



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

(11) PI 1007528-3 B1



(22) Data do Depósito: 12/03/2010

(45) Data de Concessão: 13/10/2020

(54) Título: SISTEMA PARA GERAR UM SINAL DE SAÍDA A PARTIR DE UM SINAL DE ÁUDIO DE ENTRADA USANDO UM FATOR DE TRANSPOSIÇÃO T, MÉTODO PARA TRANSPOR UM SINAL DE ÁUDIO DE ENTRADA POR UM FATOR DE TRANSPOSIÇÃO T E MEIO DE ARMAZENAMENTO.

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

(30) Prioridade Unionista: 18/09/2009 US 61/243,624; 28/01/2009 SE 0900087-8.

(73) Titular(es): DOLBY INTERNATIONAL AB.

(72) Inventor(es): PER EKSTRAND; LARS FALCK VILLEMOS.

(86) Pedido PCT: PCT EP2010053222 de 12/03/2010

(87) Publicação PCT: WO 2010/086461 de 05/08/2010

(85) Data do Início da Fase Nacional: 28/07/2011

(57) Resumo: SISTEMA PARA GERAR UM SINAL DE SAÍDA A PARTIR DE UM SINAL DE ENTRADA USANDO UM FATOR DE TRANSPOSIÇÃO T, SISTEMA PARA DECODIFICAR UM SINAL DE MULTIMÍDIA, MÉTODO PARA TRANSPOR UM SINAL DE ENTRADA POR UM FATOR DE TRANSPOSIÇÃO T, PROGRAMA DE SOFTWARE E MEIO DE ARMAZENAMENTO. A presente invenção refere-se à transposição de sinais no tempo e/ou na frequência e, em particular, à codificação de sinais de áudio. Mais particularmente, a presente invenção refere-se a métodos de reconstrução de alta frequência (HFR) incluindo um transpositor harmônico de domínio de frequência. Um método e um sistema para a geração de um sinal de saída transposto a partir de um sinal de entrada usando-se um fator de transposição T são criados. O sistema compreende uma janela de análise de comprimento L_a , a extração de um quadro do sinal de entrada, e uma unidade de transformação de análise de ordem M transformando as amostras em coeficientes M complexos. M é uma função do fator de transposição T, uma unidade de transformação de síntese de ordem M transformando os coeficientes alterados em amostras M alteradas, e uma janela de síntese de comprimento L_s , gerando um quadro do sinal de (...).

Relatório Descritivo da Patente de Invenção para "**SISTEMA PARA GERAR UM SINAL DE ÁUDIO DE SAÍDA A PARTIR DE UM SINAL DE ÁUDIO DE ENTRADA USANDO UM FATOR DE TRANSPOSIÇÃO T, MÉTODO PARA TRANSPOR UM SINAL DE ÁUDIO DE ENTRADA POR UM FATOR DE TRANSPOSIÇÃO T E MEIO DE ARMAZENAMENTO**".

CAMPO TÉCNICO

[0001] A presente invenção refere-se à transposição de sinais em frequência e/ou distensão / compressão de um sinal no tempo e, em particular, à codificação de sinais de áudio. Em outras palavras, a presente invenção se refere a uma modificação de escala de tempo e/ou de escala de frequência. Mais particularmente, a presente invenção refere-se a métodos de reconstrução de alta frequência (HFR) incluindo um transpositor harmônico de domínio de frequência.

ANTECEDENTES DA INVENÇÃO

[0002] As tecnologias de HFR, tal como a tecnologia de replicação de banda espectral (SBR), permitem melhorar significativamente a eficiência de codificação de codificadores – decodificadores de áudio perceptivo tradicionais. Em combinação com uma codificação de áudio avançado (AAC) de MPEG-4, ela forma um codificador – decodificador de áudio muito eficiente, o qual já está em uso com o sistema de rádio por satélite XM e o Digital Radio Mondiale, e também padronizado com 3GPP, Fórum de DVD e outros. A combinação de AAC e SBR é denominada aacPlus. É uma parte do padrão MPEG-4, onde é referido como Perfil de AAC de alta eficiência (HE-AAC). Em geral, a tecnologia de HFR pode ser combinada com qualquer codificador – decodificador de áudio perceptivo de uma forma compatível para trás e para frente, assim oferecendo a possibilidade de um aprimoramento em sistemas de difusão já estabelecidos como a Camada 2 de MPEG usada no sistema Eureka DAB. Os métodos de transposição de HFR também

podem ser combinados com codificadores – decodificadores de fala para se permitir uma fala de banda larga em taxas de bit ultrabaixas.

[0003] A ideia básica por trás de HFR é a observação que, usualmente, uma correlação forte entre as características da faixa de frequência alta de um sinal e as características da faixa de frequência baixa do mesmo sinal está presente. Assim, uma boa aproximação para a representação da faixa de frequência alta de entrada original de um sinal pode ser obtida por uma transposição de sinal a partir da faixa de frequência baixa para a faixa de frequência alta.

[0004] Este conceito de transposição foi estabelecido na WO 98/57436, o qual é incorporado como referência, como um método para a recriação de uma banda de frequência alta a partir de uma banda de frequência mais baixa de um sinal de áudio. Uma economia substancial na taxa de bit pode ser obtida pelo uso deste conceito na codificação de áudio e/ou na codificação de fala. A seguir, uma referência será feita a uma codificação de áudio, mas deve ser notado que os métodos e sistemas descritos são igualmente aplicáveis a uma codificação de fala em uma codificação unificada de fala e de áudio (USAC).

[0005] Em um sistema de codificação de áudio baseado em HFR, um sinal de largura de banda baixa é apresentado a um codificador de forma de onda de núcleo para codificação, e frequências mais altas são geradas de novo no lado de decodificador usando-se a transposição do sinal de largura de banda baixa e uma informação de lado adicional, a qual é tipicamente codificada em taxas de bit mais baixas e a qual descreve o formato espectral alvo. Para taxas de bit baixas, onde a largura de banda do sinal codificado de núcleo é estreita, torna-se crescentemente importante reproduzir ou sintetizar uma banda alta, isto é, a faixa de frequência alta do sinal de áudio, com características perceptivamente agradáveis.

[0006] Na técnica anterior, há vários métodos para a reconstrução

de frequência alta usando-se, por exemplo, transposição harmônica, ou uma distensão de tempo. Um método é baseado em codificadores de voz em fase operando sob o princípio de realização de uma análise de frequência com uma resolução de frequência suficientemente alta. Uma modificação de sinal é realizada no domínio de frequência antes de uma ressintetização do sinal. A modificação de sinal pode ser uma operação de distensão de tempo ou de transposição.

[0007] Um dos problemas subjacentes que existem com estes métodos são as restrições opostas de uma resolução de frequência alta pretendida, de modo a se obter uma transposição de alta qualidade para sons estacionários e a resposta no tempo do sistema para sons transientes ou percussivos. Em outras palavras, embora o uso de uma resolução de alta frequência seja benéfico para a transposição de sinais estacionários, essa resolução de frequência alta tipicamente requer grandes tamanhos de janela, os quais são prejudiciais quando se lida com porções transientes de um sinal. Uma abordagem para se lidar com este problema pode ser mudar de forma adaptativa as janelas do transpositor, por exemplo, pelo uso de uma comutação de janela como uma função de características de sinal de entrada. Tipicamente, janelas longas serão usadas para porções estacionárias de um sinal, de modo a se obter uma resolução de frequência alta, enquanto janelas curtas serão usadas para porções transientes do sinal, de modo a se implementar uma boa resposta transiente, isto é, uma boa resolução temporal do transpositor. Contudo, esta abordagem tem o inconveniente de medidas de análise de sinal, tal como uma detecção transiente ou similar, terem que ser incorporadas no sistema de transposição. Essas medidas de análise de sinal frequentemente envolvem uma etapa de decisão, por exemplo, uma decisão quanto à presença de um transiente, o que dispara uma comutação de processamento de sinal. Mais ainda, essas medidas tipicamente afetam a confiabilidade do sistema, e elas

podem introduzir artefatos de sinal quando da comutação do processamento de sinal, por exemplo, quando de uma comutação entre tamanhos de janela.

[0008] A presente invenção resolve os problemas mencionados anteriormente com referência ao desempenho transiente de transposição harmônica sem a necessidade de uma comutação de janela. Mais ainda, uma transposição harmônica melhorada é obtida a uma complexidade adicional baixa.

SUMÁRIO DA INVENÇÃO

[0009] A presente invenção refere-se ao problema de desempenho transiente melhorado para transposição harmônica, bem como melhoramentos sortidos para métodos conhecidos para transposição harmônica. Mais ainda, a presente invenção destaca como uma complexidade adicional pode ser mantida em um mínimo, enquanto se retêm os melhoramentos propostos.

[00010] Dentre outros, a presente invenção compreende pelo menos um dos aspectos a seguir:

- superamostragem na frequência por um fator que é uma função do fator de transposição do ponto de operação do transpositor;
- escolha apropriada da combinação de janelas de análise e de síntese; e
- garantia de um alinhamento de tempo de diferentes sinais transpostos para os casos em que esses sinais são combinados.

[00011] De acordo com um aspecto da invenção, um sistema para a geração de um sinal de saída transposto a partir de um sinal de entrada usando-se um fator de transposição T é descrito. O sinal de saída transposto pode ser uma versão distendida no tempo e/ou deslocada na frequência do sinal de entrada. Com relação ao sinal de entrada, o sinal de saída transposto pode ser distendido no tempo pelo fator de transposição T . Alternativamente, as componentes de frequência do

sinal de saída transposto podem ser deslocadas para cima pelo fator de transposição T.

[00012] O sistema pode compreender uma janela de análise de comprimento L, a qual extrai amostras L do sinal de entrada. Tipicamente, as amostras L dos sinais de entrada são amostras do sinal de entrada, por exemplo, um sinal de áudio, no domínio de tempo. As amostras L extraídas são referidas como um quadro do sinal de entrada. O sistema compreende, ainda, uma unidade de transformação de análise de ordem $M = F * L$ que transforma as amostras L de domínio de tempo em M coeficientes complexos com F sendo um fator de sobreamostragem de frequência. Os coeficientes complexos M tipicamente são coeficientes no domínio de frequência. A transformação de análise pode ser uma transformada de Fourier, uma transformada de Fourier Rápida, uma transformada de Fourier Discreta, uma Transformada de Wavelet ou um estágio de análise de um banco de filtro (possivelmente modulado). O fator de sobreamostragem F é baseado em ou é uma função do fator de transposição T.

[00013] A operação de sobreamostragem também pode ser referida como preenchimento com zero da janela de análise pela adição de $(F - 1) * L$ zeros. Também pode ser vista como escolhendo um tamanho de uma transformação de análise M o qual é maior do que o tamanho da janela de análise por um fator F.

[00014] O sistema também pode compreender uma unidade de processamento não linear que altera a fase dos coeficientes complexos pelo uso do fator de transposição T. A alteração da fase pode compreender a multiplicação da fase dos coeficientes complexos pelo fator de transposição T. Além disso, o sistema pode compreender uma unidade de transformação de síntese de ordem M que transforma os coeficientes alterados em M amostras alteradas e uma janela de síntese de comprimento L para a geração do sinal de saída. A transformada de

síntese pode ser uma transformada de Fourier inversa, uma transformada de Fourier Rápida inversa, uma transformada de Fourier Discreta inversa, uma Transformada de Wavelet inversa ou um estágio de síntese de um banco de filtro (possivelmente) modulado. Tipicamente, a transformada de análise e a transformada de síntese estão relacionadas a cada outra, por exemplo, de modo a se obter uma reconstrução perfeita de um sinal de entrada quando o fator de transposição $T = 1$.

