[0233] Assim, o índice da diferença da frequência alvo e da frequência inicial do terceiro sinal à diferença da frequência alvo e da frequência inicial do segundo sinal é 2,0 (um valor de número inteiro). O índice da diferença da frequência alvo e da frequência inicial do quarto sinal para a diferença da frequência alvo e da frequência inicial do segundo sinal é 3,0 (um valor de número inteiro).

[0234] Antes de fornecer descrições das aplicações específicas em mais detalhes, para melhor explicação, as bases da MDCT são descritas.

[0235] A MDCT de um sinal real $x(n)$ é definida para os segmentos do sinal em janela com $w(n)$ no tempo l, isso é $w_a(l, n) \cdot x(l, n) \in \mathbb{R}$, de comprimento N como segue:

$$\begin{aligned} & X_{MDCT}(l,m) \\ & = MDCT\{w_a(l,n)\cdot x(l,n)\} \\ & = \sqrt{\tfrac{2}{M}}\sum_{n=0}^{N-1} w_a(l,n)\cdot x(l,n)\cos(\tfrac{\pi}{M}(m+\tfrac{1}{2})(n+\tfrac{1}{2}+\tfrac{M}{2})), \\ & M=\tfrac{N}{2}, m\in\{0,1,\ldots,M-1\}, \end{aligned} \tag{1}$$

[0236] O + 1/2 em (m + 1/2) representa a mudança de frequência. O (n + 1/2 + M/2) representa a mudança de tempo.

[0237] A transformada inversa é escrita como

$$\begin{aligned} & \tilde{x}(l,n) \\ & = MDCT^{-1}\{X(l,m)\} \\ & = w_s(l,n)\sqrt{\tfrac{2}{M}}\sum_{m=0}^{M-1} X(l,m)\cos(\tfrac{\pi}{M}(m+\tfrac{1}{2})(n+\tfrac{1}{2}+\tfrac{M}{2})), \\ & N=2M, n\in\{0,1,\ldots,N-1\}. \end{aligned} \tag{2}$$

[0238] A MDCT pode ser vista como a parte real da Transformada de Cosseno Discreta Modificada Complexa (CMDCT) que é definida como

$$\begin{aligned} & X_{CMDCT}(l,m) \\ & = CMDCT\{w_a(l,n)\cdot x(l,n)\} \\ & = \sqrt{\tfrac{2}{M}}\sum_{n=0}^{N-1} w_a(l,n)\cdot x(l,n)\cos(\tfrac{\pi}{M}(m+\tfrac{1}{2})(n+\tfrac{1}{2}+\tfrac{M}{2})) \\ & -j\sqrt{\tfrac{2}{M}}\sum_{n=0}^{N-1} w_a(l,n)\cdot x(l,n)\sin(\tfrac{\pi}{M}(m+\tfrac{1}{2})(n+\tfrac{1}{2}+\tfrac{M}{2})), \\ & M=\tfrac{N}{2}, m\in\{0,1,\ldots,M-1\}. \end{aligned} \tag{3}$$

[0239] Além disso, a CMDCT pode ser expressa como uma Transformada Discreta de Fourier com Empilhamento

Diferencial (ODFT | *Oddly-Stacked Discrete Fourier Transform*) ou Transformada Discreta de Fourier (DFT | *Discrete Fourier Transform*) e termos da fase de pré- e pós-cruzamento exponencial

$$
\begin{aligned}
&X_{CMDCT}(l,m)\\
&= CMDCT\{w_a(l,n)\cdot x(l,n)\}\\
&= ODFT\{w_a(l,n)\cdot x(l,n)\}\cdot e^{-j\frac{\pi}{M}(m+\frac{1}{2})(\frac{1}{2}+\frac{M}{2})}\\
&= DFT\{w_a(l,n)\cdot x(l,n)\cdot e^{-j\frac{\pi}{2M}n}\}\cdot e^{-j\frac{\pi}{M}(m+\frac{1}{2})(\frac{1}{2}+\frac{M}{2})},\\
&M=\frac{N}{2}, m\in\{0,1,\ldots,M-1\}.
\end{aligned}
\tag{4}
$$

[0240] O $e^{-j\frac{\pi}{M}(m+\frac{1}{2})(\frac{1}{2}+\frac{M}{2})}$ representa a mudança de tempo por pós-cruzamento.

[0241] A seguir, a extração e a correção dos padrões do tom no domínio de MDCT são descritas. Agora, algumas explicações são fornecidas com relação às popularidades particulares de MDCT. Em particular, primeiramente, as provisões para a MDCT são consideradas.

[0242] Como pode ser visto a partir das Equações 4 que compreendem um exponencial chamado termo de pós-cruzamento, a CMDCT tem funções base mudadas por tempo comparado à DFT ou ODFT. Assim, é desejado desacoplar a descompensação da fase absoluta φ_0 dos sinusoides corrigidos da posição espectral da aplicação de correção, este cruzamento deve ser considerado.

[0243] As aplicações conduzem a extração padrão e a correção no domínio de ODFT e pós-processam a superposição de todos os padrões pela aplicação do referido cruzamento antes de misturar com os coeficientes de MDCT.

[0244] Cada correção é obtida pela extração dos

espectros de ODFT do complexo truncado dos sinusoides prototípicos ou varreduras geradas, de acordo com as seguintes equações. Um sinusoide com frequência instantânea variável (IF) $f(t)$ pode ser sintetizado como

$$x(t) = cos\,(2\pi\varphi(t)) \qquad (5)$$

[0245] com a fase instantânea

$$\varphi(t) = \varphi(0) + \int_0^t 2\pi f(\tau)d\tau \qquad (6)$$

[0246] Para simplificar a relação entre a descrição de MDCT discreta de tempo e sinusoide contínuo de tempo uma taxa de amostragem normalizada $fs = 1$ é assumida a seguir. A frequência instantânea (FI) $f(\tau)$ dos modelos de varredura é escolhida de modo que FI inicial e alvo são exatamente obtidos nos pontos de simetria de cancelamento de distorção do domínio de tempo (TDAC | *time domain distorção cancellation*) $t_0 = N/4 + 0,5$ e $t_1 = 3N/4 + 0,5$ de cada bloco de tempo de MDCT do comprimento N, respectivamente. Uma varredura linear da frequência f_0 para f_1 abrangendo uma faixa de frequência $\Delta f = f_1 - f_0$ em um intervalo de tempo de comprimento $M = N/2$ tem uma frequência instantânea (FI)

$$f(t_0 + t) = f_0 + \frac{\Delta f}{M}t \qquad (7)$$

[0247] levando a uma fase instantânea

$$\varphi(t_0 + t) = \varphi(t_0) + f_0 t + \frac{\Delta f}{2M}t^2 \qquad (8)$$

[0248] sinusoides com frequências inicial e final da resolução dupla (comparado a MDCT a ser empregada para síntese padrão) podem ser gerados selecionando $f_0 = k\frac{\pi}{2M}$ e $f_1 = (k+m)\frac{\pi}{2M}$, com descompensação da frequência m medida nos

índices da posição de transformação. Índices ímpares correspondem a frequências "sobre a posição" e índices pares apresentam frequências "entre posições". O progresso da fase entre as estruturas subsequentes pode ser calculado como

$$\Delta\varphi = \varphi(t_1) - \varphi(t_0) = f_0 M + \frac{\Delta f}{2M} M^2 = k\frac{\pi}{2} + m\frac{\pi}{4} \quad (9)$$

[0249] Isso significa que para interligação temporal sem descontinuidades de padrões a fase de cada correção pode ser ajustada por um múltiplo inteiro de $\frac{\pi}{4}$ dependendo do índice da frequência inicial *k* e o índice da descompensação da frequência *m* odo padrão precedente. A variável *m* pode também ser vista como a taxa de varredura, onde, por exemplo, *m* = 1 denota uma varredura de meia posição sobre a duração de um bloco de tempo.

[0250] Além disso, a compensação para mudança espectral da posição do número inteiro pode ser conduzida. A posição espectral destes sinusoides prototípicos ou varreduras é beneficamente escolhida para estar localizada no meio do espectro a fim de reduzir os erros se dobra cíclica. Dependente da distância espectral *d* do sinusoide prototípico e localização alvo de correção, a correção é adaptada pelas rotações de pós-processamento de *dπ* / 2 para sempre obter uma fase fixa predefinida independentemente da localização alvo de correção. Em outras palavras, uma rotação de pós-processamento compensa a rotação indesejada da fase que é inerentemente causada pela mudança espectral.

[0251] Agora, as considerações de eficiência e precisão são fornecidas. Primeiramente, a eficiência do

cálculo computacional é considerada:

[0252] A tabela I fornece operações para realizar diferentes pós-cruzamentos. Para manter a quantidade de padrões a ser armazenados razoavelmente pequenos e, mais importante, para poder explorar o fato que as rotações pelas determinadas frações simples de π podem ser obtidas pelas operações listadas na Tabela I, as possíveis frequências e varreduras devem ser restritas.

rotação	operação	implementação
0	1	Cópia de "0" padrões
$\frac{\pi}{2}$	i	Varredura $\Re$ e parte $-\Im$ de "0" padrões
π	-1	Negação de "0" padrões
$\frac{3\pi}{2}$	$-i$	Varredura $-\Re$ e parte $\Im$ de "0" padrões
$\frac{\pi}{4} + \frac{n\pi}{2}$	dto.	Faça o ref. acima em padrão "$\frac{\pi}{4}$"

[0253] Tabela I

[0254] (OPERAÇÕES PARA ROTAÇÕES SIMPLES)

[0255] A seguir, a resolução de frequência é considerada. Estas restrições são, ao mesmo tempo, necessárias para permitir uma reprodução perceptualmente satisfatória das partes do sinal parametricamente codificada. Visto que tal parte do sinal pode compreender uma sequência arbitrária de padrões do tom de tempo, cada grau adicional de liberdade multiplica o número de padrões a ser armazenado ou, de modo alternativo, os custos do cálculo computacional para adaptação de padrões. Assim, faz sentido escolher a resolução espectral de modo que nenhum efeito de dessintonização é percebido pelo ouvinte médio na faixa espectral alvo pretendida.

[0256] Ouvintes e músicos treinados podem perceber

as dessintonizações abaixo de 5 centos, o ouvinte médio pode aceitar os desvios de aproximadamente 10 centos (um décimo de um semitom). Assim, a substituição espectral dos tons de seno deve apenas ser feita acima de uma determinada frequência de corte que corresponde ao cenário do pior caso da dessintonização permitida. Por exemplo, em uma MDCT com banda 512, em uma frequência de amostragem de 12,8 kHz, a resolução espectral por banda é 12,5 Hz. A escolha da resolução de maia banda para os padrões do tom, as quantidades máximas de desvio de frequência a 3,125 Hz, que é igual ou abaixo de 10 cento acima de uma frequência de corte de aproximadamente 540 Hz.

[0257] Agora, o tamanho padrão é considerado. De acordo com as aplicações, os padrões a ser armazenados são truncados. O tamanho real dos padrões depende do tipo de janela que é geralmente determinado pelo codificador de transformação (por exemplo, janela de seno ou derivada de Kaiser-Bessel (KBD | *Kaiser-Bessel derived*) para AAC) e o índice sinal para ruído (SNR | *signal-to-noise ratio*). Embora os padrões com valor complexo sejam armazenados, a correção real é feita apenas utilizando a parte real do padrão adequadamente girado.

[0258] A seguir, os padrões do tom são considerados. Primeiramente, os padrões de tom estacionário são estacionários.

[0259] Para as razões previamente mencionadas, a resolução espectral deve ser escolhida duas vezes a resolução nominal da MDCT. Como uma consequência, duas versões de todos

os padrões devem ser armazenadas, uma paro sinusoides com frequências que coincidem com uma posição (padrão sobre a posição) e uma para frequências que estão localizadas entre as posições (padrão entre posições). Para as menores exigências de memória possíveis, a simetria dos padrões pode ser explicada pelo armazenamento apenas da metade dos coeficientes do padrão real.

[0260] De acordo com a Equação 9 (definindo m = 0), em qualquer sequência de tempo destes padrões de tom estacionário, as quantidades do progresso da fase envolvidas em $\Delta\varphi = \pi / 2$ ou $\Delta\varphi = -\pi / 2$ para padrões sobre a posição, e $\Delta\varphi = 0$ ou $\Delta\varphi = \pi$ para padrões entre posições. Isso é devido ao empilhamento de frequência diferencial da MDCT.

[0261] A fase envolvida absoluta pode ser calculada por $\varphi_0 + n\ \pi / 2$ *com* n como um número inteiro $\in \{1, 3\}$ para padrões sobre uma posição e $\in \{2, 4\}$ para padrões entre posições. A escolha do número inteiro real depende da paridade do número da posição (par/ímpar). φ_0 denota um valor de descompensação da fase arbitrária. Assim, para padrão puramente estacionário do tom, um pós-processamento por quatro rotações alternativas é necessário para adequar os padrões em sua posição pretendida na grade t/f de uma sequência de espectros de MDCT. Uma escolha de $\varphi_0 + n\ \pi / 2$, $n \in \mathbb{N}$ apresenta estas rotações comuns.

[0262] Agora, os padrões de varredura de frequência são considerados.

[0263] Devido à resolução espectral sendo duas vezes a resolução nominal da MDCT, também duas versões de cada

padrão de varredura precisam ser armazenadas, uma para varreduras com frequências iniciais que coincidem com uma posição e uma para as frequências iniciais que estão localizadas entre as posições. Além disso, as varreduras permitidas são definidas para ser lineares e para cobrir meia posição, uma posição completa e uma posição e meia de MDCT por bloco de tempo, cada uma em uma versão de direção para baixo e para cima, resultante em 12 padrões a ser armazenados adicionalmente. Para as menores exigências possíveis de memória, os padrões de varredura podem ser armazenados apenas em uma direção; a direção oposta pode ser derivada pelo paralelismo temporal do padrão. De acordo com a Equação 9 (definindo $m \in \{ 1, 3, 5 \ldots \}$), o padrão que envolve as distâncias de varredura de meia posição exigem rotação de pós-processamentos por $\varphi_0 + n \pi / 4$.