[00015] De acordo com um outro aspecto da invenção, o fator de sobreamostragem F é proporcional ao fator de transposição T . Em particular, o fator de sobreamostragem F pode ser maior do que ou igual a $(T + 1)/2$. Esta seleção do fator de sobreamostragem F assegura que artefatos indesejáveis, por exemplo, pré- e pós-ecos, os quais podem ser incorridos pela transposição, sejam rejeitados pela janela de síntese.

[00016] Deve ser notado que, em termos mais gerais, o comprimento da janela de análise pode ser L_a e o comprimento da janela de síntese pode ser L_s . Também nesses casos, pode ser benéfico selecionar a ordem da unidade de transformação M com base na ordem de transposição T , isto é, como uma função da ordem de transposição T . Mais ainda, pode ser benéfico selecionar M para ser maior do que o comprimento médio da janela de análise e da janela de síntese, isto é, maior do que $(L_a + L_s)/2$. Em uma modalidade, a diferença entre a ordem da unidade de transformação M e o comprimento de janela médio é proporcional a $(T - 1)$. Em uma outra modalidade, M é selecionado para ser maior do que ou igual a $(TL_a + L_s)/2$. Deve ser notado que o caso em que o comprimento da janela de análise e da janela de síntese é igual, isto é, $L_a = L_s = L$, é um caso especial do caso geral acima. Para o caso geral, o fator de sobreamostragem F pode ser

$$F \geq 1 + (T - 1) \frac{L_a}{L_s + L_a}$$

[00017] O sistema ainda pode compreender uma unidade de passo

de análise deslocando a janela de análise por um passo de análise de amostras S_a ao longo do sinal de entrada. Como resultado da unidade de passo de análise, uma sucessão de quadros do sinal de entrada é gerada. Além disso, o sistema pode compreender uma unidade de passo de síntese deslocando a janela de síntese e/ou quadros sucessivos do sinal de saída por um passo de síntese de amostras S_s . Como resultado, uma sucessão de quadros deslocados do sinal de saída é gerada, a qual pode ser sobreposta e adicionada em uma unidade de superposição – adição.

[00018] Em outras palavras, a janela de análise pode extrair ou isolar L ou, mais geralmente, L_a amostras do sinal de entrada, por exemplo, pela multiplicação de um conjunto de amostras L do sinal de entrada com coeficientes de janela não nulos. Um conjunto como esse de amostras L pode ser referido como um quadro de sinal de entrada ou como um quadro do sinal de entrada. A unidade de passo de análise desloca a janela de análise ao longo do sinal de entrada e, desse modo, seleciona um quadro diferente do sinal de entrada, isto é, gera uma sequência de quadros do sinal de entrada. A distância de amostra entre quadros sucessivos é dada pelo passo de análise. De uma maneira similar, a unidade de passo de síntese desloca a janela de síntese e/ou os quadros do sinal de saída, isto é, gera uma sequência de quadros deslocados do sinal de saída. A distância de amostra entre quadros sucessivos do sinal de saída é dada pela janela de síntese. O sinal de saída pode ser determinado pela superposição da sequência de quadros do sinal de saída e pela adição de valores de amostra os quais coincidem no tempo.

[00019] De acordo com um outro aspecto da invenção, o passo de síntese é T vezes o passo de análise. Nesses casos, o sinal de saída corresponde ao sinal de entrada, distendido no tempo pelo fator de transposição T . Em outras palavras, pela seleção do passo de síntese

para ser T vezes maior do que o passo de análise, um deslocamento no tempo ou uma distensão no tempo do sinal de saída com respeito ao sinal de entrada pode ser obtido. Este deslocamento no tempo é de ordem T .

[00020] Em outras palavras, o sistema mencionado acima pode ser descrito conforme se segue. Usando uma unidade de janela de análise, uma unidade de transformação de análise e uma unidade de passo de análise com um passo de análise S_a , uma suíte ou uma sequência de conjuntos de M coeficientes complexos pode ser determinada a partir de um sinal de entrada. O passo de análise define o número de amostras que a janela de análise é movida para frente ao longo do sinal de entrada. Como o tempo decorrido entre duas amostras sucessivas é dado pela taxa de amostragem, o passo de análise também define o tempo decorrido entre dois quadros do sinal de entrada. Como consequências, também o tempo decorrido entre dois conjuntos sucessivos de M coeficientes complexos é dado pelo passo de análise S_a .

[00021] Após passar pela unidade de processamento não linear, onde a fase dos coeficientes complexos pode ser alterada, por exemplo, pela multiplicação dela pelo fator de transposição T , a suíte ou sequência de conjuntos de M coeficientes complexos pode ser reconvertida no domínio de tempo. Cada conjunto de M coeficientes complexos alterados pode ser transformado em M amostras alteradas usando-se a unidade de transformação de síntese. Em uma operação de superposição – adição seguinte envolvendo a unidade de janela de síntese e a unidade de passo de síntese com um passo de síntese S_s , a suíte de conjuntos de M amostras alteradas pode ser sobreposta e adicionada para a formação do sinal de saída. Nesta operação de superposição – adição, conjuntos sucessivos de M amostras alteradas podem ser deslocados por S_s amostras uns com respeito aos outros,

antes de eles serem multiplicados pela janela de síntese e subsequentemente adicionados para a produção do sinal de saída. Consequentemente, se a janela de síntese S_s for T vezes a janela de análise S_a , o sinal poderá ser distendido no tempo por um fator F .

[00022] De acordo com um outro aspecto da invenção, a janela de síntese é derivada da janela de análise e do passo de síntese. Em particular, a janela de síntese pode ser dada pela fórmula:

$$v_s(n) = v_a(n) \left(\sum_{k=-\infty}^{\infty} (v_a(n - k \cdot \Delta t))^2 \right)^{-1}$$

com $v_s(n)$ sendo a janela de síntese, $v_a(n)$ sendo a janela de análise; e Δt sendo o passo de síntese S_s . A janela de análise e/ou de síntese pode ser uma janela gaussiana; uma janela de cosseno; uma janela de Hamming; uma janela de Hann; uma janela retangular; uma janela de Bartlett; uma janela de Blackman; uma janela que tem a função $v(n) = \text{sen}(\pi/L (n + 0,5))$, $0 \leq n < L$, em que, no caso de comprimentos diferentes da janela de análise e da janela de síntese, L pode ser L_a ou L_s , respectivamente.

[00023] De acordo com um outro aspecto da invenção, o sistema ainda compreende uma unidade de contração que realiza, por exemplo, uma conversão de taxa do sinal de saída pela ordem de transposição T , desse modo se produzindo um sinal de saída transposto. Pela seleção do passo de síntese para ser T vezes o passo de análise, um sinal de saída distendido no tempo pode ser obtido, conforme destacado acima. Se a taxa de amostragem do sinal distendido no tempo for aumentada por um fator T ou se o sinal distendido no tempo for de amostragem reduzida por um fator T , um sinal de saída transposto poderá ser gerado, que corresponde ao sinal de entrada, deslocado na frequência pelo fator de transposição T . A operação de redução de amostragem pode compreender a etapa de seleção de apenas um subconjunto de amostras do sinal de saída. Tipicamente, apenas toda T -ésima amostra

do sinal de saída é retida. Alternativamente, a taxa de amostragem pode ser aumentada por um fator T , isto é, a taxa de amostragem é interpretada como sendo T vezes mais alta. Em outras palavras, uma reamostragem ou uma conversão de taxa de amostragem significa que a taxa de amostragem é mudada, para um valor mais alto ou mais baixo. Uma redução de amostragem significa uma conversão de taxa para um valor mais baixo.

[00024] De acordo com um outro aspecto da invenção, o sistema pode gerar um segundo sinal de saída a partir do sinal de entrada. O sistema pode compreender uma segunda unidade de processamento não linear alterando a fase dos coeficientes complexos pelo uso de um segundo fator de transposição T_2 e uma segunda unidade de passo de síntese deslocando a janela de síntese e/ou os quadros do segundo sinal de saída por um segundo passo de síntese. A alteração da fase pode compreender a multiplicação da fase por um fator T_2 . Pela alteração da fase dos coeficientes complexos usando-se o segundo fator de transposição e pela transformação dos segundos coeficientes alterados em M segundas amostras alteradas e pela aplicação da janela de síntese, os quadros do segundo sinal de saída podem ser gerados a partir de um quadro do sinal de entrada. Pela aplicação do segundo passo de síntese à sequência de quadros do segundo sinal de saída, o segundo sinal de saída poderá ser gerado na unidade de superposição – adição.

[00025] O segundo sinal de saída pode ser contraído na segunda unidade de contração que realiza, por exemplo, uma conversão de taxa do segundo sinal de saída pela segunda ordem de transposição T_2 . Isto produz um segundo sinal de saída transposto. Em resumo, um primeiro sinal de saída transposto pode ser gerado usando-se o primeiro fator de transposição T e um segundo sinal de saída transposto pode ser gerado usando-se o segundo fator de transposição T_2 . Estes dois sinais de

saída transpostos então podem ser fundidos em uma unidade de combinação para a produção do sinal de saída transposto geral. A operação de fusão pode compreender a adição dos dois sinais de saída transpostos. Essa geração e a combinação de uma pluralidade de sinais de saída transpostos podem ser benéficas para a obtenção de boas aproximações do componente de sinal de frequência alta, a qual é para ser sintetizada. Deve ser notado que qualquer número de sinais de saída transpostos pode ser gerado usando-se uma pluralidade de ordens de transposição. Esta pluralidade de sinais de saída transpostos pode ser fundida, então, por exemplo, adicionada em uma unidade de combinação para a produção de um sinal de saída transposto geral.

[00026] Pode ser benéfico que a unidade de combinação atribua pesos aos primeiro e segundo sinais de saída transpostos, antes da fusão. A atribuição de peso pode ser realizada de modo que a energia ou a energia por largura de banda dos primeiro e segundo sinais de saída transpostos corresponda à energia ou à energia por largura de banda do sinal de saída, respectivamente.

[00027] De acordo com um aspecto adicional da invenção, o sistema pode compreender uma unidade de alinhamento a qual aplica um desvio no tempo aos primeiro e segundo sinais de saída transpostos, antes da entrada na unidade de combinação. Esse desvio no tempo pode compreender o deslocamento dos dois sinais de saída transpostos com respeito a cada outro no domínio de tempo. O desvio no tempo pode ser uma função da ordem de transposição e/ou do comprimento das janelas. Em particular, o desvio no tempo pode ser determinado como $(T - 2)L/4$.

[00028] De acordo com um outro aspecto da invenção, o sistema de transposição descrito acima pode ser embutido em um sistema para a decodificação de um sinal de multimídia recebido compreendendo um sinal de áudio. O sistema de decodificação pode compreender uma

unidade de transposição a qual corresponde ao sistema destacado acima, onde o sinal de entrada tipicamente é um componente de frequência baixa do sinal de áudio e o sinal de saída é um componente de frequência alta do sinal de áudio. Em outras palavras, o sinal de entrada tipicamente é um sinal de passa baixa com uma certa largura de banda e o sinal de saída é um sinal de passa banda de uma largura de banda tipicamente mais alta. Mais ainda, ele pode compreender um decodificador de núcleo para a decodificação do componente de frequência baixa do sinal de áudio a partir do fluxo de bit recebido. Esse decodificador de núcleo pode ser com base em um esquema de codificação, tal como Dolby E, Dolby Digital ou AAC. Em particular, esse sistema de decodificação pode ser uma caixa adaptadora para a decodificação de um sinal de multimídia recebido compreendendo um sinal de áudio e outros sinais, tal como de vídeo.

[00029] Deve ser notado que a presente invenção também descreve um método para a transposição de um sinal de entrada por um fator de transposição T . O método corresponde ao sistema destacado acima e pode compreender qualquer combinação dos aspectos mencionados acima. Ele pode compreender as etapas de extração de amostras do sinal de entrada usando-se uma janela de análise de comprimento L , e de seleção de um fator de sobreamostragem F como uma função do fator de transposição T . Ele ainda pode compreender as etapas de transformação das amostras L do domínio de tempo para o domínio de frequência produzindo coeficientes complexos $F * L$, e de alteração dos coeficientes complexos com o fator de transposição T . Em etapas adicionais, o método pode transformar os coeficientes complexos $F * L$ alterados no domínio de tempo produzindo as $F * L$ amostras alteradas, e pode gerar o sinal de saída usando uma janela de síntese de comprimento L . Deve ser notado que o método também pode ser adaptado para os comprimentos gerais da janela de análise e de

síntese, isto é, para L_a e L_s gerais, conforme destacado acima.

[00030] De acordo com um outro aspecto da invenção, o método pode compreender as etapas de deslocamento da janela de análise por um passo de análise de S_a amostras ao longo do sinal de entrada, e/ou pelo deslocamento da janela de síntese e/ou dos quadros do sinal de saída por um passo de síntese de S_s amostras. Pela seleção do passo de análise para ser T vezes o passo de análise, o sinal de saída pode ser distendido no tempo com respeito ao sinal de entrada por um fator T . Quando da execução de uma etapa adicional de realização de uma conversão de taxa do sinal de saída pela ordem de transposição T , um sinal de saída transposto pode ser obtido. Esse sinal de saída transposto pode compreender componentes de frequência que são deslocadas para cima por um fator T com respeito às componentes de frequência correspondentes do sinal de entrada.

[00031] O método ainda pode compreender as etapas para a geração de um segundo sinal de saída. Isto pode ser implementado pela alteração da fase dos coeficientes complexos pelo uso de um segundo fator de transposição T_2 , pelo deslocamento da janela de síntese e/ou dos quadros do segundo sinal de saída por um segundo passo de síntese, um segundo sinal de saída pode ser gerado usando-se o segundo fator de transposição T_2 e o segundo passo de síntese. Pela realização de uma conversão de taxa do segundo sinal de saída pela segunda ordem de transposição T_2 , um segundo sinal de saída transposto pode ser gerado. Eventualmente, pela fusão dos primeiro e segundo sinais de saída transpostos, um sinal de saída transposto fundido ou geral incluindo as componentes de sinal de frequência alta geradas por duas ou mais transposições com diferentes fatores de transposição pode ser obtido.

[00032] De acordo com outros aspectos da invenção, a invenção descreve um programa de software adaptado para execução em um

processador e para a realização das etapas de método da presente invenção, quando realizadas em um dispositivo de computação. A invenção também descreve um meio de armazenamento que compreende um programa de software adaptado para execução em um processador e para a realização das etapas de método da invenção, quando realizadas em um dispositivo de computação. Mais ainda, a invenção descreve um produto de programa de computador que compreende instruções executáveis para a realização do método da invenção, quando executado em um computador.

[00033] De acordo com um outro aspecto, um outro método e um sistema para a transposição de um sinal de entrada por um fator de transposição T são descritos. Este método e o sistema podem ser usados independentemente ou em combinação com os métodos e sistemas destacados acima. Qualquer um dos recursos destacados no presente documento pode ser aplicado a este método / sistema e vice-versa.

[00034] O método pode compreender a etapa de extração de um quadro de amostras do sinal de entrada usando-se uma janela de análise de comprimento L . Então, o quadro do sinal de entrada pode ser transformado a partir do domínio de tempo para o domínio de frequência produzindo M coeficientes complexos. A fase dos coeficientes complexos pode ser alterada com o fator de transposição T e os M coeficientes complexos alterados podem ser transformados no domínio de tempo, produzindo M amostras alteradas. Eventualmente, um quadro de um sinal de saída pode ser gerado usando-se uma janela de síntese de comprimento L . O método e o sistema podem usar uma janela de análise e uma janela de síntese, as quais são diferentes de cada outra. A janela de análise e a de síntese podem ser diferentes com respeito ao seu formato, seu comprimento, ao número de coeficientes definindo as janelas e/ou os valores dos coeficientes definindo as janelas. Ao se

fazer isto, graus adicionais de liberdade na seleção das janelas de análise e de síntese podem ser obtidos, de modo que uma descontinuidade do sinal de saída transposto possa ser reduzida ou removida.

[00035] De acordo com um outro aspecto, a janela de análise e a janela de síntese são biortogonais uma com respeito à outra. A janela de síntese $v_s(n)$ pode ser dada por:

$$v_s(n) = c \frac{v_a(n)}{s(n(\text{mod } \Delta t_s))}, \quad 0 \leq n < L_s$$

com c sendo uma constante, $v_a(n)$ sendo a janela de análise (311), Δt_s sendo um passo de tempo da janela de síntese e $s(n)$ sendo dado por:

$$s(m) = \sum_{i=0}^{L_s(\Delta t_s-1)} v_a^2(m + \Delta t_s i), \quad 0 \leq m < \Delta t_s.$$

[00036] O passo de tempo da janela de síntese Δt_s tipicamente corresponde ao passo de síntese S_s .

[00037] De acordo com um aspecto adicional, a janela de análise pode ser selecionada de modo que sua transformada z tenha zeros duplos no círculo unitário. Preferencialmente, a transformada z da janela de análise apenas tem zeros duplos no círculo unitário. A título de exemplo, a janela de análise pode ser uma janela de seno ao quadrado. Em um outro exemplo, a janela de análise de comprimento L pode ser determinada pela convolução de duas janelas de seno de comprimento L , produzindo-se uma janela de seno ao quadrado de comprimento $2L-1$. Em uma etapa adicional, um zero é anexado à janela de seno ao quadrado, produzindo-se uma janela de base de comprimento $2L$. Eventualmente, a janela de base pode ser reamostrada usando-se interpolação linear, desse modo se produzindo uma janela simétrica par de comprimento L como a janela de análise.

[00038] Os métodos e sistemas descritos no presente documento

podem ser implementados como um software, um firmware e/ou um hardware. Certos componentes podem ser implementados, por exemplo, como um software rodando em um processador de sinal digital ou um microprocessador. Outros componentes podem ser implementados, por exemplo, como um hardware ou como circuitos integrados específicos de aplicação. Os sinais encontrados nos métodos e sistemas descritos podem ser armazenados em meios, tais como uma memória de acesso randômico ou meios de armazenamento óticos. Eles podem ser transferidos através de redes, tais como redes de rádio, redes por satélite, redes sem fio ou redes com fio, por exemplo, a internet. Os dispositivos típicos que fazem uso do método e do sistema descritos no presente documento são caixas adaptadoras ou outro equipamento de instalações prediais de consumidor, os quais decodifiquem sinais de áudio. No lado de codificação, o método e o sistema podem ser usados em estações de difusão, por exemplo, sistemas de extremidade de entrada de vídeo ou TV.

[00039] Deve ser notado que as modalidades e os aspectos da invenção descritos aqui neste documento podem ser combinados arbitrariamente. Em particular, deve ser notado que os aspectos destacados para o sistema também são aplicáveis ao método correspondente englobado pela presente invenção. Mais ainda, deve ser notado que a exposição da invenção também cobre outras combinações de concretização além das combinações de concretização as quais são explicitamente dadas pelas referências prévias nas concretizações, isto é, as concretizações e suas características técnicas podem ser combinadas em qualquer ordem e qualquer formação.

BREVE DESCRIÇÃO DOS DESENHOS

[00040] A presente invenção será descrita, agora, a título de exemplos ilustrativos, não limitando o escopo ou o espírito da invenção, com referência aos desenhos associados, nos quais:

a figura 1 ilustra um Dirac em uma posição em particular conforme aparece nas janelas de análise e de síntese de um transpositor harmônico;

a figura 2 ilustra um Dirac em uma posição diferente, conforme aparece nas janelas de análise e de síntese de um transpositor harmônico;

a figura 3 ilustra um Dirac para a posição da figura 2, conforme ele aparece de acordo com a presente invenção;

a figura 4 ilustra a operação de um decodificador de áudio melhorado de HFR;

a figura 5 ilustra a operação de um transpositor harmônico usando várias ordens;

a figura 6 ilustra a operação de um transpositor harmônico de domínio de frequência (FD);

a figura 7 mostra uma sucessão de janela de análise e de síntese;

a figura 8 ilustra janelas de análise e de síntese em passos diferentes;

a figura 9 ilustra o efeito de reamostragem do passo de síntese de janelas;

as figura 10 e 11 ilustram modalidades de um codificador e de um decodificador, respectivamente, usando os esquemas de transposição harmônica destacados no presente documento; e

a figura 12 ilustra uma modalidade de uma unidade de transposição mostrada nas figuras 10 e 11.

DESCRIÇÃO DETALHADA

[00041] As modalidades descritas abaixo são meramente ilustrativas para os princípios da presente invenção para uma Transposição Harmônica Melhorada. É entendido que modificações e variações dos arranjos e detalhes descritos aqui serão evidentes para outros versados

na técnica. Portanto, há a intenção de ser limitado apenas pelo escopo das concretizações de patente iminentes e não pelos detalhes específicos apresentados a título de descrição e de explicação das modalidades aqui.

[00042] A seguir, os princípios de transposição harmônica no domínio de frequência e os melhoramentos propostos conforme ensinado pela presente invenção são destacados. Um componente-chave da transposição harmônica é uma distensão de tempo por um fator de transposição T inteiro, o que preserva a frequência de senoides. Em outras palavras, a transposição harmônica é baseada na distensão no tempo do sinal subjacente pelo fator T . A distensão no tempo é realizada de modo que as frequências de senoides, as quais compõem o sinal de entrada sejam mantidas. Essa distensão no tempo pode ser realizada usando-se um codificador de voz de fase. O codificador de voz de fase é baseado em uma representação de domínio de frequência fornecida por um banco de filtro de DFT em janela com uma janela de análise $v_a(n)$ e uma janela de síntese $v_s(n)$. Essa transformada de análise / síntese também é referida como uma Transformada de Fourier de tempo curto (STFT).

[00043] Uma transformada de Fourier de tempo curto é realizada em um sinal de entrada de domínio de tempo para a obtenção de uma sucessão de quadros espectrais sobrepostos. De modo a se minimizarem possíveis efeitos de banda lateral, janelas apropriadas de análise / síntese, por exemplo, janelas gaussianas, janelas de cosseno, janelas de Hamming, janelas de Hann, janelas retangulares, janelas de Bartlett, janelas de Blackman e outras devem ser selecionadas. O atraso de tempo no qual todo quadro espectral é capturado a partir do sinal de entrada é referido como o tamanho de salto ou passo. A STFT do sinal de entrada é referida como o estágio de análise e leva a uma representação de domínio de frequência do sinal de entrada. A

representação de domínio de frequência compreende uma pluralidade de sinais de sub-banda, onde cada sinal de sub-banda representa um certo componente de frequência do sinal de entrada.