[0264] A seguir, a interligação dos padrões é considerada. Para esta finalidade, a referência é feita à figura 2. A figura 2 ilustra o alinhamento do parâmetro do padrão sinusoidal com relação ao bloco de tempo de MDCT. Se os padrões são interligados em uma sequência temporal, uma fase inicial para o padrão real no ponto n_0 da figura 2 deve ser escolhida (utilizando as rotações previamente mencionadas) e a fase alvo (fase de parada) no ponto n_1 deve ser armazenada para continuação sem descontinuidades com o padrão subsequente.

[0265] Varreduras que abrangem as distâncias da varredura de meia posição são pós-processadas pela rotação de pós-processamentos por $\varphi_0 + n \pi / 4$, para tanto os padrões de

varredura quanto para padrões estacionários, visto que as varreduras e partes estacionárias podem ser arbitrariamente interligadas em uma sequência de tempo. Uma escolha de $\varphi_0 + n\,\pi / 4$, $n \in N$ resulta em uma rotação que também é fácil de calcular pela soma/diferença da parte real e da parte imaginária do padrão e uma escala subsequente por $\frac{\sqrt{2}}{2}$. De modo alternativo, todos os padrões podem ser adicionalmente armazenados em uma versão pré-girada $\pi / 4$ e podem ser aplicados juntos com uma rotação de pós-processamento comum por $n\,\pi / 2$, $n = 1, 2, 3$ (veja a Tabela 1).

[0266] A figura 3 ilustra um processo de correção de padrões do tom exemplar, em que (a-b) ilustram a geração do padrão prototípico, em que (c) ilustra a truncagem padrão, em que (d) ilustra a adaptação padrão na localização e fase alvo, e em que (e-f) ilustram a correção padrão.

[0267] Em particular, no painel da figura 3 (a)-(zf), todo o processo, conforme descrito acima com relação às particularidades da MDCT, da medição padrão até a adaptação padrão e correção é descrito. Primeiramente, um padrão é construído gerando um seno ou uma varredura, de acordo com as Equações 5 e 6. Então, o sinal gerado é transformado em domínio de frequência de ODFT (a) para obter um espectro complexo (b). Depois, o padrão complexo é truncado no seu comprimento (c) pretendido e armazenado em uma tabela.

[0268] Sempre que o padrão é necessário para sintetizar uma parte de sinal tonal, ele é adaptado à sua fase alvo, conforme descrito acima, com relação à interligação de padrões, e adicionalmente é compensado para a

rotação de fase induzida pela mudança espectral, conforme descrito acima com relação à compensação para a mudança espectral da posição do número inteiro (d). Ainda, a mudança de tempo que está presente na CMDCT com relação à ODFT é implementada aplicando um pós-cruzamento, conforme descrito acima. A aplicação do pós-cruzamento pode ser feita de maneira eficiente após o somar a contribuição de todos os padrões a serem corrigidos no espectro (e). Finalmente, a correção real acontece no domínio da MDCT utilizando somente a parte real do padrão adaptado. Uma MDCT produz o sinal de domínio de tempo desejado, cujo espectro está descrito no painel (f).

[0269] A Figura 4 ilustra padrões de tom espectral normalizados de acordo com uma aplicação, em particular, seno sobre a posição, seno entre posições, varredura sobre a posição, varredura entre posições (do painel superior ao inferior). Mais especificamente, a Figura 4 exemplarmente descreve uma seleção de diferentes padrões de tom para um cenário típico de codec de transformação da taxa de bit baixa utilizando uma MDCT de banda 512, com janela sinusoidal, em uma frequência de amostragem de 12,8 kHz, e uma resolução de meia posição para os padrões de tom. Da parte superior à parte inferior do painel, vários padrões de tom da ODFT espectrais normalizados são traçados: seno sobre a posição, seno entre posições, varredura sobre a posição e varredura entre posições. Vários padrões como esses devem ser armazenados em uma tabela.

[0270] Todos os tipos de padrão estão armazenados em

4 variantes:

- sobre a posição e entre posições
- fase inicial 0 e fase inicial π / 4 (pré-girada, conforme descrito acima com relação à interligação acima dos padrões).

[0271] Padrões de varredura tem 6 variantes adicionais:

- meia, complete e uma e meia varredura da posição
- direção de varredura para cima e para baixo

[0272]

[0273] O número total de padrões a serem armazenados é 4 vezes (1 estacionário + 6 varreduras) e quantidades de até 28 padrões complexos.

[0274] Para as menores exigências de memória possíveis, os padrões de varredura podem alternativamente ser armazenados somente em uma direção; a direção oposta pode ser derivada por um processamento espectral que é duplo para o paralelismo temporal do padrão. Tipicamente, isso pode ser obtido em um domínio de frequência pelo cálculo computacional do conjugado complexo (pela multiplicação da parte imaginária por -1) do padrão e aplicando um termo da fase complexa (cruzado) que depende do domínio real (ODFT, CMDCT, etc.).

[0275] A qualidade do sinal que pode ser obtida sintetizando padrões espectrais truncados depende do tipo de janela, que geralmente já é determinado pelo codec de transformação, e sobre a escolha real do comprimento padrão, que pode ser adaptado à qualidade perceptual geral do codec e

aos recursos disponíveis (memória, complexidade de cálculo computacional).

[0276] A Figura 5 ilustra um índice do sinal para ruído (SNR) de um padrão de tom truncado como uma função de comprimento padrão para uma janela sinusoidal. Em particular, a Figura 5 mostra o SNR médio como uma função de comprimento padrão para a janela sinusoidal. No cenário descrito com relação à Figura 3, truncar os padrões para, por exemplo, 19 posições produz uma média de SNR de aproximadamente 65 dB. Se um SNR mais baixo é aceitável, por exemplo, em um codec de bit muito baixo, já um comprimento padrão de 5 posições pode ser suficiente.

[0277] A figura 6a descreve uma variação da ilustração da figura 2, em que a figura 6a ilustra uma frequência instantânea em pontos no tempo para blocos de sobreposição, de acordo com as aplicações.

[0278] A figura 6b ilustra um progresso de fase para funções de base de DCT e DCT IV, de acordo com as aplicações com relação ao diagrama fornecido pela figura 6a.

[0279] A figura 6c ilustra um espectro de potência 670, um espectro de MDCT substituído 675, um espectro de MDCT quantizado 680 e um espectro de MDCT com padrões 685, de acordo com uma aplicação.

[0280] O espectro de MDCT quantizado 680 foi gerado em um lado do codificador através da quantização do espectro de MDCT substituído 675. O espectro de MDCT substituído 675 foi gerado com base em um espectro de entrada do sinal de áudio (não mostrado), conforme descrito a partir do

codificador acima e com base em um espectro de potência 670.

[0281] O espectro de MDCT quantizado 680 será obtido sobre um lado do decodificador processando um espectro do sinal de áudio codificado (não mostrado) para obter o espectro de MDCT quantizado 680 como um espectro do sinal de áudio decodificado.

[0282] Como pode ser visto na figura 6c, os pseudocoeficientes 691, 692, 693, 694, 695 e 696 no espectro do sinal de áudio decodificado 680 são substituídos pelos padrões espectrais 651, 652, 653, 654, 655 e 656, respectivamente.

[0283] Para o mesmo cenário de codec de taxa de bit baixa, conforme acima, a complexidade de cálculo computacional da síntese do padrão do tom recém-proposta foi comparada com a complexidade de cálculo computacional de um banco de osciladores plano no domínio de tempo. Foi assumido que um máximo de 20 faixas sinusoidais estão ativas durante a codificação de um item monofônico em uma configuração de codec perceptual completo em uma taxa de bit bastante baixa de 13,2 kbps. A carga de trabalho do cálculo computacional foi medida na implementação C do codec. Os itens utilizados para as medições continham pelo menos um instrumento tonal dominante cada um com conteúdo harmônico rico (por exemplo, diapasão, violino, cravo, pop saxofone, ensemble de metais). Em média, a complexidade de cálculo computacional da síntese baseada no padrão de tom é somente 10% da implementação simples utilizando um banco de osciladores no domínio de tempo.

[0284] As aplicações descritas acima fornecem conceitos para melhorar os codificadores de áudio baseados em taxas de bit de MDCT baixas através da geração de sinusoides paramétricos e varredura de seno. Aplicando os conceitos fornecidos, tais sinais podem ser gerados de maneira muito eficiente no decodificador utilizando padrões de tom que são adaptados pelas rotações da fase de pós-processamento. Para a síntese real desses padrões de tom, o banco de filtro da IMDCT do codificador pode ser coutilizado. Conforme descrito acima, a escolha inicial da resolução espectral determina uma frequência de corte inferior para a geração de tom perceptualmente apropriada, a demanda de memória de armazenamento e a complexidade de cálculo computacional do pós-processamento padrão exigido. Em um cenário exemplar de codec de áudio de taxa de bit baixa, uma redução da complexidade de cálculo computacional de 90% em um SNR de 65 dB foi alcançada comparada à implementação de um banco de osciladores de domínio de tempo.

[0285] Enquanto uma solução empregaria um banco de osciladores no domínio de tempo em uma taxa de amostra completa, tal solução permitiria uma interpolação suave entre parâmetros subsequentes. Entretanto, essa solução é computacionalmente pesada.

[0286] É vantajoso para uma complexidade de cálculo computacionalmente baixa empregar padrões espectrais ToneFilling (TF) de MDCT. Aqui, o espectro pode ser corrigido com padrões TF em uma taxa de amostra de bloco. Padrões espectrais truncados podem ser armazenados, por exemplo, em

uma tabela, por exemplo, uma tabela de uma base de dados ou de uma memória.

[0287] Nas aplicações, uma "interpolação" de faixas sinusoidais de uma amplitude por uma janela de síntese de sobreposição de 50% e de uma frequência por escolha de padrões de varredura com inclinação apropriada é fornecida, que é computacionalmente muito eficiente.

[0288] Aplicações fornecem um desenho do padrão de domínio de tempo para alinhamento mínimo. A fase e a frequência instantânea (IF) se equiparam em pontos no momento onde os blocos de sobreposição tem ponderações iguais.

[0289] Como pode ser visto na figura 6a, pontos de simetria estão localizados em:

[0290] n_0: ¼*b_comprimento+0.5; e

[0291] n_1: ¾* b_comprimento+0.5.

[0292] Para adequar continuamente uma faixa sinusoidal, de acordo com uma aplicação, padrões são escolhidos a partir do padrão da posição de número inteiro ("sobre a posição"), padrão de posição de número fracionário ("entre posições") e varreduras lineares: meia, completa e uma e meia varredura de posição.

[0293] Os padrões escolhidos são adaptados para a localização pretendida na grade t/f de MDCT conduzindo o dimensionamento de amplitude, e, com relação à fase, conduzindo uma rotação complexa (cruzado) como uma função de localização de fonte padrão, localização alvo, fase do predecessor temporal.

[0294] Devido à resolução de frequência limitada,

somente um conjunto discreto de rotações predefinidas é necessário, em particular:

- N*π/2 rotações através da permutação da parte real e imaginária e do sinal; e
- N* π/4 rotações implementadas por padrões pré-girados π/4.

[0295] Implementar uma mudança de tempo de MDCT exige padrões/correção no domínio da ODFT. Uma resolução de meia posição é realizada por uma granularidade de fase de π/2, e dois tipos de padrão diferentes.

[0296] Uma mudança de frequência da ODFT/DCT-IV é realizada por uma fase de progresso de padrões de posição de número inteiro por +π/2 ou -π/2, por uma fase de progresso de posição de número fracionário por 0 ou π, e é dependente de uma paridade de número de posição (par/ímpar). Isso está ilustrado na figura 6b.

[0297] Nas aplicações, todos os padrões são armazenados em 4 variantes, abrangendo as combinações das alternativas:

- posição de número inteiro e posição de número fracionário;
- $\varphi = 0$ ou $\varphi = \pi/4$ (pré-girados, necessário para movimentar varreduras de meia posição).

[0298] Nas aplicações, padrões de varredura têm 6 variantes adicionais abrangendo as combinações das alternativas;

- meia, completa ou uma varredura de posição e meia; e

- para cima ou para baixo

[0299] Isso resulta em um número total de: 4* (1 estacionário + 6 varreduras) = 28 padrões complexos. A correção real é a parte real do padrão final (girado).

[0300] Os conceitos fornecidos podem, por exemplo, ser empregados para USAC, em particular no caminho do sinal de codificação de transformação.

[0301] Resumindo o que está acima, a MDCT é crítica para sinais tonais de codificação em taxas de bit baixas devido à ocorrência de perturbações melódicas. O modelo psicoacústico clássico, entretanto, não é responsável por isso. Assim, um modelo menos desagradável é necessário. Instrumentos de codificação paramétrica podem ajudar nas taxas de bit baixas. Perturbações de ToneFilling podem ser menos desagradáveis do que as melódicas.

[0302] A implementação eficiente de osciladores de ToneFilling pode ser alcançada através da correção de padrões de MDCT adaptados t/f. Empregando ToneFilling, a qualidade decente em taxa de bit baixa e na codificação de atraso baixo de música tonal é obtida.

[0303] A seguir, uma descrição com relação a algumas outras aplicações é fornecida.

[0304] A figura 10 ilustra um aparelho para gerar um sinal de saída de áudio com base em um espectro do sinal de áudio codificado.

[0305] O aparelho compreende uma unidade de processamento 110 para processar o espectro do sinal de áudio codificado para obter um espectro do sinal de áudio

decodificado. O espectro do sinal de áudio decodificado compreende uma pluralidade de coeficientes espectrais, em que cada um dos coeficientes espectrais tem uma localização espectral dentro do espectro do sinal de áudio codificado e um valor espectral, em que os coeficientes espectrais são sequencialmente organizados, de acordo com sua localização espectral dentro do espectro do sinal de áudio codificado de modo que os coeficientes espectrais formam uma sequência de coeficientes espectrais.