[00044] A representação de domínio de frequência do sinal de entrada então pode ser processada de uma forma desejada. Para fins de distensão de tempo do sinal de entrada, cada sinal de sub-banda pode ser distendido no tempo, por exemplo, por um atraso das amostras de sinal de sub-banda. Isto pode ser obtido pelo uso de um tamanho de salto de síntese, o qual é maior do que o tamanho de salto de análise. O sinal de domínio de tempo pode ser reconstruído pela realização de uma transformada de Fourier (rápida) inversa em todos os quadros, seguida por uma acumulação sucessiva dos quadros. Esta operação do estágio de síntese é referida como uma operação de superposição – adição. O sinal de saída resultante é uma versão distendida no tempo do sinal de entrada compreendendo as mesmas componentes de frequência que o sinal de entrada. Em outras palavras, o sinal de saída resultante tem a mesma composição espectral que o sinal de entrada, mas é mais lento do que o sinal de entrada, isto é, sua progressão é distendida no tempo.

[00045] A transposição para frequências mais altas então pode ser obtida subsequentemente, ou de uma maneira integrada, através de uma redução da amostragem dos sinais distendidos. Como resultado, o sinal transposto tem a extensão no tempo do sinal inicial, mas compreende componentes de frequência os quais são deslocados para cima por um fator de transposição predefinido.

[00046] Em termos matemáticos, o codificador de voz de fase pode ser descrito conforme se segue. Um sinal de entrada $x(t)$ é amostrado a uma taxa de amostragem R para a produção do sinal de entrada discreto $x(n)$. Durante o estágio de análise, uma STFT é determinada para o sinal de entrada $x(n)$ em instantes de tempo de análise em

particular t_a^k para valores sucessivos k. Os instantes de tempo de análise preferencialmente são selecionados uniformemente através de $t_a^k = k \cdot \Delta t_a$, onde Δt_a é o fator de salto de análise ou o passo de análise.

Em cada um destes instantes de tempo de análise t_a^k , uma transformada de Fourier é calculada por uma porção em janela do sinal original $x(n)$, em que a janela de análise $v_a(t)$ é centralizada em torno de t_a^k , isto é, $v_a(t - t_a^k)$. Esta porção em janela do sinal de entrada $x(n)$ é referida como um quadro. O resultado é a representação de STFT do sinal de entrada $x(n)$, a qual pode ser denotada como:

$$X(t_a^k, \Omega_m) = \sum_{n=-\infty}^{\infty} v_a(n - t_a^k) x(n) \exp(-j\Omega_m n),$$

onde $\Omega_m = 2\pi m/M$ é a frequência central do m-ésimo sinal de sub-banda da análise de STFT e M é o tamanho da transformada de Fourier discreta (DFT). Na prática, a função de janela $v_a(n)$ tem um intervalo de tempo limitado, isto é, cobre apenas um número limitado de amostras L, o que é tipicamente igual ao tamanho M da DFT. Como uma consequência, a soma acima tem um número finito de termos. Os sinais de sub-banda $X(t_a^k, \Omega_m)$ são ambos uma função do tempo, através do índice k, e da frequência, através da frequência central de sub-banda Ω_m .

[00047] O estágio de síntese pode ser realizado nos instantes de tempo de síntese t_s^k , os quais de forma típica são uniformemente distribuídos de acordo com $t_s^k = k \cdot \Delta t_s$, onde Δt_s é o fator de salto de síntese ou o passo de síntese. Em cada um destes instantes de tempo de síntese, um sinal de tempo curto $y_k(n)$ é obtido por uma transformação inversa de Fourier do sinal de sub-banda de STFT $Y(t_s^k, \Omega_m)$, o qual pode ser idêntico a $X(t_a^k, \Omega_m)$, nos instantes de tempo

de síntese t_s^k . Contudo, tipicamente, os sinais de sub-banda de STFT são modificados, por exemplo, distendidos no tempo e/ou de fase modulada, e/ou de amplitude modulada, de modo que o sinal de sub-banda de análise $X(t_a^k, \Omega_m)$ difira do sinal de sub-banda de síntese $Y(t_s^k, \Omega_m)$. Em uma modalidade preferida, os sinais de sub-banda de STFT são de fase modulada, isto é, a fase dos sinais de sub-banda de STFT é modificada. O sinal de síntese de termo curto $y_k(n)$ pode ser denotado como:

$$y_k(n) = \frac{1}{M} \sum_{m=0}^{M-1} Y(t_s^k, \Omega_m) \exp(j\Omega_m n).$$

[00048] O sinal de síntese de termo curto $y_k(n)$ pode ser visto como um componente do sinal de saída geral $y(n)$ compreendendo os sinais de sub-banda de síntese $Y(t_s^k, \Omega_m)$ para $m = 0, \dots, M - 1$, no instante de tempo de síntese t_s^k . Isto é, o sinal de termo curto $y_k(n)$ é a DFT inversa para um quadro de sinal específico. O sinal de saída geral $y(n)$ pode ser obtido pela superposição e pela adição de sinais de tempo curto em janela $y_k(n)$ em todos os instantes de tempo de síntese t_s^k . Isto é, o sinal de saída $y(n)$ pode ser denotado como:

$$y(n) = \sum_{k=-\infty}^{\infty} v_s(n - t_s^k) y_k(n - t_s^k),$$

onde $v_s(n - t_s^k)$ é a janela de síntese centralizada em torno do instante de tempo de síntese t_s^k . Deve ser notado que a janela de síntese tipicamente tem um número limitado de amostras L , de modo que a soma mencionada acima apenas compreenda um número limitado de termos.

[00049] A seguir, a implementação de distensão no tempo no domínio de frequência é destacada. Um ponto de começo adequado de modo a se descreverem os aspectos do distensor no tempo é considerar

o caso em que $T = 1$, isto é, o caso em que o fator de transposição T equivale a 1 e onde nenhuma distensão ocorre. Assumindo que o passo de tempo de análise Δt_a e o passo de tempo de síntese Δt_s do banco de filtro de DFT sejam iguais, isto é, $\Delta t_a = \Delta t_s = \Delta t$, o efeito combinado de análise seguida pela síntese é aquele de uma modulação de amplitude com uma função periódica em Δt :

$$K(n) = \sum_{k=-\infty}^{\infty} q(n - k\Delta t), \quad (1)$$

onde $q(n) = v_a(n) v_s(n)$ é o produto pontual das duas janelas, isto é, o produto pontual da janela de análise e da janela de síntese. É vantajoso escolher a janela de modo que $K(n) = 1$ ou um outro valor constante, uma vez que, então, o banco de filtro de DFT em janela obtém uma reconstrução perfeita. Se a janela de análise $v_a(n)$ for dada, e se a janela de análise for de duração suficientemente longa, se comparada com o passo Δt , poder-se-á obter uma reconstrução perfeita pela escolha da janela de síntese de acordo com:

$$v_s(n) = v_a(n) \left(\sum_{k=-\infty}^{\infty} (v_a(n - k \cdot \Delta t))^2 \right)^{-1}. \quad (2)$$

[00050] Para $T > 1$, isto é, para um fator de transposição maior do que 1, uma distensão de tempo pode ser obtida pela realização da análise no passo $\Delta t_a = \Delta t/T$, ao passo que o passo de síntese é mantida em $\Delta t_s = \Delta t$. Em outras palavras, uma distensão de tempo por um fator T pode ser obtida pela aplicação de um fator de salto ou um passo no estágio de síntese. Conforme pode ser visto a partir das fórmulas providas acima, o uso de um passo de síntese a qual é T vezes maior do que o passo de análise deslocará os sinais de síntese de termo curto $y_k(n)$ por intervalos T vezes maiores na operação de superposição – adição. Isto eventualmente resultará em uma distensão no tempo do sinal de saída $y(n)$.

[00051] Deve ser notado que a distensão no tempo pelo fator T pode

envolver, adicionalmente, uma multiplicação de fase por um fator T entre a análise e a síntese. Em outras palavras, uma distensão no tempo por um fator T envolve uma multiplicação de fase por um fator T dos sinais de sub-banda.

[00052] A seguir, é destacado como a operação de distensão no tempo descrita acima pode ser traduzida em uma operação de transposição harmônica. A modificação de escala de passo ou transposição harmônica pode ser obtida pela realização de uma conversão de taxa de amostra do sinal de saída distendido no tempo $y(n)$. Para a realização de uma transposição harmônica por um fator T , um sinal de saída $y(n)$ o qual é uma versão distendida no tempo pelo fator T do sinal de entrada $x(n)$ pode ser obtido usando-se o método de codificação de voz de fase descrito acima. A transposição harmônica então pode ser obtida pela redução da amostragem do sinal de saída $y(n)$ por um fator T ou pela conversão da taxa de amostragem de R para TR . Em outras palavras, em vez de interpretar o sinal de saída $y(n)$ como tendo a mesma taxa de amostragem que o sinal de entrada $x(n)$, mas de duração de T vezes, o sinal de saída $y(n)$ pode ser interpretado como sendo da mesma duração, mas de T vezes a taxa de amostragem. A redução de amostragem subsequente de T então pode ser interpretada como tornando a taxa de amostragem de saída igual à taxa de amostragem de entrada, de modo que os sinais eventualmente possam ser adicionados. Durante estas operações, deve-se ter cuidado quando da redução da amostragem do sinal transposto, de modo que nenhuma descontinuidade ocorra.

[00053] Assumindo que o sinal de entrada $x(n)$ seja uma senoide e quando se assume uma janela de análise simétrica $v_a(n)$, o método de distensão no tempo com base no codificador de voz de fase descrito acima funcionará perfeitamente para valores ímpares de T , e resultará em uma versão distendida no tempo do sinal de entrada $x(n)$ tendo a

mesma frequência. Em combinação com uma redução de amostragem subsequente, uma senoide $y(n)$ com uma frequência a qual é T vezes a frequência do sinal de entrada $x(n)$ será obtida.

[00054] Para valores pares de T , o método de distensão no tempo / transposição harmônica destacado acima será mais aproximado, uma vez que lobos laterais de valor negativo da resposta de frequência da janela de análise $v_a(n)$ serão reproduzidos com fidelidade diferente pela multiplicação de fase. Os lobos de lado negativo tipicamente vêm do fato de que a maioria das janelas práticas (ou filtros de protótipo) tem numerosos zeros discretos localizados no círculo unitário, resultando em deslocamentos de fase de 180 graus. Quando da multiplicação dos ângulos de fase usando fatores de transposição pares, os deslocamentos de fase tipicamente serão transladados para 0 (em vez de múltiplos de 360) grau, dependendo do fator de transposição usado. Em outras palavras, quando se usam fatores de transposição pares, os deslocamentos de fase se anulam. Isto tipicamente dará origem a uma descontinuidade no sinal de saída transposto $y(n)$. Um cenário particularmente desvantajoso pode surgir quando uma senoide estiver localizada em uma frequência correspondente ao topo do primeiro lobo lateral do filtro de análise. Dependendo da rejeição deste lobo na resposta de magnitude, a descontinuidade será mais ou menos audível no sinal de saída. Deve ser notado que, para fatores pares T , uma diminuição no passo geral Δt tipicamente melhora o desempenho do distensor no tempo à custa de uma complexidade computacional mais alta.