[0306] Além disso, o aparelho compreende um determinador dos pseudocoeficientes 120 para determinar um ou mais pseudocoeficiente(s) do espectro do sinal de áudio decodificado utilizando a informação adicional (informação adicional), cada um dos pseudocoeficientes tendo uma localização espectral e um valor espectral.

[0307] Além disso, o aparelho compreende uma unidade de modificação de espectro 130 para definir um ou mais pseudocoeficiente(s) para um valor predefinido para obter um espectro do sinal de áudio modificado.

[0308] Além disso, o aparelho compreende uma unidade de conversão de tempo do espectro 140 para converter o espectro do sinal de áudio modificado para um domínio de tempo para obter um sinal de conversão de domínio de tempo.

[0309] Além disso, o aparelho compreende um oscilador controlável 150 para gerar um sinal oscilador de domínio de tempo, o oscilador controlável sendo controlado pela localização espectral e pelo valor espectral de pelo menos um de um ou mais pseudocoeficiente(s).

[0310] Além disso, o aparelho compreende um misturador 160 para misturar o sinal de conversão de domínio de tempo e o sinal oscilador de domínio de tempo para obter o sinal de saída de áudio.

[0311] Em uma aplicação, o misturador pode ser configurado para misturar o sinal de conversão de domínio de tempo e o sinal oscilador de domínio de tempo adicionando o sinal de conversão de domínio de tempo ao sinal oscilador de domínio de tempo no domínio de tempo.

[0312] A unidade de processamento 110 pode, por exemplo, ser qualquer tipo de decodificador de áudio, por exemplo, um decodificador de áudio de MP3, um decodificador de áudio para WMA, um decodificador de áudio para arquivos em WAVE, um decodificador de áudio de AAC ou um decodificador de áudio de USAC.

[0313] A unidade de processamento 110 pode, por exemplo, ser um decodificador de áudio conforme descrito em [8] (ISO/IEC 14496-3:2005 - Tecnologia de Informação - Codificação de objetos de audiovisual - Parte 3: Áudio, Subparte 4) ou conforme descrito em [9] (ISO/IEC 14496-3:2005 - Tecnologia de Informação - Codificação de objetos de audiovisual - Parte 3: Áudio, Subparte 4). Por exemplo, a unidade de processamento 430 pode compreender um redimensionamento de valores quantizados ("de-quantização"), e/ou ferramenta de modelagem de ruído temporal, conforme descrito, por exemplo, em [8] e/ou a unidade de processamento 430 pode compreender uma ferramenta de substituição de ruído perceptual como descrito, por exemplo, em [8].

[0314] De acordo com uma aplicação, cada um dos coeficientes espectrais pode ter pelo menos um de um precursor imediato e de um sucessor imediato, em que o precursor imediato do referido coeficiente espectral pode ser um dos coeficientes espectrais que imediatamente precede o referido coeficiente espectral dentro da sequência, em que o sucessor imediato do referido coeficiente espectral pode ser um dos coeficientes espectrais que imediatamente sucedem o referido coeficiente espectral dentro da sequência.

[0315] O determinador dos pseudocoeficientes 120 pode ser configurado para determinar um ou mais pseudocoeficiente(s) do espectro do sinal de áudio decodificado determinando pelo menos um coeficiente espectral da sequência, que tem um valor espectral que é diferente do valor predefinido, que tem um precursor imediato do valor espectral que é igual ao valor predefinido, e que tem um sucessor imediato ao valor espectral que é igual ao valor predefinido. Em uma aplicação, o valor predefinido pode ser zero e o valor predefinido pode ser zero.

[0316] Em outras palavras: O determinador dos pseudocoeficientes 120 determina para alguns ou todos os coeficientes do espectro do sinal de áudio decodificado se o coeficiente respectivamente considerado é diferente do valor predefinido (preferivelmente: diferente de 0), se o valor espectral do coeficiente precedente for igual ao valor predefinido (preferivelmente: igual a 0) e se o valor espectral do coeficiente sucessor for igual ao valor predefinido (preferivelmente: igual a 0).

[0317] Em algumas aplicações, tal coeficiente determinado é (sempre) um pseudocoeficiente.

[0318] Em outras aplicações, entretanto, tal coeficiente determinado é (somente) um pseudocoeficiente candidato e pode ou não ser um pseudocoeficiente. Nessas aplicações, o determinador dos pseudocoeficientes 120 é configurado para determinar pelo menos um pseudocoeficiente candidato, que tem um valor espectral que é diferente do valor predefinido, que tem um precursor imediato, cujo valor espectral é igual ao valor predefinido, e que pode ter um sucessor imediato, cujo valor espectral é igual ao valor predefinido.

[0319] O determinador dos pseudocoeficientes 120 é então configurado para determinar se o pseudocoeficiente candidato é um pseudocoeficiente determinando se a informação adicional indica que o referido pseudocoeficiente candidato é um pseudocoeficiente.

[0320] Por exemplo, tal informação adicional pode ser recebida pelo determinador dos pseudocoeficientes 120 em um campo de bit, que indica para cada um dos coeficientes espectrais do espectro do sinal de áudio quantizado que tem um precursor imediato cujo valor espectral é igual ao valor predefinido e um sucessor imediato, cujo valor espectral é igual ao valor predefinido, se o referido coeficiente for um dos coeficientes extremos (por exemplo, por um valor de bit 1) ou se o referido coeficiente não é um dos coeficientes extremos (por exemplo, para um valor de bit 0).

[0321] Por exemplo, um campo de bit [000111111] pode

indicar que os primeiros três coeficientes "independentes" (seus valores espectrais não são iguais ao valor predefinido, mas os valores espectrais dos seus precursores e dos seus sucessores são iguais ao valor predefinido) que aparece (sequencialmente ordenado) no espectro (quantizado) do sinal de áudio não são coeficientes extremos, mas os próximos seis coeficientes "independentes" são coeficientes extremos. Esse campo de bit descreve a situação que pode ser vista no espectro MDCT quantizado 635 na figura 9, onde os primeiros três coeficientes "independentes" 5, 8, 25 não são coeficientes extremos, mas onde os próximos seis coeficientes "independentes" 59, 71, 83, 94, 116, 141 são coeficientes extremos.

[0322] A unidade de modificação de espectro 130 pode ser configurada para "apagar" os pseudocoeficientes do espectro do sinal de áudio decodificado. De fato, a unidade de modificação de espectro define o valor espectral dos pseudocoeficientes do espectro do sinal de áudio decodificado ao valor predefinido (preferivelmente a 0). Isso é razoável, visto que (pelo menos um) pseudocoeficientes somente serão necessários para controlar o oscilador controlável (pelo menos um) 150. Assim, considere, por exemplo, o espectro MDCT quantizado 635 na figura 9. Se o espectro 635 for considerado como o espectro do sinal de áudio decodificado, a unidade de modificação de espectro 130 definiria os valores espectrais dos coeficientes extremos 59, 71, 83, 94, 116 e 141 para obter o espectro do sinal de áudio modificado e deixaria os outros coeficientes do espectro não modificados.

[0323] A unidade de conversão de tempo do espectro 140 converte o espectro do sinal de áudio modificado de um domínio espectral para um domínio de tempo. Por exemplo, o espectro do sinal de áudio modificado pode ser um espectro de MDCT, e a unidade de conversão de tempo do espectro 140 pode ser um banco de filtro da Transformada de Cosseno Discreta Modificada Inversa (IMDCT). Em outras aplicações, o espectro pode ser um espectro da MDST e a unidade de conversão de tempo do espectro 140 pode ser um banco de filtro da Transformada de Seno Discreta Modificada Inversa (IMDST. Ou, em outras aplicações, o espectro pode ser um espectro da DFT e a unidade de conversão de tempo do espectro 140 pode ser um banco de filtro da Transformada de Fourier Discreta (IDFT).

[0324] O oscilador controlável 150 pode ser configurado para gerar o sinal oscilador de domínio de tempo tendo uma frequência do sinal oscilador de modo que a frequência do sinal oscilador do sinal oscilador possa depender da localização espectral de um de um ou mais pseudocoeficiente(s). O sinal oscilador gerado pelo oscilador pode ser um sinal de seno de domínio de tempo. O oscilado controlável 150 pode ser configurado para controlar a amplitude do sinal de seno de domínio de tempo dependendo do valor espectral de um de um ou mais pseudocoeficiente(s).

[0325] De acordo com uma aplicação, os pseudocoeficientes são valores com sinal, cada um compreendendo um componente do sinal. O oscilador controlável 150 pode ser configurado para gerar o sinal oscilador de domínio de tempo de modo que a frequência do sinal oscilador,

além disso, possa depender do componente do sinal de um de um ou mais pseudocoeficiente(s) de modo que a frequência do sinal oscilador possa ter um primeiro valor de frequência, quando o componente do sinal tem um primeiro valor do sinal, e de modo que a frequência do sinal oscilador possa ter um segundo valor de frequência de diferente, quando o componente do sinal tem um segundo valor diferente.

[0326] Por exemplo, considere o pseudocoeficiente na localização espectral 59 no espectro de MDCT 635 da figura 9. Se a frequência de 8200 Hz seria atribuída à localização espectral 59 e se a frequência 8400 Hz seria atribuída à localização espectral 60, então, o oscilador controlável pode, por exemplo, ser configurado para definir a frequência do oscilador a 8200 Hz, se o sinal do valor espectral do pseudocoeficiente for positivo, e pode, por exemplo, ser configurado para definir a frequência do oscilador a 8300 Hz, se o sinal do valor espectral do pseudocoeficiente for negativo.

[0327] Assim, o sinal do valor espectral do pseudocoeficiente pode ser utilizado para controlar, se oscilador controlável definir a frequência do oscilador a uma frequência (por exemplo, 8200 Hz) atribuída à localização espectral derivada do pseudocoeficiente (por exemplo, localização espectral 59) ou a uma frequência (por exemplo, 8300 Hz) entre a frequência (por exemplo, 8200 Hz) atribuída à localização espectral derivada do pseudocoeficiente (por exemplo, localização espectral 59) e a frequência (por exemplo, 8400 Hz) atribuída à localização espectral que

imediatamente segue a localização espectral derivada do pseudocoeficiente (por exemplo, localização espectral 60).

[0328] A figura 11 ilustra uma aplicação, em que o aparelho compreende ainda osciladores controláveis 252, 254, 256 para gerar outros sinais osciladores de domínio de tempo controlados pelos valores espectrais de outros pseudocoeficientes de um ou mais pseudocoeficiente(s).

[0329] Cada um dos outros osciladores controláveis 252, 254, 256 gera um dos outros sinais osciladores de domínio de tempo. Cada um dos osciladores controláveis 252, 254, 256 é configurado para conduzir a frequência do sinal oscilador com base na localização espectral derivada de um dos pseudocoeficientes. E/ou cada um dos osciladores controláveis 252, 254, 256 é configurado para conduzir a amplitude do sinal oscilador com base no valor espectral de um dos pseudocoeficientes.

[0330] Cada um dos outros osciladores controláveis 252, 254, 256 gera um dos outros sinais osciladores de domínio de tempo. Cada um dos osciladores controláveis 252, 254, 256 é configurado para conduzir a frequência do sinal oscilador com base na localização espectral de um dos pseudocoeficientes. E/ou cada um dos osciladores controláveis 252, 254, 256 é configurado para conduzir a amplitude do sinal oscilador com base no valor espectral de um dos pseudocoeficientes.

[0331] O misturador 160 da figura 10 e da figura 11 é configurado para misturar o sinal de conversão de domínio de tempo gerado pela unidade de conversão de tempo do

espectro 140 e um ou mais sinais osciladores de domínio de tempo gerados por um ou mais osciladores controláveis 150, 252, 254, 256 para obter o sinal de saída de áudio. O misturador 160 pode gerar o sinal de saída de áudio por uma superposição do sinal de conversão de domínio de tempo e um ou mais sinais osciladores de domínio de tempo.

[0332] A figura 12 ilustra dois diagramas comparando sinusoides originais (esquerda) e sinusoides depois de processados por uma interligação MDCT/IMDCT (direita). Depois de ser processado pela interligação MDCT/ IMDCT, o sinusoide compreende perturbações melódicas. Os conceitos fornecidos acima evitam que as sinusoides sejam processadas pela interligação MDCT/IMDCT, mas ao invés disso, a informação sinusoidal é codificada por um pseudocoeficiente e/ou o sinusoide é reproduzido por um oscilador controlável.

[0333] Embora alguns aspectos tenham sido descritos no contexto de um aparelho, é evidente que estes aspectos também representam uma descrição do método correspondente, onde um bloco ou dispositivo corresponde a uma etapa do método ou uma característica de uma etapa do método. De forma análoga, os aspectos descritos no contexto de uma etapa do método também representam uma descrição de um bloco correspondente ou item ou característica de um aparelho correspondente.

[0334] O sinal inventivo decomposto pode ser armazenado em um meio de armazenamento digital ou pode ser transmitido por um meio de transmissão, tal como um meio de transmissão sem fio ou um meio de transmissão cabeado, tal

como a Internet.

[0335] Dependendo de certas exigências de implementação, as aplicações da invenção podem ser implementadas em hardware ou em software. A implementação pode ser realizada utilizando um meio de armazenamento digital, por exemplo, um disquete, um DVD, um CD, uma memória ROM, uma PROM, uma EPROM, uma EEPROM ou uma memória FLASH, tendo sinais de controle eletronicamente legíveis armazenados nele (ou são capazes de cooperar) com um sistema de computador programável, de modo que o respectivo método seja realizado.

[0336] Algumas aplicações, de acordo com a invenção, compreendem um transportador de dados não transitório tendo sinais de controle eletronicamente legíveis que são capazes de cooperar com um sistema de computador programável, de modo que um dos métodos descritos neste documento seja realizado.