[00055] Na EP0940015B1 / no WO09/57436 intitulado "Source coding enhancement using spectral band replication", o qual é incorporado como referência, foi descrito um método sobre como se evitar que uma descontinuidade emerja de um transpositor harmônico, quando se usarem fatores de transposição pares. Este método,

denominado travamento de fase relativa, avalia a diferença de fase relativa entre canais adjacentes, e determina se uma senoide é de fase invertida em qualquer canal. A detecção é realizada pelo uso da equação (32) da EP0940015B1. Os canais detectados como de fase invertida são corrigidos, após os ângulos de fase serem multiplicados pelo fator de transposição real.

[00056] A seguir, um novo método para se evitar uma descontinuidade quando se usam fatores de transposição T pares e/ou ímpares é descrito. Ao contrário do método de travamento de fase relativa do EP0940015B1, este método não requer a detecção e a correção de ângulos de fase. A solução nova para o problema acima faz uso de janelas de transformada de análise e de síntese que não são idênticas. No caso de uma reconstrução perfeita (PR), isto corresponde a uma transformada / um banco de filtro biortogonal, em vez de uma transformada / um banco de filtro ortogonal.

[00057] Para a obtenção de uma transformada biortogonal dada uma certa janela de análise $v_a(n)$, a janela de síntese $v_s(n)$ é escolhida para seguir:

$$\sum_{i=0}^{L/(\Delta t_s)-1} v_a(m + \Delta t_s, i) v_s(m + \Delta t_s, i) = c, \quad 0 \leq m < \Delta t_s,$$

onde c é uma constante, Δt_s é o passo de tempo de síntese e L é a extensão da janela. Se a sequência $s(n)$ for definida como:

$$s(m) = \sum_{i=0}^{L/(\Delta t_s)-1} v_a^2(m + \Delta t_s, i), \quad 0 \leq m < \Delta t_s,$$

isto é, $v_a(n) = v_s(n)$ for usado para uma formação de janela de análise e de síntese, então, a condição para uma transformada ortogonal será:

$$s(m) = c, \quad 0 \leq m < \Delta t_s.$$

[00058] Contudo, a seguir, uma outra sequência $w(n)$ é introduzida, onde $w(n)$ é uma medida sobre quanto a janela de síntese $v_s(n)$ se

desvia da janela de análise $v_a(n)$, isto é, quanto a transformada biortogonal difere do caso ortogonal. A sequência $w(n)$ é dada por:

$$w(n) = \frac{v_s(n)}{v_a(n)}, \quad 0 \leq n < L.$$

[00059] A condição para uma reconstrução perfeita, então, é dada por:

$$\sum_{i=0}^{L/(\Delta t_s)-1} v_a^2(m + \Delta t_s, i) w(m + \Delta t_s, i) = c, \quad 0 \leq m < \Delta t_s.$$

[00060] Para uma solução possível, $w(n)$ poderia ser restrita a ser periódica com o passo de tempo de síntese Δt_s , isto é, $w(n) = w(n + \Delta t_s i)$, $\forall i, n$. Então, obtém-se:

$$\sum_{i=0}^{L/(\Delta t_s)-1} v_a^2(m + \Delta t_s, i) w(m + \Delta t_s, i) = w(m) \sum_{i=0}^{L/(\Delta t_s)-1} v_a^2(m + \Delta t_s, i) = w(m) s(m) = c, \\ 0 \leq m < \Delta t_s.$$

[00061] A condição na janela de síntese $v_s(n)$, daí, é:

$$v_s(n) = w(n \bmod \Delta t_s) v_a(n) = c \frac{v_a(n)}{s(n \bmod \Delta t_s)}, \quad 0 \leq n < L.$$

[00062] Pela derivação da janela de síntese $v_s(n)$, conforme destacado acima, uma liberdade muito maior quando do projeto da janela de análise $v_a(n)$ é provida. Esta liberdade adicional pode ser usada para o projeto de um par de janelas de análise / síntese, o que não exhibe uma descontinuidade do sinal transposto.

[00063] Para a obtenção de um par de janelas de análise / síntese que suprima uma descontinuidade para fatores de transposição pares, várias modalidades serão destacadas a seguir. De acordo com uma primeira modalidade, as janelas ou os filtros de protótipo são feitos longos o bastante para a atenuação do nível do primeiro lobo lateral na resposta de frequência abaixo de um certo nível de "descontinuidade". O passo de tempo de análise Δt_a neste caso será apenas uma fração (pequena) do comprimento de janela L . Isto tipicamente resultará em um espalhamento de transientes, por exemplo, em sinais percussivos.

[00064] De acordo com uma segunda modalidade, a janela de análise $v_a(n)$ é escolhida para ter zeros duplos no círculo unitário. A resposta de fase resultante de um zero duplo é um deslocamento de fase de 360 graus. Estes deslocamentos de fase são retidos, quando os ângulos de fase são multiplicados pelos fatores de transposição, independentemente de os fatores de transposição serem ímpares ou pares. Quando um filtro de análise apropriado e atenuado $v_a(n)$, tendo zeros duplos no círculo unitário é obtido, a janela de síntese é obtida a partir das equações destacadas acima.

[00065] Em um exemplo da segunda modalidade, o filtro / a janela de análise $v_a(n)$ é a "janela de seno ao quadrado", isto é, a janela seno:

$$v(n) = \sin\left(\frac{\pi}{L}(n+0.5)\right), 0 \leq n < L$$

convoluta com ela mesma como $v_a(n) = v(n) \otimes v(n)$. Contudo, deve ser notado que o filtro / a janela resultante $v_a(n)$ será simétrica ímpar com comprimento, $L_a = 2L - 1$, isto é, um número ímpar de coeficientes de filtro / janela. Quando um filtro / uma janela com comprimento par é mais apropriado, em particular um filtro simétrico par, o filtro pode ser obtido primeiramente pela convolução de duas janelas de seno de comprimento L . Então, um zero é anexado ao fim do filtro resultante. Subsequentemente, o filtro de $2L$ de comprimento é reamostrado usando-se uma interpolação linear até um filtro simétrico par de comprimento L , o qual ainda tem zeros duplos no círculo unitário.

[00066] Em geral, foi destacado como um par de janelas de análise e de síntese pode ser selecionado de modo que uma descontinuidade no sinal de saída transposto possa ser evitada ou reduzida significativamente. O método é particularmente relevante quando se usam fatores de transposição pares.

[00067] Um outro aspecto a considerar no contexto de transpositores harmônicos é um envolvimento de fase. Deve ser notado que, ao passo

que grande cuidado deve ser tomado com relação a questões de desenvolvimento em codificadores de voz de fase de finalidade geral, o transpositor harmônico tem operações de fase definidas de forma não ambígua, quando fatores de transposição inteiros T são usados. Assim, em modalidades preferidas a ordem de transposição T é um valor inteiro. Caso contrário, técnicas de desenvolvimento de fase poderiam ser aplicadas, onde um desenvolvimento de fase é um processo por meio do qual o incremento de fase entre dois quadros consecutivos é usado para a estimativa da frequência instantânea de uma quase senoide em cada canal.

[00068] Ainda um outro aspecto a considerar, quando se lida com a transposição de sinais de áudio e/ou de voz, é o processamento de seções de sinal estacionárias e/ou transientes. Tipicamente, de modo a se ser capaz de transpor sinais de áudio estacionários sem artefatos intermodulação, a resolução de frequência do banco de filtro de DFT tem que ser bem alta e, portanto, as janelas são comparadas longamente com transientes nos sinais de entrada $x(n)$, notadamente, sinais de áudio e/ou de voz. Como resultado, o transpositor tem uma resposta transiente ruim. Contudo, conforme será apreciado a seguir, este problema pode ser resolvido por uma modificação do projeto de janela, o tamanho de transformada e os parâmetros de passo de tempo. Daí, diferentemente de muitos métodos do estado da técnica para melhoramento da resposta transiente de um codificador de voz de fase, a solução proposta não se baseia em qualquer operação adaptativa de sinal, tal como uma detecção de transiente.

[00069] A seguir, a transposição harmônica de sinais transientes usando-se codificadores de voz é destacada. Como um ponto de começo, um sinal transiente de protótipo, um pulso de Dirac discreto no tempo em um instante de tempo $t = t_0$,

$$\delta(t - t_0) = \begin{cases} 1, & t = t_0 \\ 0, & t \neq t_0 \end{cases}$$

é considerado. A transformada de Fourier de um pulso de Dirac como esse tem magnitude unitária e uma fase linear com uma inclinação proporcional a t_0 :

$$X(\Omega_m) = \sum_{n=-\infty}^{\infty} \delta(n - t_0) \exp(-j\Omega_m n) = \exp(-j\Omega_m t_0).$$

[00070] Essa transformada de Fourier pode ser considerada como o estágio de análise do codificador de voz de fase descrito acima, onde uma janela de análise plana $v_a(n)$ de duração infinita é usada. De modo a se gerar um sinal de saída $y(n)$ o qual é distendido no tempo por um fator T , isto é, um pulso de Dirac $\delta(t - Tt_0)$ no instante de tempo $t = Tt_0$, a fase dos sinais de sub-banda de análise deve ser multiplicada pelo fator T , de modo a se obter o sinal de sub-banda de síntese $Y(\Omega_m) = \exp(-j\Omega_m Tt_0)$, o qual produz o pulso de Dirac $\delta(t - Tt_0)$ como uma saída de uma transformada de Fourier inversa.

[00071] Isto mostra que a operação da multiplicação de fase dos sinais de sub-banda de análise por um fator T leva ao deslocamento de tempo desejado de um pulso de Dirac, isto é, de um sinal de entrada transiente. Deve ser notado que para sinais transientes mais realistas compreendendo mais de uma amostra não nula, as operações adicionais de distensão no tempo dos sinais de sub-banda de análise por um fator T deve ser realizada. Em outras palavras, diferentes tamanhos de salto devem ser usados na análise e no lado de síntese.

[00072] Contudo, deve ser notado que as considerações acima se referem a um estágio de análise / síntese usando janelas de análise e de síntese de comprimentos infinitos. De fato, um transpositor teórico com uma janela de duração infinita proporcionaria a distensão correta de um pulso de Dirac $\delta(t - t_0)$. Para uma análise em janela de duração finita, a situação é embaralhada pelo fato de cada bloco de análise ser para ser interpretado como um intervalo de período de um sinal periódico com um período igual ao tamanho da DFT.

[00073] Isto é ilustrado na figura 1, a qual mostra a análise e a síntese 100 de um pulso de Dirac $\delta(t - t_0)$. A parte superior da figura 1 mostra a entrada do estágio de análise 110 e a parte inferior da figura 1 mostra a saída do estágio de síntese 120. Os gráficos superior e inferior representam o domínio de tempo. A janela de análise estilizada 111 e a janela de síntese 121 são descritas como janelas triangulares (Barlett). O pulso de entrada $\delta(t - t_0)$ 112 no instante no tempo $t = t_0$ é descrito no gráfico de topo 110 como uma seta vertical. É assumido que o bloco de transformada DFT é de um tamanho de $M = L$, isto é, o tamanho da transformada DFT é escolhido para ser igual ao tamanho das janelas. A multiplicação de fase dos sinais de sub-banda pelo fator T produzirá a análise de DFT de um pulso de Dirac $\delta(t - Tt_0)$ em $t = Tt_0$, embora periodizado para um trem de pulso de Dirac com período L . Isto é devido ao comprimento finito da janela aplicada e à transformada de Fourier. O trem de pulso periodizado com período L é descrito pelas setas tracejadas 123, 124 no gráfico inferior.

[00074] Em um sistema de mundo real, em que ambas as janelas de análise e de síntese são de comprimento finito, o trem de pulso realmente contém apenas uns poucos pulsos (dependendo do fator de transposição), um pulso principal, isto é, o termo desejado, uns poucos pré-pulsos, e uns poucos pós-pulsos, isto é, os termos indesejados. Os pré- e pós-pulsos emergem porque a DFT é periódica (com L). Quando um pulso está localizado em uma janela de análise, de modo que a fase complexa fique envolvida quando multiplicada por T (isto é o pulso é deslocado para fora do fim da janela e envolve de volta até o começo), um pulso indesejado emerge. Os pulsos indesejados podem ter ou não a mesma polaridade que o pulso de entrada, dependendo da localização da janela de análise e do fator de transposição.

[00075] Isto pode ser visto matematicamente quando da transformada do pulso de Dirac $\delta(t - t_0)$ situado no intervalo $-L/2 \leq t_0 <$

$L/2$ usando-se uma DFT com comprimento L centralizado em torno de $t = 0$,

$$X(\Omega_m) = \sum_{n=-L/2}^{L/2-1} \delta(n - t_0) \exp(-j\Omega_m n) = \exp(-j\Omega_m t_0) \dots$$

[00076] Os sinais de sub-banda de análise são de fase multiplicada por um fator T para a obtenção dos sinais de sub-banda de síntese $Y(\Omega_m) = \exp(-j\Omega_m T t_0)$. Então, a DFT inversa é aplicada para a obtenção do sinal de síntese periódico:

$$y(n) = \frac{1}{L} \sum_{m=-L/2}^{L/2-1} \exp(-j\Omega_m T t_0) \exp(j\Omega_m n) = \sum_{k=-\infty}^{\infty} \delta(n - T t_0 + kL).$$

isto é, um trem de pulso de Dirac com período L .

[00077] No exemplo da figura 1, a janela de síntese usa uma janela finita $v_s(n)$ 121. A janela de síntese finita 121 captura o pulso desejado $\delta(t - T t_0)$ em $t = T t_0$, o qual é descrito como uma seta sólida 122 e cancela as outras contribuições as quais são mostradas como setas tracejadas 123, 124.

[00078] Conforme o estágio de análise e de síntese se move ao longo do eixo de tempo de acordo com o fator de salto ou o passo de tempo Δt , o pulso $\delta(t - t_0)$ 112 terá uma outra posição em relação ao centro da respectiva janela de análise 111. Conforme destacado acima, a operação para a obtenção da distensão no tempo consiste no movimento do pulso 112 para T vezes sua posição em relação ao centro da janela. Desde que esta posição esteja na janela 121, esta operação de distensão no tempo garantirá que todas as contribuições se somem a um pulso sintetizado distendido no tempo único $\delta(t - T t_0)$ em $t = T t_0$.

[00079] Contudo, ocorre um problema para a situação da figura 2, onde o pulso $\delta(t - t_0)$ 212 se move mais para fora em direção à borda do bloco de DFT. A figura 2 ilustra uma configuração de análise / síntese similar 200 como a figura 1. O gráfico superior 210 mostra a entrada no estágio de análise e a janela de análise 211, e o gráfico inferior 220

ilustra a saída do estágio de síntese e a janela de síntese 221. Quando da distensão no tempo do pulso de Dirac de entrada 212 por um fator T , o pulso de Dirac distendido no tempo 222, isto é, $\delta(t - Tt_0)$ está fora da janela de síntese 221. Ao mesmo tempo, um outro pulso de Dirac 224 do trem de pulso, isto é, $\delta(t - Tt_0 + L)$ no instante de tempo $t = Tt_0 - L$, é capturado pela janela de síntese. Em outras palavras, o pulso de Dirac de entrada 212 não é atrasado para um instante de tempo T vezes posterior, mas é movido para frente para um instante de tempo que fica antes do pulso de Dirac de entrada 212. O efeito final sobre o sinal de áudio é a ocorrência de um pré-eco a uma distância no tempo da escala das janelas de transpositor bastante longas, isto é, em um instante de tempo $t = Tt_0 - L$, o qual é $L - (T - 1)t_0$ anterior ao pulso de Dirac de entrada 212.

[00080] O princípio da solução proposta pela presente invenção é descrito com referência à figura 3. A figura 3 ilustra um cenário de análise / síntese 300 similar à figura 2. O gráfico superior 310 mostra a entrada no estágio de análise com a janela de análise 311, e o gráfico inferior 320 mostra a saída do estágio de síntese com a janela de síntese 321. A ideia básica da invenção é adaptar o tamanho de DFT de modo a se evitem pré-ecos. Isto pode ser obtido pela regulação do tamanho M da DFT, de modo que nenhuma imagem de pulso de Dirac indesejada a partir do trem de pulso resultante seja capturada pela janela de síntese. O tamanho da transformada DFT 301 é aumentado para $M = FL$, onde L é o comprimento da função de janela 302 e o fator F é um fator de sobreamostragem de domínio de frequência. Em outras palavras, o tamanho da transformada DFT 301 é selecionado para ser maior do que o tamanho de janela 302. Em particular, o tamanho da transformada DFT 301 pode ser selecionado para ser maior do que o tamanho de janela 302 da janela de síntese. Devido ao comprimento aumentado 301 da transformada DFT, o período do trem de pulso

compreendendo os pulsos de Dirac 322, 324 é FL. Pela seleção de um valor suficientemente grande de F, isto é, pela seleção de um fator de sobreamostragem de domínio de frequência suficientemente grande, contribuições indesejadas para a distensão de pulso podem ser canceladas. Isto é mostrado na figura 3, onde o pulso de Dirac 324 no instante $t = Tt_0 - FL$ fica fora da janela de síntese 321. Portanto, o pulso de Dirac 324 não é capturado pela janela de síntese 321 e, como uma consequência, pré-ecos podem ser evitados.

[00081] Deve ser notado que em uma modalidade preferida a janela de síntese e a janela de análise têm comprimentos "nominais" iguais. Contudo, quando se usa uma reamostragem implícita do sinal de saída pelo descarte ou pela inserção de amostras nas bandas de frequência da transformada ou do banco de filtro, o tamanho de janela de síntese tipicamente será diferente do tamanho de análise, dependendo da reamostragem ou do fator de transposição.

[00082] O valor mínimo de F, isto é, o fator de sobreamostragem de domínio de frequência mínimo pode ser deduzido a partir da figura 3. A condição para a não captura de imagens de pulso de Dirac indesejadas pode ser formulada conforme se segue: para qualquer pulso de entrada $\delta(t - t_0)$ na posição $t = t_0 < L/2$, isto é, para qualquer pulso de entrada compreendido na janela de análise 311, a imagem indesejada $\delta(t - Tt_0 + FL)$ no instante $t = Tt_0 - FL$ deve estar localizada na borda esquerda da janela de síntese em $t = -L/2$. De forma equivalente, a condição $T L/2 - FL \leq -L/2$ deve ser atendida, o que leva à regra:

$$F \geq \frac{T+1}{2}. \quad (3)$$

[00083] Conforme pode ser visto a partir da fórmula (3), o fator de sobreamostragem de domínio de frequência mínimo F é uma função do fator de transposição / de distensão no tempo T. Mais especificamente, o fator de sobreamostragem de domínio de frequência mínimo F é

proporcional ao fator de transposição / de distensão no tempo T.

[00084] Pela repetição da linha de pensamento acima para o caso em que as janelas de análise e de síntese têm comprimentos diferentes, obtém-se uma forma mais geral. Sejam L_A e L_S os comprimentos das janelas de análise e de síntese, respectivamente, e seja M o tamanho de DFT empregado. A fórmula de extensão da regra (3) então é:

$$M \geq \frac{TL_A + L_S}{2}. \quad (4)$$

[00085] Que esta regra de fato é uma extensão de (3) pode ser verificado pela inserção de $M = FL$ e $L_A = L_S = L$ em (4) e dividindo-se por L em ambos os lados da equação resultante.

[00086] A análise acima é realizada para um modelo bem especial de um transiente, isto é, um pulso de Dirac. Contudo, o raciocínio pode ser estendido para mostrar quando se usa o esquema de distensão de tempo descrito acima, os sinais de entrada os quais têm uma envoltória espectral quase plana e os quais se anulam fora do intervalo de tempo [a, b] serão distendidos para a extração de sinais os quais são pequenos fora do intervalo [Ta, Tb]. Também pode ser checado pelo estudo de espectrogramas de sinais de áudio e/ou de fala reais que pré-ecos desaparecem nos sinais distendidos, quando a regra acima para a seleção de um fator de sobreamostragem de domínio de frequência apropriado for respeitada. Uma análise mais quantitativa também revela que pré-ecos são mais reduzidos quando se usam fatores de sobreamostragem de domínio de frequência os quais são ligeiramente inferiores para o valor imposto pela condição da fórmula (3). Isto é devido ao fato de que as funções de janela típicas $v_s(n)$ são pequenas perto de suas bordas, desse modo se atenuando pré-ecos indesejados, os quais são posicionados perto das bordas das funções de janela.

[00087] Em resumo, a presente invenção ensina uma nova forma de melhoramento da resposta transiente de transpositores harmônicos de domínio de frequência, ou distensores de tempo, pela introdução de

uma transformada sobreamostrada, onde a quantidade de sobreamostragem é uma função do fluxo de trabalho escolhido.

[00088] A seguir, a aplicação de uma transposição harmônica de acordo com a invenção em decodificadores de áudio é descrita em maiores detalhes. Um caso de uso comum para um transpositor harmônico é em um sistema de codificador – decodificador de áudio / fala empregando uma assim denominada extensão de largura de banda ou reconstrução de frequência alta (HFR). Deve ser notado que, embora uma referência possa ser feita a uma codificação de áudio, os métodos e sistemas descritos são igualmente aplicáveis a uma codificação de fala e em uma codificação unificada de fala e de áudio (USAC).

[00089] Nesses sistemas de HFR, o transpositor pode ser usado para a geração de um componente de sinal de frequência alta a partir de um componente de sinal de frequência baixa provido pelo assim denominado decodificador de núcleo. O envoltório do componente de frequência alta pode ser conformado no tempo e na frequência com base em uma informação de lado portada no fluxo de bit.

[00090] A figura 4 ilustra a operação de um decodificador de áudio melhorado de HFR. O decodificador de áudio de núcleo 401 extrai um sinal de áudio de largura de banda baixa o qual é alimentado para um aumentador de amostra 404, o qual pode ser requerido, de modo a se produzir uma contribuição de saída de áudio final na taxa de amostragem plena desejada. Esse aumento de amostragem é requerido para sistemas de taxa dupla, onde o codificador – decodificador de áudio de núcleo de banda limitada está operando à metade da taxa de amostragem de áudio externa, enquanto a parte de HFR é processada à frequência de amostragem plena. Conseqüentemente, para um sistema de taxa única, este aumentador de amostra 404 é omitido. A saída de largura de banda baixa de 401 também é enviada para o transpositor ou a unidade de transposição 402, a qual extrai um sinal

transposto, isto é, um sinal que compreende a faixa de frequência alta desejada. Este sinal transposto pode ser conformado no tempo e na frequência pelo ajustador de envoltória 403. A saída de áudio final é a soma de um sinal de núcleo de largura de banda e do sinal transposto de envoltória ajustada.