[0337] De forma geral, as aplicações da presente invenção podem ser implementadas como um produto do programa de computador com um código do programa, o código do programa sendo operativo para realizar um dos métodos quando o produto do programa de computador é executado em um computador. O código do programa pode, por exemplo, ser armazenado em um transportador legível por máquina.

[0338] Outras aplicações compreendem o programa de computador para realizar um dos métodos descritos neste documento, armazenados em um transportador legível por máquina.

[0339] Em outras palavras, uma aplicação do método

inventivo é, assim, um programa de computador tendo um código do programa para realizar um dos métodos descritos neste documento, quando o programa de computador é executado em um computador.

[0340] Outra aplicação dos métodos inventivos é, portanto, um transportador de dados (ou um meio de armazenamento digital, ou um meio legível por computador) compreendendo, gravado nele, o programa de computador para realizar um dos métodos descritos neste documento.

[0341] Outra aplicação do método inventivo é, portanto, um fluxo de dados ou uma sequência de sinais que representam o programa de computador para realizar um dos métodos descritos neste documento. O fluxo de dados ou a sequência de sinais pode, por exemplo, ser configurado para ser transferido através de uma conexão de comunicação de dados, por exemplo, através da Internet.

[0342] Outra aplicação compreende um meio de processamento, por exemplo, um computador ou um dispositivo lógico programável, configurado para ou adaptado para realizar um dos métodos descritos neste documento.

[0343] Outra aplicação compreende um computador tendo instalado nele o programa de computador para realizar um dos métodos descritos neste documento.

[0344] Em algumas aplicações, um dispositivo lógico programável (por exemplo, um arranjo de portas de campo programáveis) pode ser utilizado para realizar algumas ou todas as funcionalidades dos métodos descritos neste documento. Em algumas aplicações, um arranjo de portas de

campo programáveis pode cooperar com um microprocessador, a fim de realizar um dos métodos descritos neste documento. De forma geral, os métodos são preferivelmente realizados por qualquer aparelho de hardware.

[0345] As aplicações descritas acima são meramente ilustrativas para os princípios da presente invenção. Entende-se que modificações e variações das disposições e os detalhes descritos no presente documento serão evidentes a outros especialistas na técnica. É intenção da invenção, portanto, ser limitada apenas pelo escopo das reivindicações da patente anexas e não pelos detalhes específicos apresentados em forma de descrição e explicação das aplicações no presente documento

REFERÊNCIAS:

[0346] [1] Daudet, L.; Sandler, M.; , "MDCT analysis of sinusoids: exact results and applications to coding artifacts reduction," Speech and Audio Processing, IEEE Transactions on, vol.12, no.3, pp. 302-312, May 2004

[0347] [2] Purnhagen, H.; Meine, N.;, "HILN-the MPEG-4 parametric audio coding tools," *Circuits and Systems, 2000. Proceedings. ISCAS 2000 Geneva. The 2000 IEEE International Symposium an,* vol.3, no., pp.201-204 vol.3, 2000

[0348] [3] Oomen, Werner; Schuijers, Erik; den Brinker, Bert; Breebaart, Jeroen:," Advances in Parametrie Coding for High-Quality Audio," Audio Engineering Society Convention 114, preprint, Amsterdam/NL, March 2003

[0349] [4] van Schijndel, N.H. ; van de Par, S.; ,

"Rate-distortion optimized hybrid sound coding," Applications of Signal Processing to Audio and Acoustics, 2005. IEEE Workshop on, vol., no., pp. 235-238, 16-19 Oct. 2005

[0350] [5] Bessette, 8.; Lefebvre, R.; Salami, R. ; , "Universal speech/audio coding using hybrid ACELP/TCX techniques," Acoustics, Speech, and Signal Processing, 2005. Proceedings. (ICASSP '05). IEEE International Conference on, vol.3, no., pp. iii/301- iii/304 Val. 3, 18-23 March 2005

[0351] [6] Ferreira, A.J.S. "Combined spectral envelope normalization and subtraction of sinusoidal components in the ODFT and MDCT frequency domains," Applications of Signal Processing to Audio and Acoustics, 2001 IEEE Workshop on the, vol., no., pp. 51-54

Accurate spectral replacement. In *Audio Engineering Society Convention JJ 8,* 5 2005.

[0357] [12] Rade Kutil. Optimized sinusoid synthesis via inverse truncated fourier transforrn. *Trans. Audio. Speech and Lang. Proc.,* 17(2):221-230, February 2009.

[0358] [13] Nikolaus Meine and Heiko Purnhagen. Fast sinusoid synthesis for mpeg-4 hiln parametric audio decoding. *Proc. of the 5 th Int. Conference on Digital Audio Effects (DAFx-02), Hamburg, Germany, September 26-28, 2002,* 0(0), 2002.

REIVINDICAÇÕES

1. Um aparelho para gerar um sinal de saída de áudio com base em um espectro do sinal de áudio codificado, caracterizado pelo aparelho compreender:

uma unidade de processamento para processar o espectro de sinal de áudio codificado para adquirir um espectro de sinal de áudio decodificado compreendendo uma pluralidade de coeficientes espectrais, em que cada um dos coeficientes espectrais compreende uma localização espectral dentro do espectro de sinal de áudio codificado e um valor espectral, em que os coeficientes espectrais são ordenados sequencialmente de acordo com sua localização espectral dentro do espectro de sinal de áudio codificado de modo que os coeficientes espectrais formem uma sequência de coeficientes espectrais,

um determinante de pseudo coeficientes para determinar um ou mais pseudo coeficientes do espectro de sinal de áudio decodificado, em que cada um dos pseudo coeficientes é um dos coeficientes espectrais,

uma unidade de substituição para substituir pelo menos um ou mais pseudo coeficientes por um determinado padrão espectral para adquirir um espectro de sinal de áudio modificado, em que o padrão espectral determinado compreende pelo menos dois coeficientes padrão, em que cada um dos pelo menos dois coeficientes padrão compreende um valor espectral , e

uma unidade de conversão de espectro no tempo para converter o espectro do sinal de áudio modificado em um

domínio do tempo para adquirir o sinal de saída de áudio.

2. Aparelho, de acordo com a reivindicação 1, em que o aparelho é caracterizado por compreender, além disso, uma unidade de armazenamento que compreende um banco de dados ou uma memória tendo armazenado dentro do banco de dados ou dentro da memória uma pluralidade de padrões espectrais armazenados, em que cada um dos padrões espectrais armazenados compreende uma propriedade espectral,

em que a unidade de substituição é configurada para solicitar um dos padrões espectrais armazenados da unidade de armazenamento como um padrão espectral solicitado,

em que a unidade de armazenamento é configurada para fornecer o padrão espectral solicitado, e

em que a unidade de substituição está configurada para substituir o pelo menos um ou mais pseudo coeficientes pelo padrão espectral determinado com base no padrão espectral solicitado.

3. Aparelho, de acordo com a reivindicação 2, caracterizado pelo fato de que a unidade de substituição é configurada para solicitar o referido um dos padrões espectrais armazenados da unidade de armazenamento dependendo de uma primeira localização espectral derivada derivada de pelo menos um de um ou mais pseudo coeficientes determinados pelo determinador de pseudo coeficientes.

4. Aparelho, de acordo com a reivindicação 3, caracterizado em que um ou mais pseudo coeficientes são valores com sinais, cada um compreendendo um componente de

sinal, e em que a unidade de substituição é configurada para determinar a primeira localização espectral derivada com base na localização espectral de um pseudo coeficiente de um ou mais pseudo coeficientes e com base no componente de sinal do referido pseudo coeficiente, de modo que a primeira localização espectral derivada seja igual a a localização espectral do referido pseudo coeficiente quando o componente de sinal compreende um primeiro valor de sinal, e de modo que a primeira localização espectral derivada seja igual a uma localização modificada, a localização modificada resultante do deslocamento da localização espectral do referido pseudo coeficiente por um valor predefinido quando o componente do sinal compreende um segundo valor diferente.

5. Aparelho, de acordo com a reivindicação 3, caracterizado em que a pluralidade de padrões espectrais armazenados sendo armazenados no banco de dados ou na memória da unidade de armazenamento são padrões de tom estacionários ou padrões de varredura de frequência, em que o determinante de pseudo coeficientes é configurado para determinar dois ou mais pseudo coeficientes temporariamente consecutivos do espectro de sinal de áudio decodificado, em que a unidade de substituição é configurada para atribuir um primeiro pseudo-coeficiente e um segundo pseudo-coeficiente dos dois ou mais pseudo-coeficientes temporalmente consecutivos a uma pista dependendo se uma diferença absoluta entre a primeira localização espectral derivada derivada do primeiro pseudo-coeficiente e um segundo localização espectral derivada derivada do segundo pseudo coeficiente é menor do que um

valor limite, e em que a unidade de substituição está configurada para solicitar um dos padrões de tom estacionários da unidade de armazenamento quando a primeira localização espectral derivada do primeiro pseudo-coeficiente da pista é igual à segunda localização espectral derivada do segundo pseudo-coeficiente da pista , e em que a unidade de substituição é configurada para solicitar um dos padrões de varredura de frequência da unidade de armazenamento quando a primeira localização espectral derivada do primeiro pseudo coeficiente da pista é diferente da segunda localização espectral derivada derivada do segundo pseudo coeficiente de a pista.

6. Aparelho, de acordo com a reivindicação 5, caracterizado em que a unidade de substituição é configurada para solicitar um primeiro padrão de varredura de frequência dos padrões de varredura de frequência da unidade de armazenamento quando uma diferença de frequência entre a segunda localização espectral derivada do segundo pseudo coeficiente da pista e a primeira localização espectral derivada derivada do o primeiro pseudo coeficiente da trilha é igual a metade de um valor predefinido, em que a unidade de substituição é configurada para solicitar um segundo padrão de varredura de frequência, sendo diferente do primeiro padrão de varredura de frequência, dos padrões de varredura de frequência da unidade de armazenamento quando a diferença de frequência entre a segunda localização espectral derivada derivada do segundo pseudo coeficiente do pista e a primeira localização espectral derivada derivada do primeiro pseudo

coeficiente da pista é igual ao valor predefinido, e em que a unidade de substituição está configurada para solicitar um terceiro padrão de varredura de frequência, sendo diferente do primeiro padrão de varredura de frequência e do segundo padrão de varredura de frequência, dos padrões de varredura de frequência da unidade de armazenamento quando a diferença de frequência entre a segunda localização espectral derivada derivada do o segundo pseudo coeficiente da pista e a primeira localização espectral derivada derivada do primeiro pseudo coeficiente da pista é igual a uma vez e meia o valor predefinido.

7. Aparelho, de acordo com a reivindicação 2, caracterizado pelo fato de que a unidade de substituição compreende uma unidade de adaptação de padrão sendo configurada para modificar o padrão espectral solicitado fornecido pela unidade de armazenamento para adquirir o padrão espectral determinado.

8. Aparelho, de acordo com a reivindicação 7, caracterizado pelo fato de que a unidade de adaptação de padrão é configurada para modificar o padrão espectral solicitado fornecido pela unidade de armazenamento, redimensionando os valores espectrais dos coeficientes de padrão do padrão espectral solicitado dependendo do valor espectral de um dos um ou mais pseudo coeficientes.

9. Aparelho, de acordo com a reivindicação 7, caracterizado pelo fato de que a unidade de adaptação de padrão é configurada para modificar o padrão espectral solicitado fornecido pela unidade de armazenamento dependendo

de uma fase inicial de modo que o valor espectral de cada um dos coeficientes de padrão do padrão espectral solicitado seja modificado de uma primeira forma, quando a fase de início compreende um primeiro valor de fase de início, e de modo que o valor espectral de cada um dos coeficientes de padrão do padrão espectral solicitado é modificado de uma segunda maneira diferente, quando a fase de início compreende um segundo início diferente valor de fase.

10. Aparelho, de acordo com a reivindicação 7, caracterizado em que o valor espectral de cada um dos coeficientes de padrão do padrão espectral solicitado é um coeficiente complexo que compreende uma parte real e uma parte imaginária, e em que a unidade de adaptação de padrão é configurada para modificar o padrão espectral solicitado modificando a parte real e a parte imaginária de cada um dos coeficientes de padrão do padrão espectral solicitado fornecido pela unidade de armazenamento aplicando um fator de rotação complexo $e^{j \cdot \varphi}$, em que φ é um valor de ângulo.

11. Aparelho, de acordo com a reivindicação 7, caracterizado em que o valor espectral de cada um dos coeficientes de padrão do padrão espectral solicitado é um coeficiente complexo que compreende uma parte real e uma parte imaginária, e em que a unidade de adaptação de padrão é configurada para modificar o padrão espectral solicitado fornecido pela unidade de armazenamento negando a parte real e imaginária do valor espectral de cada um dos coeficientes de padrão do padrão espectral solicitado, ou trocando a parte real ou um parte real negada e a parte imaginária ou uma

parte imaginária negada do valor espectral de cada um dos coeficientes de padrão do padrão espectral solicitado.

12. Aparelho, de acordo com a reivindicação 7, caracterizado pelo fato de que a unidade de adaptação de padrão é configurada para modificar o padrão espectral solicitado fornecido pela unidade de armazenamento, realizando um espelhamento temporal do padrão, computando o conjugado complexo do padrão e aplicando um termo de fase complexo.

13. Aparelho, de acordo com a reivindicação 7, caracterizado em que o espectro de sinal de áudio decodificado é representado em um domínio MDCT, em que a unidade de adaptação de padrão é configurada para modificar o padrão espectral solicitado fornecido pela unidade de armazenamento, modificando os valores espectrais dos coeficientes de padrão do padrão espectral solicitado para adquirir um padrão espectral modificado, em que os valores espectrais são representados em um domínio de Transformada Discreta de Fourier com Empilhamento Diferencial, em que a unidade de adaptação de padrão é configurada para transformar os valores espectrais dos coeficientes de padrão do padrão espectral modificado do domínio da Transformação de Fourier Discreta Estranhamente Empilhada para o domínio de MDCT para adquirir o padrão espectral determinado, e em que a unidade de substituição é configurada para substituir o pelo menos um ou mais pseudo coeficientes pelo padrão espectral determinado sendo representado no domínio MDCT para adquirir o espectro de sinal de áudio modificado sendo representado no domínio

MDCT.