[00091] Conforme destacado no contexto da figura 4, o sinal de saída de decodificador de núcleo pode ter a amostra aumentada como uma etapa de pré-processamento por um fator 2 na unidade de transposição 402. Uma transposição por um fator T resulta em um sinal que tem T vezes o comprimento do sinal não transposto, em um caso de distensão no tempo. De modo a se obter o deslocamento de passo desejado ou a transposição de frequência para frequências T vezes mais altas, uma redução de amostragem ou conversão de taxa do sinal distendido no tempo é subsequentemente realizada. Conforme mencionado acima, esta operação pode ser obtida através do uso de diferentes passos de análise e de síntese no codificador de voz de fase.

[00092] A ordem de transposição geral pode ser obtida de formas diferentes. Uma primeira possibilidade é aumentar a amostra do sinal de saída de decodificador pelo fator 2 na entrada para o transpositor, conforme destacado acima. Nesses casos, o sinal distendido no tempo precisaria ter a amostra reduzida por um fator T, de modo a se obter o sinal de saída desejado, o qual é transposto na frequência por um fator T. Uma segunda possibilidade seria omitir a etapa de pré-processamento e realizar diretamente as operações de distensão no tempo no sinal de saída de decodificador de núcleo. Nesses casos, os sinais transpostos devem ter a amostra reduzida por um fator T/2, para a retenção do fator de aumento de amostragem global de 2, e de modo a se obter uma transposição de frequência por um fator T. Em outras palavras, o aumento de amostra do sinal de decodificador de núcleo pode ser omitido, quando da realização de uma redução de amostragem

do sinal de saída do transpositor 402 de $T/2$, em vez de T . Deve ser notado, contudo, que o sinal de núcleo ainda precisa ter a amostragem aumentada no aumentador de amostra 404, antes da combinação do sinal com o sinal transposto.

[00093] Também deve ser notado que o transpositor 402 pode usar vários fatores de transposição inteiros diferentes, de modo a gerar o componente de frequência alta. Isto é mostrado na figura 5, a qual ilustra a operação de um transpositor harmônico 501, o qual corresponde ao transpositor 402 da figura 4, compreendendo vários transpositores de diferente ordem de transposição ou fator de transposição T . O sinal a ser transposto é passado para o banco de transpositores individuais 501-2, 501-3, ..., 501- T_{\max} tendo ordens de transposição $T = 2, 3, \dots, T_{\max}$, respectivamente. De forma típica, uma ordem de transposição $T_{\max} = 4$ é suficiente para a maioria das aplicações de codificação de áudio. As contribuições dos diferentes transpositores 501-2, 501-3, ..., 501- T_{\max} são somadas em 502, para a produção da saída de transpositor combinada. Em uma primeira modalidade, esta operação de soma pode compreender a adição das contribuições individuais. Em uma outra modalidade, as contribuições recebem pesos diferentes, de modo que o efeito de adição de múltiplas contribuições a certas frequências seja mitigado. Por exemplo, a contribuição de terceira ordem pode ser adicionada com um ganho menor do que a contribuição de segunda ordem. Finalmente, a unidade de soma 502 pode adicionar as contribuições seletivamente, dependendo da frequência de saída. Por exemplo, a transposição de segunda ordem pode ser usada para uma primeira faixa de frequência alvo mais baixa, e a transposição de terceira ordem pode ser usada para uma segunda faixa de frequência alvo mais alta.

[00094] A figura 6 ilustra a operação de um transpositor harmônico, tal como um dos blocos individuais de 501, isto é, um dos transpositores

501-T da ordem de transposição T. Uma unidade de passo de análise 601 seleciona quadros sucessivos do sinal de entrada, o qual é para ser transposto. Estes quadros são sobrepostos, por exemplo, multiplicados, em uma unidade de janela de análise 602 com uma janela de análise. Deve ser notado que as operações de seleção de quadros de um sinal de entrada e multiplicação das amostras do sinal de entrada por uma função de janela de análise podem ser realizadas em uma etapa única, por exemplo, pelo uso de uma função de janela a qual é deslocada ao longo do sinal de entrada pelo passo de análise. Na unidade de transformação de análise 603, os quadros em janela do sinal de entrada são transformados no domínio de frequência. A unidade de transformação de análise 603 pode realizar, por exemplo, uma DFT. O tamanho da DFT é selecionado para ser F vezes maior do que o tamanho L da janela de análise, desse modo se gerando $M = F * L$ coeficientes complexos de domínio de frequência. Estes coeficientes complexos são alterados na unidade de processamento não linear 604, por exemplo, pela multiplicação de sua fase pelo fator de transposição T. A sequência de coeficientes complexos de domínio de frequência, isto é, os coeficientes complexos da sequência de quadros do sinal de entrada podem ser vistos como sinais de sub-banda. A combinação de unidade de passo de análise 601, unidade de janela de análise 602 e unidade de transformação de análise 603 pode ser vista como um estágio de análise combinado ou um banco de filtro de análise.

[00095] Os coeficientes alterados ou os sinais de sub-banda alterados são retransformados no domínio de tempo usando-se a unidade de transformação de síntese 605. Para cada conjunto de coeficientes complexos alterados, isto produz um quadro de amostras alteradas, isto é, um conjunto de M amostras alteradas. Usando a primeira abertura 606, amostras L podem ser extraídas a partir de cada conjunto de amostras alteradas, desse modo se produzindo um quadro

do sinal de saída. Em geral, uma sequência de quadros do sinal de saída pode ser gerada para a sequência de quadros do sinal de entrada. Esta sequência de quadros é deslocada uns com respeito aos outros pelo passo de síntese na unidade de passo de síntese 607. O passo de síntese pode ser T vezes maior do que o passo de análise. O sinal de saída é gerado na unidade de superposição – adição 608, onde os quadros deslocados do sinal de saída são sobrepostos e as amostras no mesmo instante de tempo são adicionadas. Ao atravessar o sistema acima, o sinal de entrada pode ser distendido no tempo por um fator T , isto é, o sinal de saída pode ser uma versão distendida no tempo do sinal de entrada.

[00096] Finalmente, o sinal de saída pode ser contraído no tempo usando-se a unidade de contração 609. A unidade de contração 609 pode realizar uma conversão de taxa de amostragem de ordem T , isto é, pode aumentar a taxa de amostragem do sinal de saída por um fator T , enquanto mantém o número de amostras não modificado. Isto produz um sinal de saída transposto que tem o mesmo comprimento no tempo que o sinal de entrada, mas compreendendo componentes de frequência os quais são deslocados para cima por um fator T com respeito ao sinal de entrada. A unidade de combinação 609 também pode realizar uma operação de redução de amostragem por um fator T , isto é, pode reter apenas toda T -ésima amostra enquanto descarta as outras amostras. Esta operação de redução de amostragem também pode ser acompanhada por uma operação de filtro de passa baixa. Se a taxa de amostragem geral permanecer não modificada, então, o sinal de saída transposto compreenderá componentes de frequência os quais serão deslocados para cima por um fator T com respeito aos componentes de frequência do sinal de entrada.

[00097] Deve ser notado que a unidade de contração 609 pode realizar uma combinação de conversão de taxa e redução de

amostragem. A título de exemplo, a taxa de amostragem pode ser aumentada por um fator de 2. Ao mesmo tempo, o sinal pode ter a amostragem reduzida por um fator $T/2$. Em geral, essa combinação de conversão de taxa e redução de amostragem também leva a um sinal de saída o qual é uma transposição harmônica do sinal de entrada por um fator T . Em geral, pode ser declarado que a unidade de contração 609 realiza uma combinação de conversão de taxa e/ou de redução de amostragem, de modo a se produzir uma transposição harmônica pela ordem de transposição T . Isto é particularmente útil quando na realização de uma transposição harmônica da saída de largura de banda baixa do decodificador de áudio de núcleo 401. Conforme destacado acima, essa saída de largura de banda baixa pode ter tido a amostragem reduzida por um fator de 2 no codificador e pode requerer, portanto, um aumento de amostragem na unidade de aumento de amostragem 404, antes da fusão dela com o componente de frequência alta reconstruída. Não obstante, pode ser benéfico para a redução da complexidade de computação realizar uma transposição harmônica na unidade de transposição 402 usando a saída de largura de banda baixa "sem a amostragem aumentada". Nesses casos, a unidade de contração 609 da unidade de transposição 402 pode realizar uma conversão de taxa de ordem 2 e, desse modo, implicitamente realizar a operação de aumento de amostragem requerida do componente de frequência alta. Como uma consequência, os sinais de saída transpostos de ordem T têm a amostragem reduzida na unidade de contração 609 pelo fator $T/2$.

[00098] No caso de múltiplos transpositores paralelos de ordens de transposição diferentes, tal como mostrado na figura 5, algumas operações de transformação ou de banco de filtro podem ser compartilhadas entre diferentes transpositores 501-2, 501-3, ..., 501- T_{\max} . O compartilhamento de operações de banco de filtro deve ser feito

preferencialmente para a análise, de modo a se obterem implementações mais efetivas de unidades de transposição 402. Deve ser notado que uma forma preferida de reamostragem das saídas de diferentes transpositores é descartar intervalos de DFT ou canais de sub-banda antes do estágio de síntese. Desta forma, os filtros de reamostragem podem ser omitidos e a complexidade pode ser reduzida quando se realiza uma DFT inversa / um banco de filtro de síntese de tamanho menor.

[00099] Conforme recém-mencionado, a janela de análise pode ser comum aos sinais de diferentes fatores de transposição. Quando se usa uma janela de análise comum, um exemplo do passo de janelas 700 aplicada ao sinal de banda baixa é descrito na figura 7. A figura 7 mostra um passo de janelas de análise 701, 702, 703 e 704, as quais são deslocadas umas com respeito às outras pelo fator de salto de análise ou pelo passo de tempo de análise Δt_a .

[000100] Um exemplo do passo de janelas aplicadas ao sinal de banda baixa, por exemplo, ao sinal de saída do decodificador de núcleo, é descrito na figura 8(a). O passo com a qual a janela de análise de comprimento L é movida para cada transformada de análise é denotada Δt_a . Cada transformada de análise como essa e a porção em janela do sinal de entrada também são referidas como um quadro. A transformada de análise converte / transforma o quadro de amostras de entrada em um conjunto de coeficientes complexos de FFT. Após a transformada de análise, os coeficientes complexos de FFT podem ser transformados de coordenadas cartesianas para polares. A suíte de coeficientes de FFT para quadros subsequentes constitui os sinais de sub-banda de análise. Para cada um dos fatores de transposição $T = 2, 3, \dots, T_{\max}$ usados, os ângulos de fase dos coeficientes de FFT são multiplicados pelo respectivo fator de transposição T e transformados de volta para coordenadas cartesianas. Daí, haverá um conjunto diferente de

coeficientes complexos de FFT representando um quadro em particular para todo fator de transposição T . Em outras palavras, para cada um dos fatores de transposição $T = 2, 3, \dots, T_{\max}$ e para cada quadro, um conjunto em separado de coeficientes de FFT é determinado. Como uma consequência, para toda ordem de transposição T , um conjunto diferente de sinais de sub-banda de síntese $Y(t_s^k, \Omega_m)$ é gerado.