14. Um aparelho para gerar uma pluralidade de padrões espectrais, compreendendo:

um gerador de sinal (165) para gerar uma pluralidade de sinais em um primeiro domínio,

uma unidade de transformação do sinal (175) para transformar cada sinal da pluralidade de sinais a partir do primeiro domínio em um segundo domínio para obter uma pluralidade de padrões espectrais, cada padrão da pluralidade de padrões espectrais transformados, compreendendo uma pluralidade de coeficientes,

uma unidade de pós-processamento (185) para truncar os padrões espectrais transformados pela remoção de um ou mais dos coeficientes dos padrões espectrais transformados para obter uma pluralidade de padrões processados, e

uma unidade de armazenamento (195), compreendendo uma base de dados ou uma memória, em que a unidade de armazenamento (195) é configurada para armazenar cada padrão processado da pluralidade de padrões processados na base de dados ou na memória, caracterizado pelo gerador de sinal (165) ser configurado para gerar cada sinal da pluralidade de sinais com base nas fórmulas

$$x(t) = cos\,(2\pi\varphi(t))$$

e

$$\varphi(t) = \varphi(0) + \int_0^t 2\pi f(\tau)d\tau \quad ,$$

em que t e τ indicam tempo,

em que $\varphi(t)$ é uma fase instantânea em t, e

em que $f(\tau)$ é uma frequência instantânea em τ,

em que cada sinal da pluralidade de sinais tem uma frequência inicial (f_0), sendo uma frequência instantânea do referido sinal em um primeiro ponto no tempo, e uma frequência alvo (f_1), sendo uma frequência instantânea do referido sinal em um diferente segundo ponto no tempo,

em que o gerador de sinal (165) é configurado para gerar um primeiro sinal da pluralidade de sinais, de modo que a frequência alvo do primeiro sinal seja igual à frequência inicial, e

em que o gerador de sinal (165) é configurado para gerar um segundo sinal diferente da pluralidade de sinais, de modo que a frequência alvo do primeiro sinal seja diferente da frequência inicial.

15. Aparelho, de acordo com a reivindicação 14, caracterizado em que a unidade de transformação de sinal é configurada para transformar cada sinal da pluralidade de sinais do primeiro domínio, sendo um domínio do tempo, para um segundo domínio, sendo um domínio espectral, em que a unidade de transformação de sinal é configurada para gerar um primeiro de uma pluralidade de blocos de tempo para transformar o referido sinal, em que cada bloco de tempo da pluralidade de blocos de tempo compreende uma pluralidade de amostras ponderadas, em que cada uma das referidas amostras ponderadas é uma amostra de sinal do referido sinal sendo ponderado por um peso de uma pluralidade de pesos, em que a

pluralidade de pesos é atribuída ao referido bloco de tempo, e em que cada peso da pluralidade de pesos é atribuído a um ponto no tempo, em que a frequência inicial de cada sinal da pluralidade de sinais é uma frequência instantânea do referido sinal no primeiro ponto no tempo, onde um primeiro dos pesos do primeiro dos blocos de tempo é atribuído ao primeiro ponto no tempo, em que um segundo dos pesos de um segundo diferente dos blocos de tempo é atribuído ao primeiro ponto no tempo, em que o primeiro dos blocos de tempo e o segundo dos blocos de tempo se sobrepõem, e em que o primeiro dos pesos é igual ao segundo dos pesos, e em que a frequência alvo de cada sinal da pluralidade de sinais é uma frequência instantânea do referido sinal no segundo ponto no tempo, onde um terceiro dos pesos do primeiro dos blocos de tempo é atribuído ao segundo ponto no tempo, onde um quarto dos pesos de um terceiro diferente dos blocos de tempo, é atribuído ao segundo ponto no tempo, em que o primeiro dos blocos de tempo e o terceiro dos blocos de tempo se sobrepõem, e em que o terceiro dos pesos é igual ao quarto dos pesos.

16. Aparelho, de acordo com a reivindicação 14, caracterizado em que cada sinal da pluralidade de sinais compreende uma fase de início, sendo uma fase do referido sinal em um primeiro ponto no tempo, em que o gerador de sinal é configurado para gerar a pluralidade de sinais de modo que a fase inicial de um primeiro da pluralidade de sinais seja igual à fase inicial de um segundo diferente da pluralidade dos sinais.

17. Aparelho, de acordo com a reivindicação

14, caracterizado pelo fato de que a unidade de pós-processamento é, além disso, configurada para conduzir uma rotação por um ângulo de fase arbitrário nos coeficientes espectrais de cada um dos padrões espectrais transformados para adquirir uma pluralidade de padrões espectrais girados arbitrariamente.

18. Aparelho, de acordo com a reivindicação 14, caracterizado pelo fato de que a unidade de pós-processamento é, além disso, configurada para conduzir uma rotação de π / 4 nos coeficientes espectrais de cada um dos padrões espectrais transformados para adquirir uma pluralidade de padrões espectrais girados.

19. Aparelho, de acordo com a reivindicação 14, caracterizado pelo fato de que o gerador de sinal é configurado para gerar o primeiro sinal, o segundo sinal e um ou mais sinais adicionais como a pluralidade de sinais, de modo que cada diferença da frequência alvo e a frequência inicial de cada dos sinais adicionais é um múltiplo inteiro de uma diferença da frequência alvo e a frequência inicial do segundo sinal.

20. Método para gerar um sinal de saída de áudio com base em um espectro de sinal de áudio codificado, caracterizado pelo fato de que o método compreende: processar o espectro de sinal de áudio codificado para adquirir um espectro de sinal de áudio decodificado compreendendo uma pluralidade de coeficientes espectrais, em que cada um dos coeficientes espectrais compreende uma localização espectral dentro do espectro de sinal de áudio codificado e um valor

espectral, em que os coeficientes espectrais são ordenados sequencialmente de acordo com sua localização espectral dentro do espectro de sinal de áudio codificado de modo que os coeficientes espectrais formem uma sequência de coeficientes espectrais, determinar um ou mais pseudo coeficientes do espectro de sinal de áudio decodificado, em que cada um dos pseudo coeficientes é um dos coeficientes espectrais, substituir pelo menos um ou mais pseudo coeficientes por um padrão espectral determinado para adquirir um espectro de sinal de áudio modificado, em que o padrão espectral determinado compreende pelo menos dois coeficientes padrão, em que cada um dos pelo menos dois coeficientes padrão compreende um valor espectral, e converter o espectro de sinal de áudio modificado para um domínio de tempo para adquirir o sinal de saída de áudio.

21. Um método para gerar uma pluralidade de padrões espectrais, compreendendo:

a geração de uma pluralidade de sinais em um primeiro domínio,

a transformação de cada sinal da pluralidade de sinais a partir do primeiro domínio em um segundo domínio para obter uma pluralidade de padrões espectrais, cada padrão da pluralidade de padrões espectrais transformados, compreendendo uma pluralidade de coeficientes,

a truncagem dos padrões espectrais transformados pela remoção de um ou mais dos coeficientes dos padrões espectrais transformados para obter uma pluralidade de padrões processados, e

o armazenamento de cada padrão processado da pluralidade de padrões processados em uma base de dados ou uma memória,

caracterizado pela geração de cada sinal da pluralidade de sinais ser conduzida com base nas fórmulas

$$x(t) = cos\,(2\pi\varphi(t))$$

e

$$\varphi(t) = \varphi(0) + \int_0^t 2\pi f(\tau)d\tau \quad ,$$

em que t e τ indicam tempo,

em que φ(t) é uma fase instantânea em t, e

em que f(τ) é uma frequência instantânea em τ,

em que cada sinal da pluralidade de sinais tem uma frequência inicial (f_0), sendo uma frequência instantânea do referido sinal em um primeiro ponto no tempo, e uma frequência alvo (f_1), sendo uma frequência instantânea do referido sinal em um diferente segundo ponto no tempo,

em que a geração da pluralidade de sinais é conduzida ao gerar um primeiro sinal da pluralidade de sinais, de modo que a frequência alvo (f_1) do primeiro sinal seja igual a frequência inicial (f_0), e

em que a geração da pluralidade de sinais é conduzida ao gerar um segundo sinal diferente da pluralidade de sinais, de modo que a frequência alvo (f_1) do primeiro sinal seja diferente da frequência inicial (f_0).

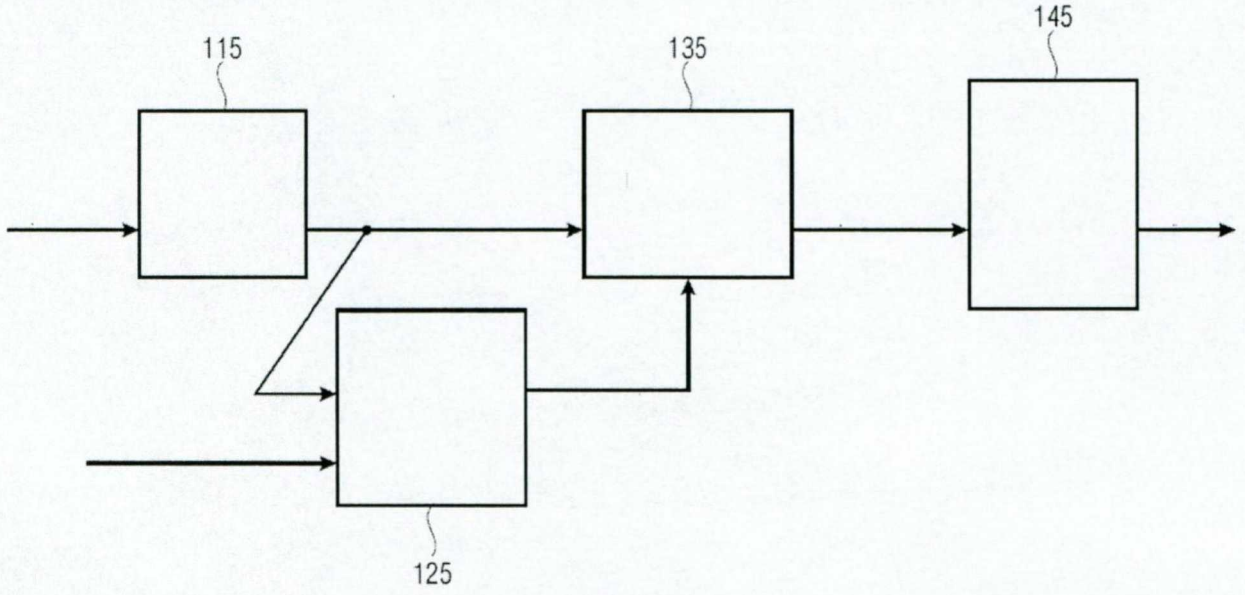


FIG 1A

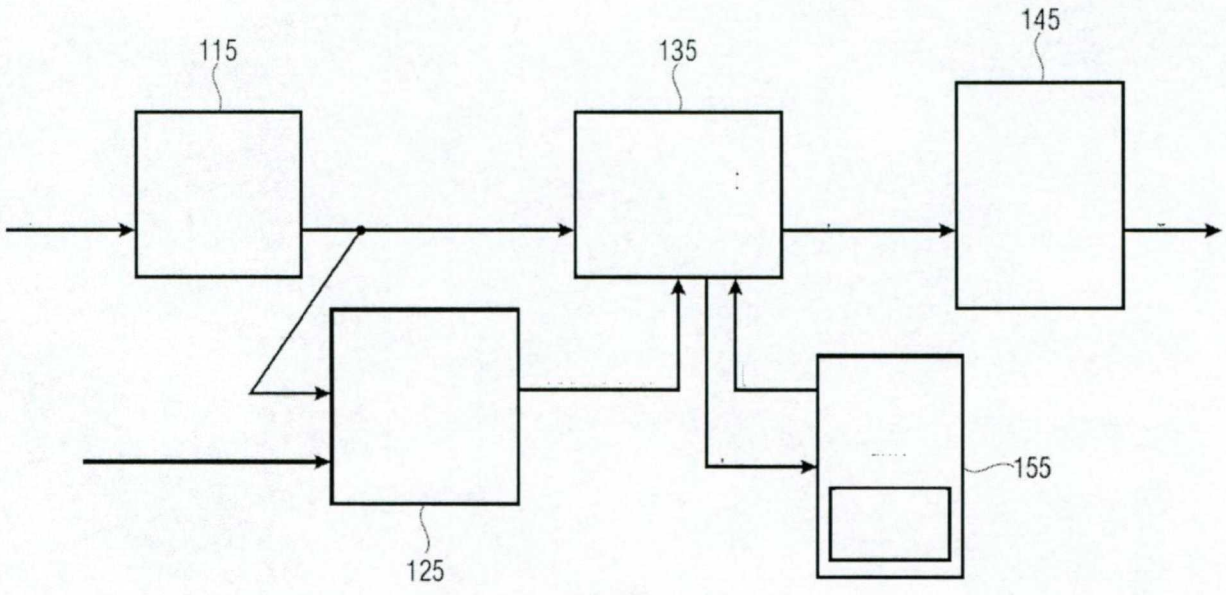


FIG 1B

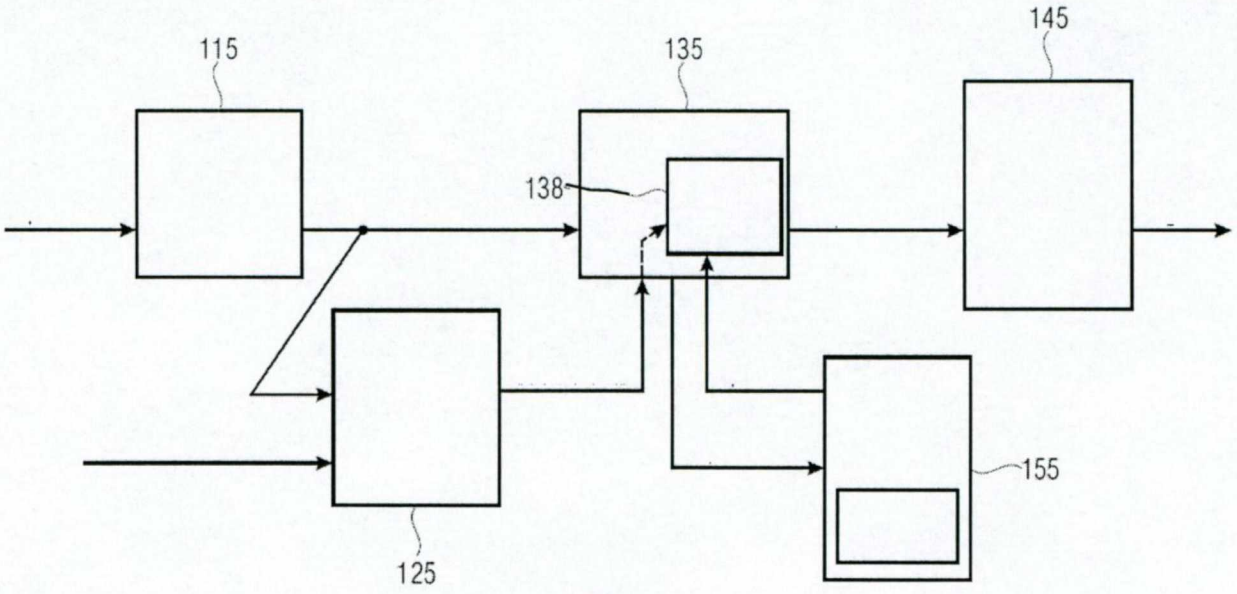


FIG 1C

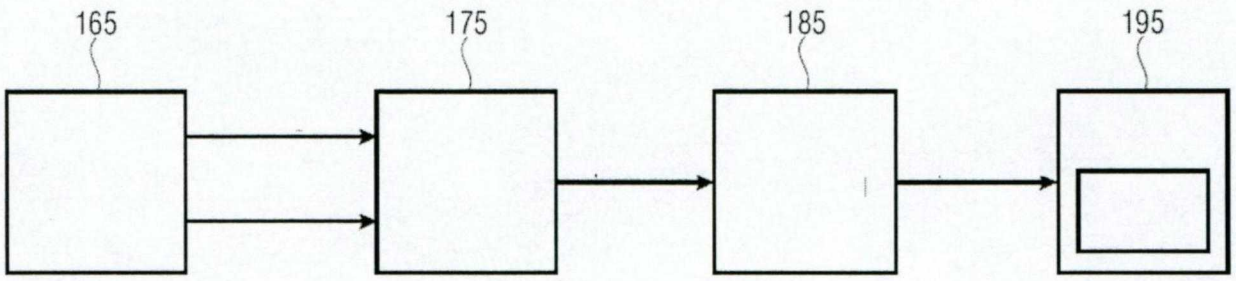


FIG 1D

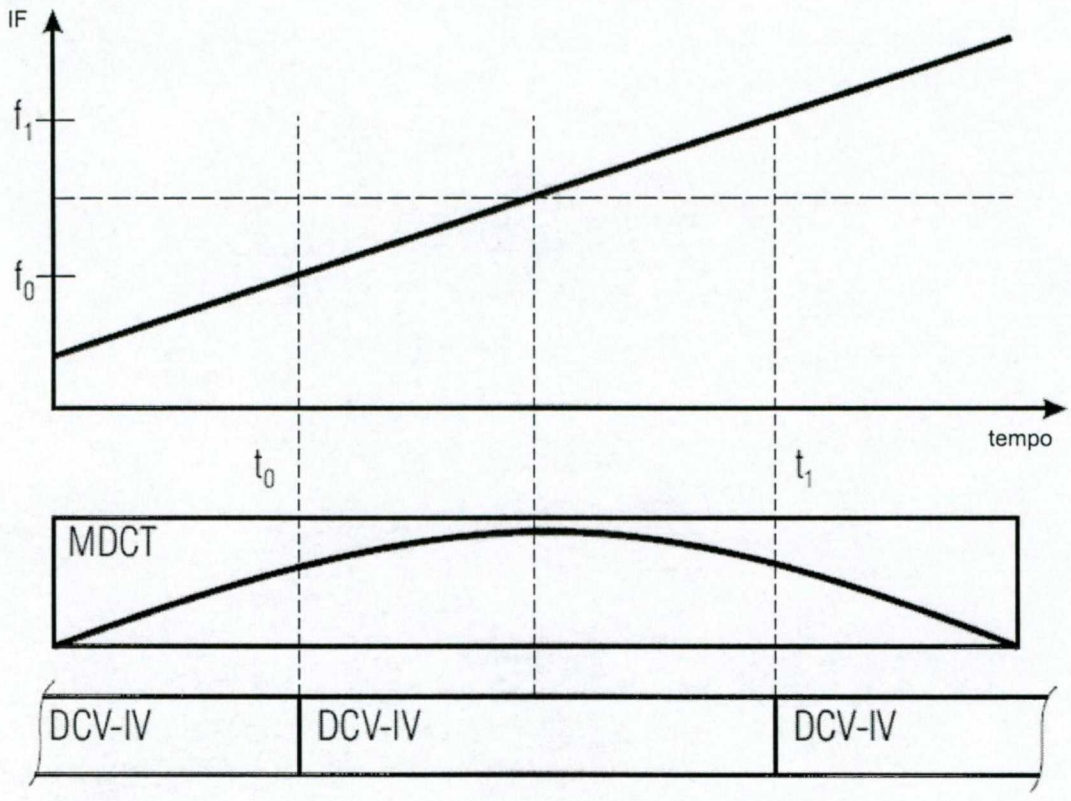


FIG 2

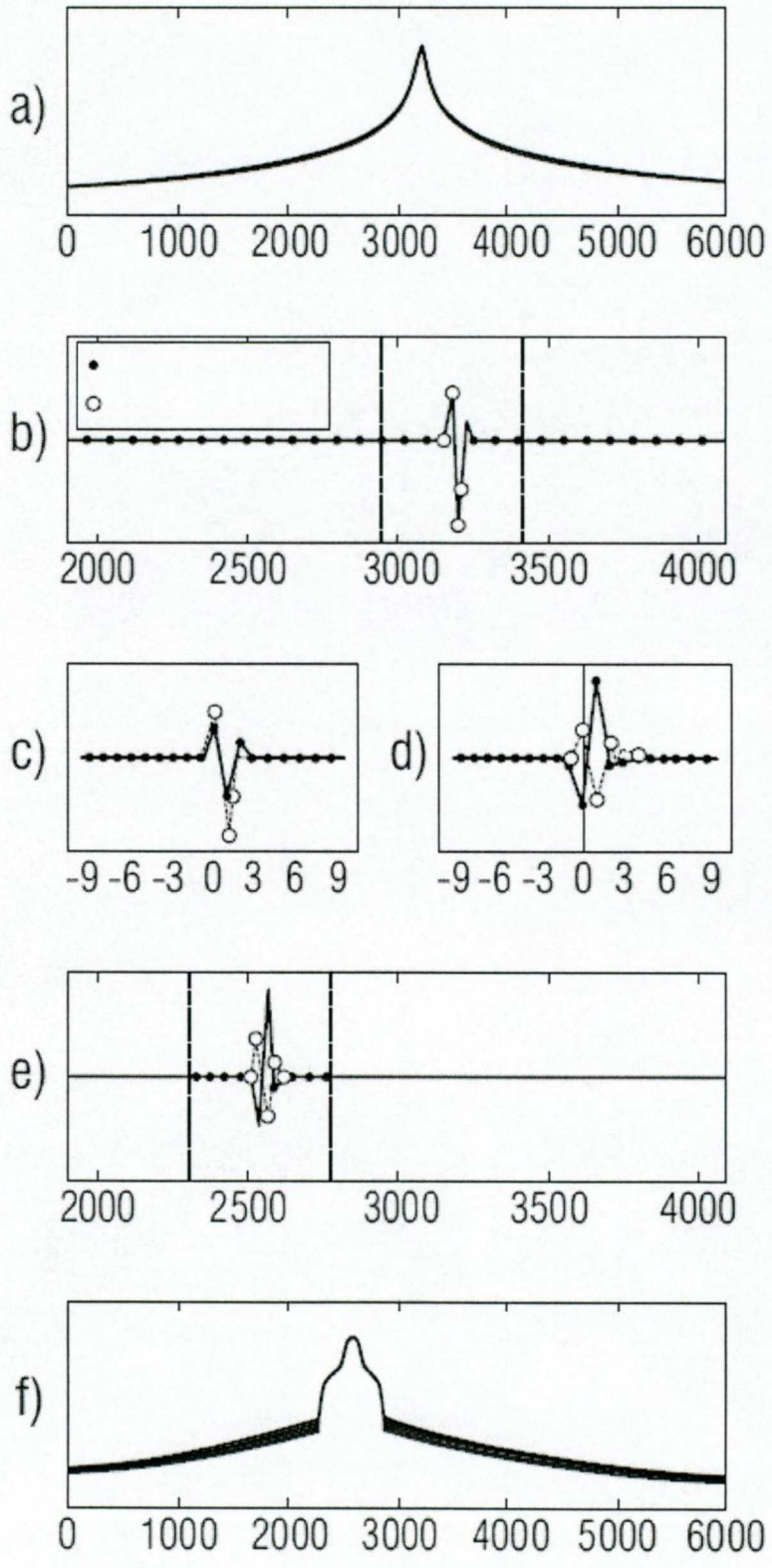

a)
0 1000 2000 3000 4000 5000 6000
b)
2000 2500 3000 3500 4000
c)
-9 -6 -3 0 3 6 9
d)
-9 -6 -3 0 3 6 9
e)
2000 2500 3000 3500 4000
f)
0 1000 2000 3000 4000 5000 6000

FIG 3

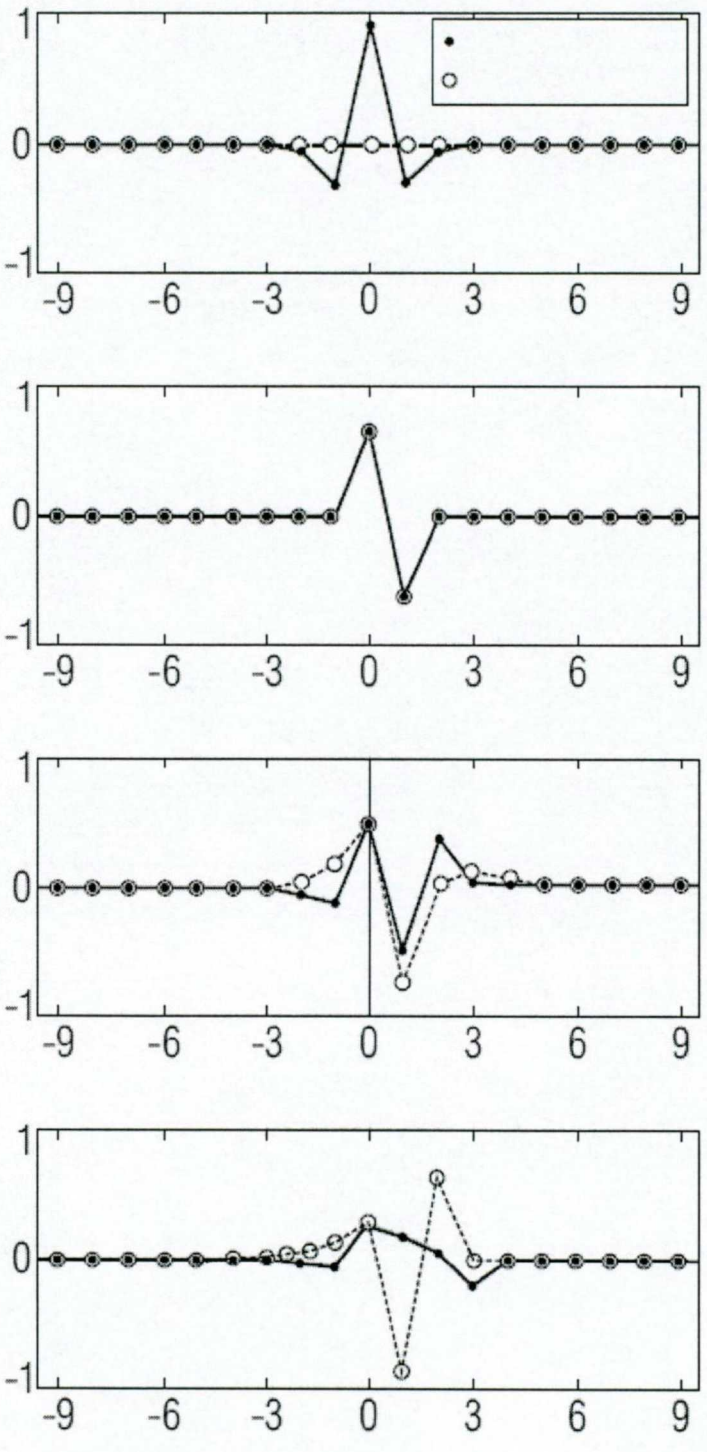


FIG 4

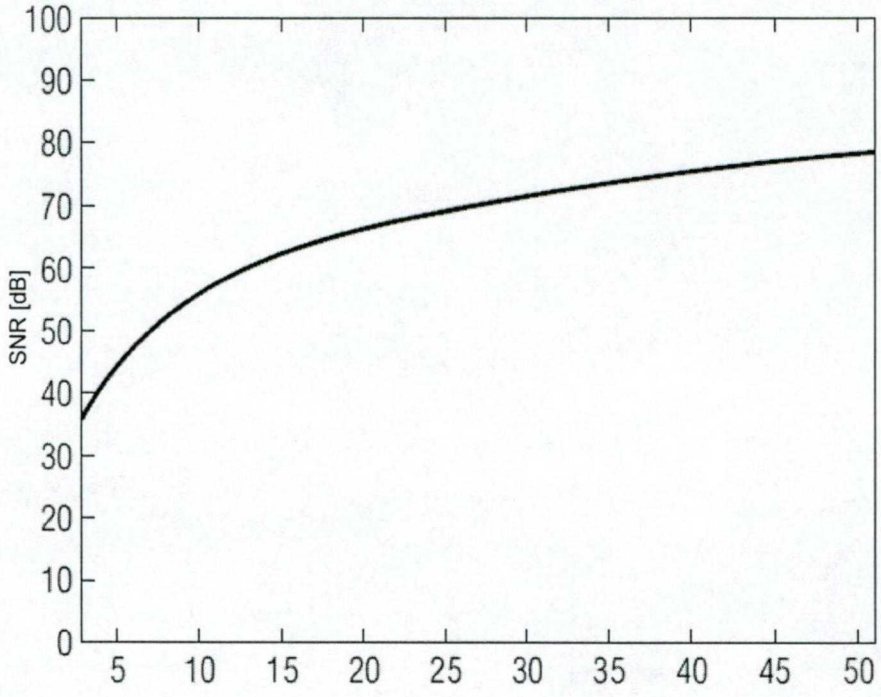


FIG 5

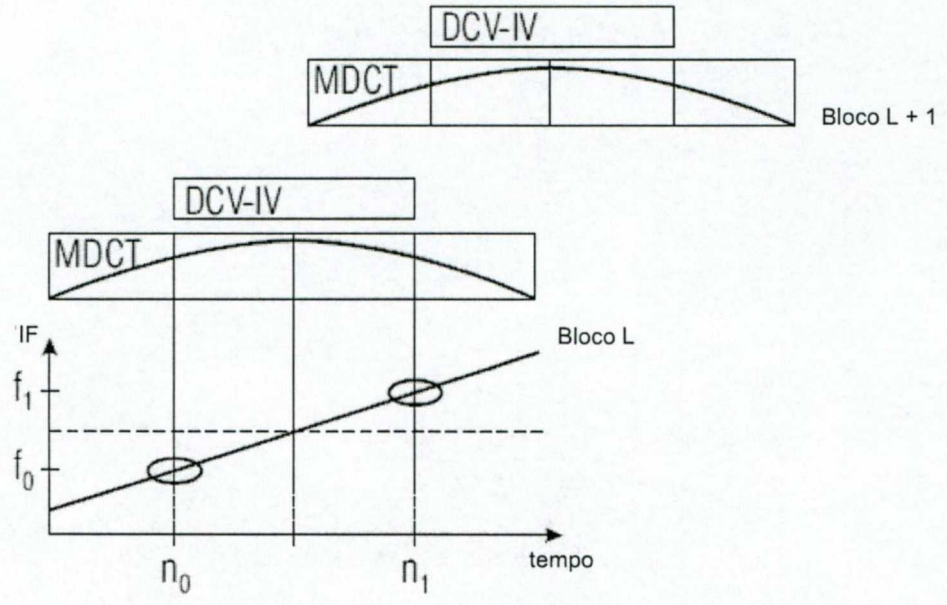


FIG 6A

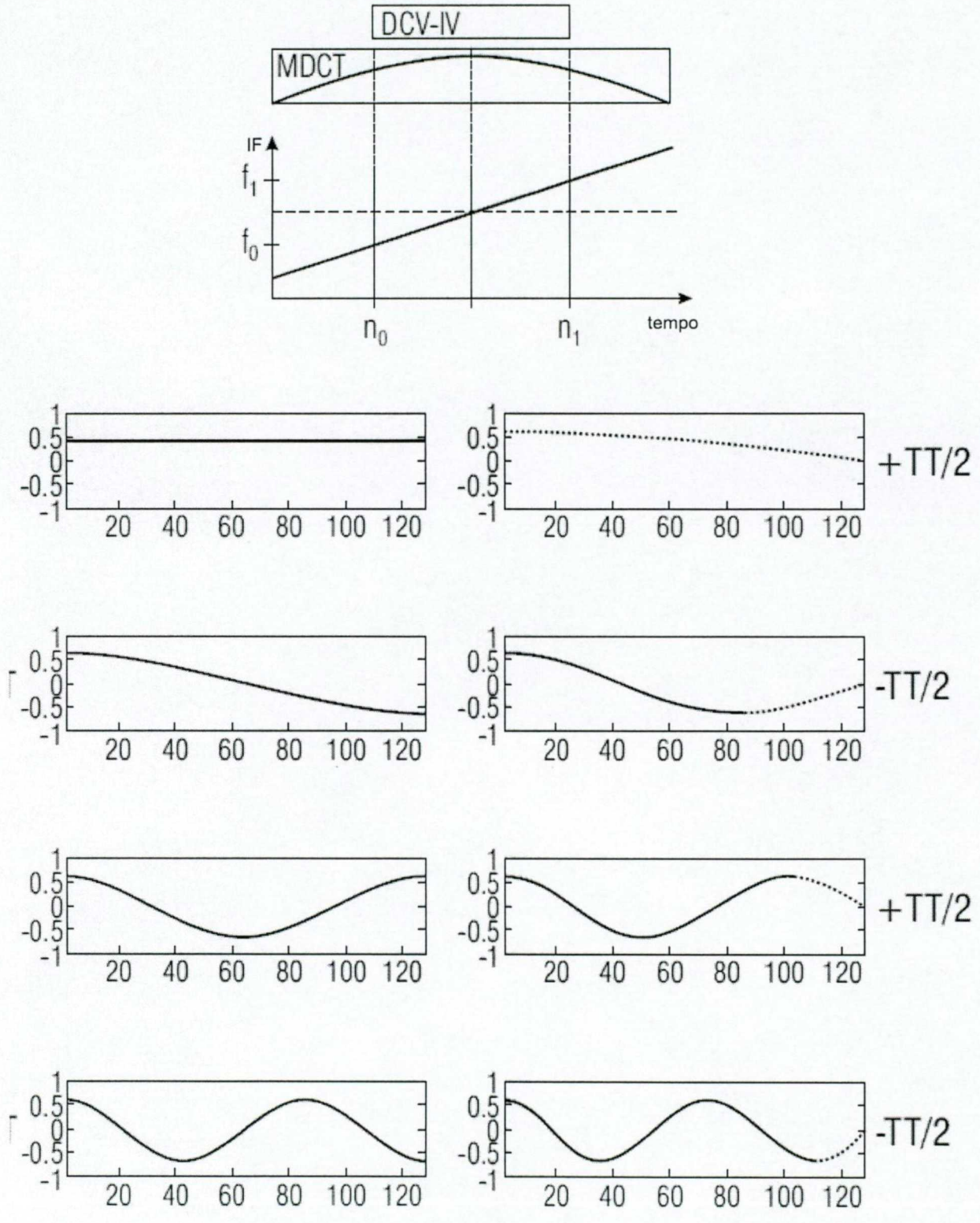


FIG 6B

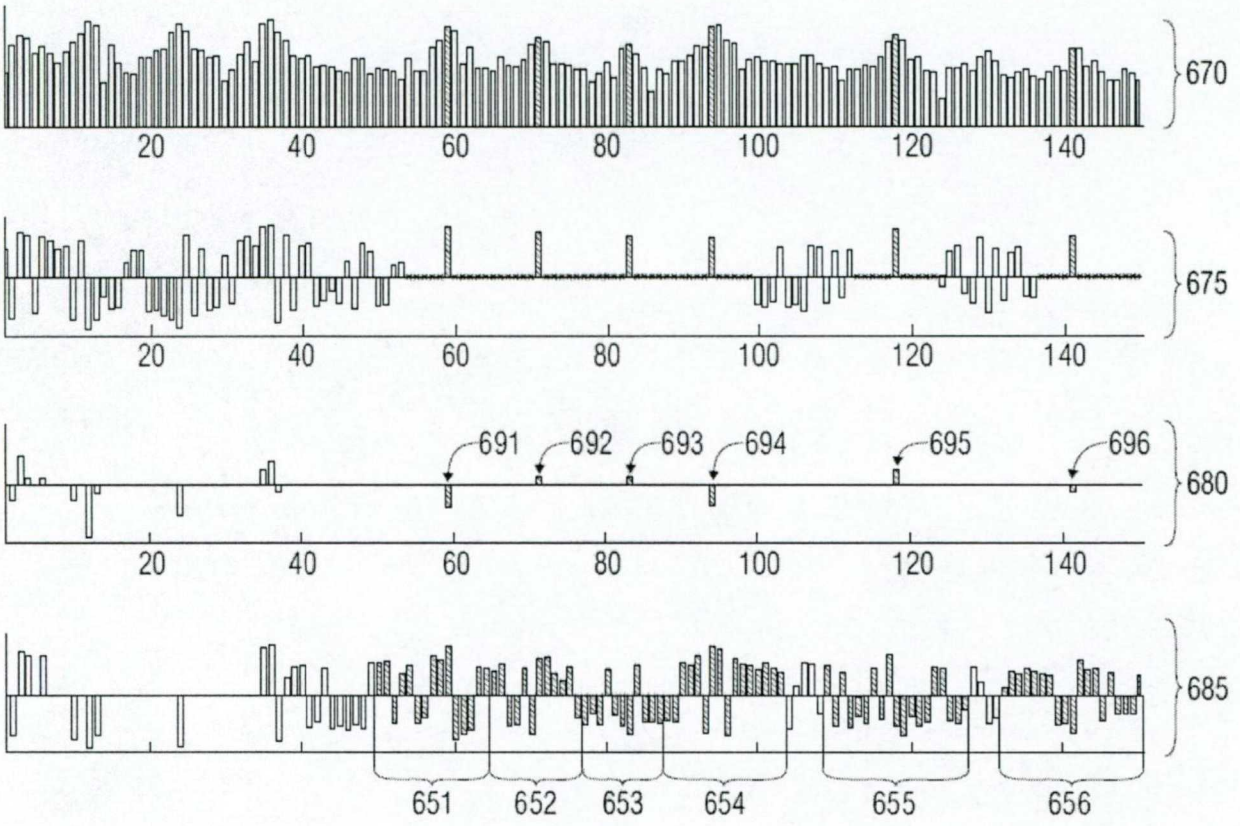

670
20
40
60
80
100
120
140
675
691
692
693
694
695
696
680
685
651
652
653
654
655
656

FIG 6C

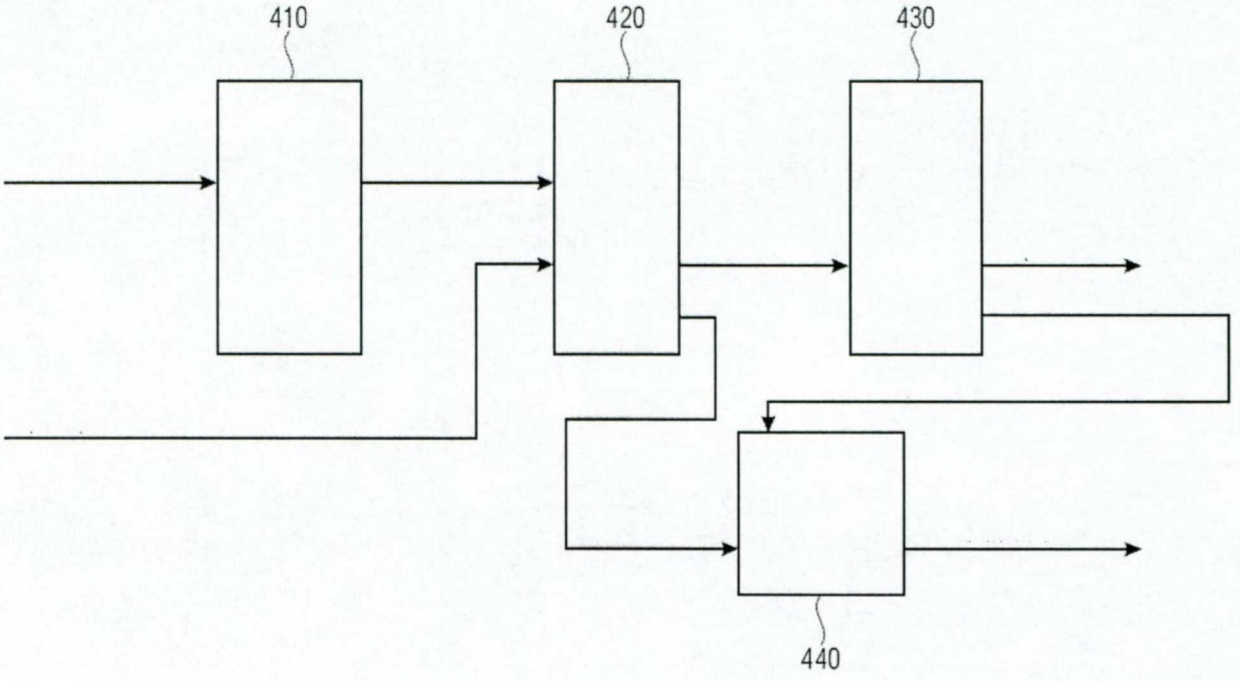


FIG 7

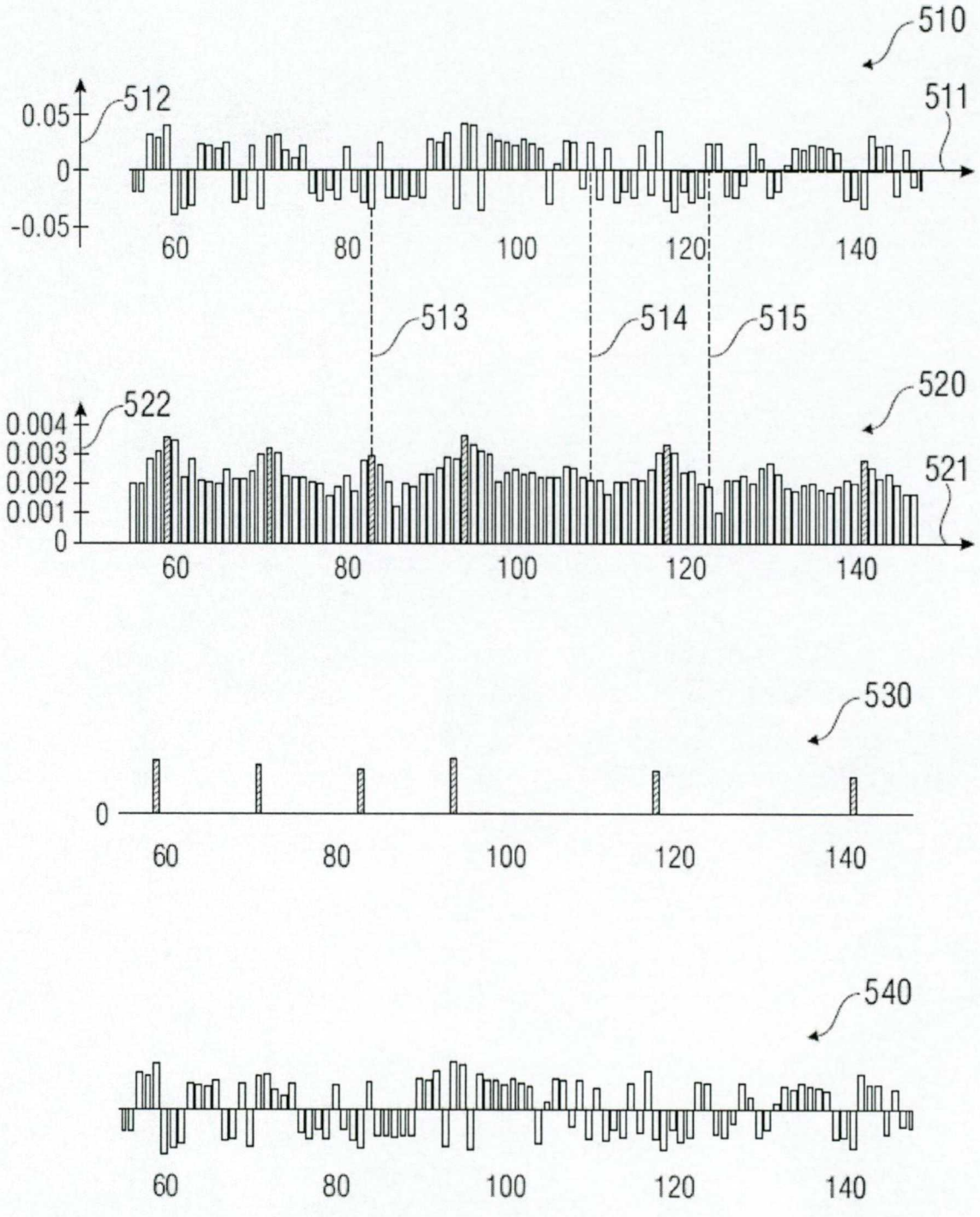

510
512
511
0.05
0
-0.05
60
80
100
120
140
513
514
515
520
522
521
0.004
0.003
0.002
0.001
0
530
540

FIG 8

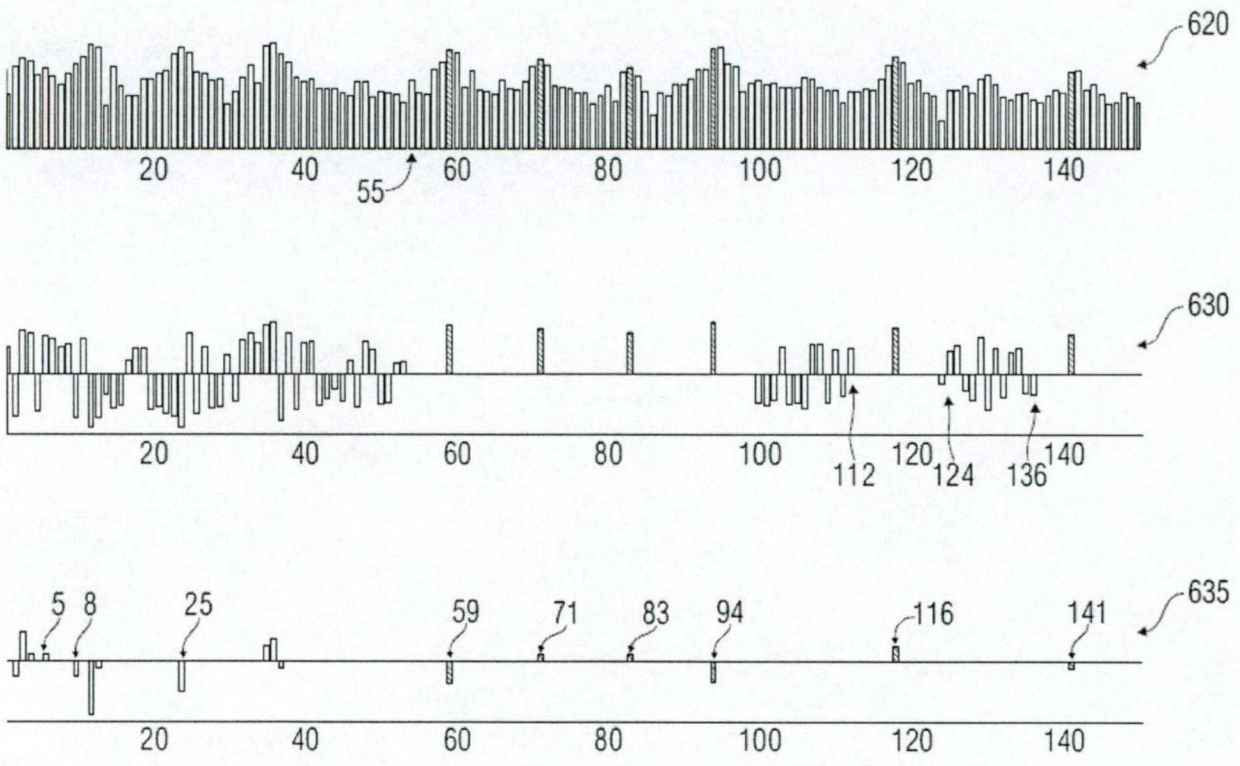

620
20
40
55
60
80
100
120
140
630
20
40
60
80
100
112
120
124
136
140
635
5
8
25
59
71
83
94
116
141
20
40
60
80
100
120
140

FIG 9

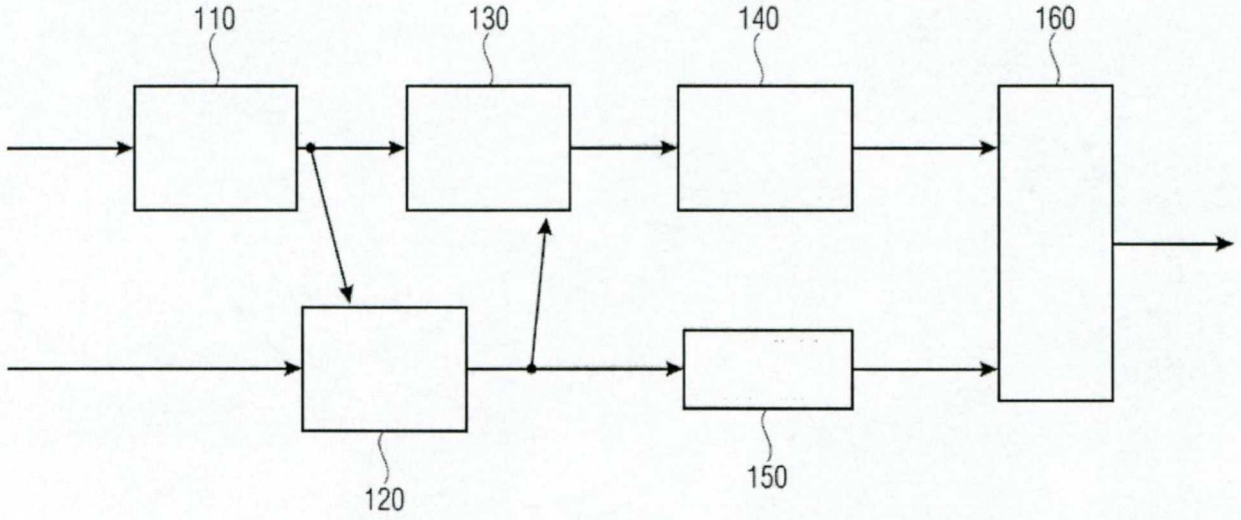


FIG 10

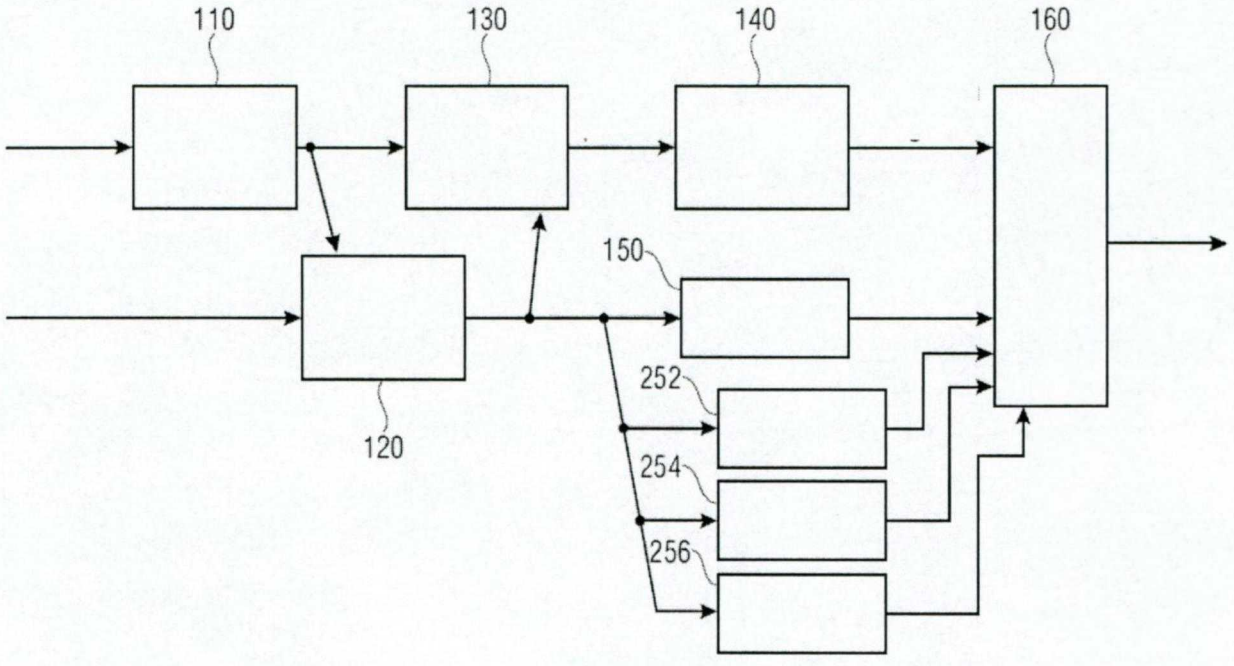


FIG 11

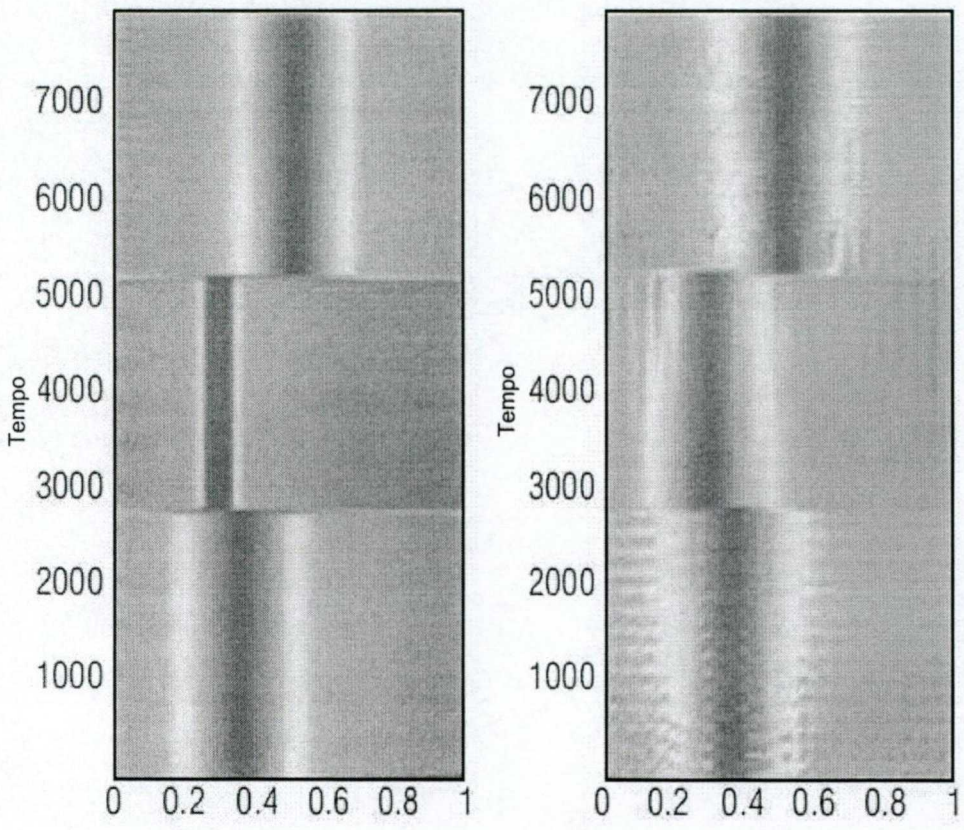

Tempo
7000
6000
5000
4000
3000
2000
1000
0
0.2
0.4
0.6
0.8
1
Tempo
7000
6000
5000
4000
3000
2000
1000
0
0.2
0.4
0.6
0.8
1

FIG 12