[000101] Nos estágios de síntese, os passos de síntese Δt_s das janelas de síntese são determinadas como uma função da ordem de transposição T usada no respectivo transpositor. Conforme destacado acima, a operação de distensão no tempo também envolve a distensão no tempo dos sinais de sub-banda, isto é, a distensão no tempo da suíte de quadros. Esta operação pode ser realizada pela escolha de um fator de salto de síntese ou passo de síntese Δt_s , o que é aumentado em relação ao passo de análise Δt_a por um fator T . Consequentemente, o passo de síntese Δt_{sT} para o transpositor de ordem T é dada por $\Delta t_{sT} = T\Delta t_a$. As figura 8(b) e 8(c) mostram o passo de síntese Δt_{sT} de janelas de síntese para os fatores de transposição $T = 2$ e $T = 3$, respectivamente, onde $\Delta t_{s2} = 2\Delta t_a$ e $\Delta t_{s3} = 3\Delta t_a$.

[000102] A figura 8 também indica o tempo de referência t_T , o qual foi "distendido" por um fator $T = 2$ e $T = 3$ nas figuras 8(b) e 8(c), se comparadas com a figura 8(a), respectivamente. Contudo, nas saídas este tempo de referência t_T precisa ser alinhado para os dois fatores de transposição. Para alinhamento da saída, o sinal transposto de terceira ordem, isto é, a figura 8(c) precisa ter a amostragem reduzida ou ter a taxa convertida com o fator $3/2$. Esta redução de amostragem leva a uma transposição harmônica com respeito ao sinal transposto de segunda ordem. A figura 9 ilustra o efeito da reamostragem sobre o passo de síntese de janelas para $T = 3$. Se for assumido que o sinal analisado é o sinal de saída de um decodificador de núcleo o qual não teve a amostragem aumentada, então, o sinal da figura 8(b) terá tido

efetivamente a frequência transposta por um fator de 2, e o sinal da figura 8(c) terá tido efetivamente a frequência transposta por um fator de 3.

[000103] A seguir, o aspecto de alinhamento de tempo de sequências transpostas de diferentes fatores de transposição quando se usam janelas de análise comuns é considerado. Em outras palavras, o aspecto de alinhamento de sinais de saída de transpositores de frequência empregando uma ordem de transposição diferente é considerado. Quando se usam os métodos destacados acima, funções de Dirac $\delta(t - t_0)$ são distendidas no tempo, isto é, movidas ao longo do eixo de tempo, pela quantidade de tempo dada pelo fator de transposição T aplicado. De modo a se converter a operação de distensão no tempo em uma operação de deslocamento de frequência, uma decimação ou redução de amostragem usando-se o mesmo fator de transposição T é realizada. Se essa decimação pelo fator de transposição ou pela ordem de transposição T for realizada na função de Dirac distendida no tempo $\delta(t - Tt_0)$, o pulso de Dirac de amostragem reduzida será alinhado no tempo com respeito ao tempo de referência zero 710 na metade da primeira janela de análise 701. Isto é ilustrado na figura 7.

[000104] Contudo, quando se usam diferentes ordens de transposição T , as decimações resultarão em diferentes desvios para a referência zero, a menos que a referência zero esteja alinhada com o tempo "zero" do sinal de entrada. Como uma consequência, um ajuste de desvio de tempo dos sinais transpostos decimados precisa ser realizado, antes de eles podem ser somados na unidade de soma 502. Como um exemplo, um primeiro transpositor de ordem $T = 3$ e um segundo transpositor de ordem $T = 4$ são assumidos. Mais ainda, é assumido que o sinal de saída do decodificador de núcleo não tenha a amostragem aumentada. Então, o transpositor decima o sinal distendido no tempo de terceira

ordem por um fator de $3/2$, e o sinal distendido no tempo de quarta ordem por um fator 2. O sinal distendido no tempo de segunda ordem, isto é, $T = 2$, será exatamente interpretado como tendo uma frequência de amostragem mais alta, se comparada com o sinal de entrada, isto é, um fator de 2 mais alto de frequência de amostragem, efetivamente tornando o sinal de saída deslocado no passo por um fator de 2.

[000105] Pode ser mostrado que, de modo a alinhar os sinais transpostos e de amostragem reduzida, desvios de tempo de $(T-2)L/4$ precisam ser aplicados aos sinais transpostos, antes de uma decimação, isto é, para as transposições de terceira e quarta ordens, desvios de $L/4$ e $L/2$ têm que ser aplicados, respectivamente. Para se verificar isto em um exemplo concreto, será assumido que a referência zero para o sinal distendido no tempo de segunda ordem corresponda ao instante de tempo ou à amostra $L/2$, isto é, à referência zero 710 na figura 7. Isto é assim porque nenhuma decimação é usada. Para um sinal distendido no tempo de terceira ordem, a referência transladará para $L/2 (2/3) = L/3$, devido à redução de amostragem por um fator de $3/2$. Se o desvio de tempo de acordo com a regra mencionada acima for adicionado antes de uma decimação, a referência transladará para $(L/2 + L/4) (2/3) = L/2$. Isto significa que a referência do sinal transposto de amostragem reduzida está alinhada com a referência zero 710. De uma maneira similar, para a transposição de quarta ordem sem desvio, a referência zero corresponde a $L/2 (1/2) = L/4$, mas, quando se usa o desvio proposto, a referência translada para $(L/2 + L/2) (1/2) = L/2$, o que de novo está alinhado com a referência zero de segunda ordem 710, isto é, a referência zero para o sinal transposto usando-se $T = 2$.

[000106] Um outro aspecto a ser considerado quando se usam simultaneamente múltiplas ordens de transposição refere-se aos ganhos aplicados às sequências transpostas de diferentes fatores de transposição. Em outras palavras, o aspecto de combinação dos sinais

de saída de transpositores de ordem de transposição diferente pode ser considerado. Há dois princípios quando na seleção do ganho dos sinais transpostos, os quais podem ser considerados sob diferentes abordagens teóricas. Em qualquer um, os sinais transpostos são supostos como sendo de conservação de energia, significando que a energia total no sinal de banda baixa, o qual é subsequentemente transposto para constituir um sinal de banda alta transposto por um fator T , é preservada. Neste caso, a energia por largura de banda deve ser reduzida pelo fator de transposição T , uma vez que o sinal é distendido pela mesma quantidade T na frequência. Contudo, senoides, as quais têm sua energia em uma largura de banda infinitesimalmente pequena, reterão sua energia após uma transposição. Isto é devido ao fato de que, da mesma forma como um pulso de Dirac é movido no tempo pelo transpositor quando de uma distensão no tempo, isto é, da mesma forma que a duração no tempo do pulso não é mudada pela operação de distensão no tempo, uma senoide é movida na frequência, quando na transposição, isto é, a duração na frequência (em outras palavras, a largura de banda) não é mudada pela operação de transposição de frequência. Isto é, embora a energia por largura de banda seja reduzida em T , a senoide tem toda sua energia em um ponto na frequência, de modo que a energia no sentido de ponto seja preservada.

[000107] A outra opção quando na seleção do ganho dos sinais transpostos é manter a energia por largura de banda após uma transposição. Neste caso, ruído branco de banda larga e transientes exibirão uma resposta de frequência plana após uma transposição, enquanto a energia das senoides aumentará por um fator T .

[000108] Um aspecto adicional da invenção é a escolha de janelas de codificador de voz de fase de análise e de síntese quando se usam janelas de análise comuns. É benéfico escolher cuidadosamente as janelas de codificador de voz de fase de análise e de síntese, isto é,

$v_a(n)$ e $v_s(n)$. Não apenas a janela de síntese $v_s(n)$ deve aderir à fórmula 2 acima, de modo a se permitir uma reconstrução perfeita. Mais ainda, a janela de análise $v_a(n)$ também deve ter uma rejeição adequada dos níveis de lobo lateral. Caso contrário, termos de "descontinuidade" indesejados tipicamente serão audíveis como uma interferência com os termos principais para senoides variando na frequência. Esses termos indesejados de "descontinuidade" também podem aparecer para senoides estacionárias, no caso de fatores de transposição pares, conforme mencionado acima. A presente invenção propõe o uso de janelas de seno por causa de sua boa relação de rejeição de lobo lateral. Daí, a janela de análise é proposta como sendo:

$$v_a(n) = \sin\left(\frac{\pi}{L}(n+0.5)\right), 0 \leq n < L \quad (4)$$

[000109] A janela de síntese $v_s(n)$ pode ser idêntica à janela de análise $v_a(n)$ ou dada pela fórmula (2) acima, se o tamanho de salto de síntese Δt_s não for um fator do comprimento de janela de análise L , isto é, se o comprimento de janela de análise L não for um inteiro divisível pelo tamanho de salto de síntese. A título de exemplo se $L = 1024$ e $\Delta t_s = 384$, então, $1024 / 384 = 2,667$ e não é um número inteiro. Deve ser notado que também é possível selecionar um par de janelas de análise e de síntese biortogonais, conforme destacado acima. Isto pode ser benéfico para a redução de descontinuidade no sinal de saída, notadamente quando se usam ordens de transposição T pares.

[000110] A seguir, uma referência é feita à figura 10 e à figura 11, as quais ilustram um codificador de exemplo 1000 e um decodificador de exemplo 1100, respectivamente, para uma codificação unificada de fala e áudio (USAC). A estrutura geral do codificador 1000 e do decodificador 1100 de USAC é descrita conforme se segue: em primeiro lugar, pode haver um pré-/pós-processamento comum consistindo em uma unidade funcional de MPEG Surround (MPEGS) para lidar com processamento em estéreo ou de canal múltiplo e uma unidade de

Replicação de Banda Espectral melhorada (eSBR) 1001 e 1101, respectivamente, a qual lida com a representação paramétrica das frequências de áudio mais altas no sinal de entrada, e a qual pode fazer uso dos métodos de transposição harmônica destacados no presente documento. Então, há duas ramificações, uma consistindo em um percurso de ferramenta de codificação de áudio avançada (AAC) modificada e a outra consistindo em percurso baseado em codificação de predição linear (domínio de LP ou LPC), o que por sua vez caracteriza uma representação de domínio de frequência ou uma representação de domínio de tempo da LPC residual. Todos os espectros transmitidos para ambos, AAC e LPC, podem ser representados no domínio de MDCT seguidos por uma quantificação e uma codificação aritmética. A representação de domínio de tempo pode usar um esquema de codificação de excitação ACELP.

[000111] A unidade de Replicação de Banda Espectral melhorada (eSBR) 1001 do codificador 1000 pode compreender componentes de reconstrução de frequência alta destacados no presente documento. Em algumas modalidades, a unidade de sSBR 1001 pode compreender uma unidade de transposição destacada no contexto das figuras 4, 5 e 6. Os dados codificados relacionados a uma transposição harmônica, por exemplo, a ordem de transposição usada, a quantidade de superamostragem de domínio de frequência necessária, ou os ganhos empregados, podem ser derivados no codificador 1000 e fundidos com a outra informação codificada em um multiplexador de fluxo de bit e encaminhados como um fluxo de áudio codificado para um decodificador correspondente 1100.

[000112] O decodificador 1100 mostrado na figura 11 também compreende uma unidade de Replicação de Banda Espectral melhorada (eSBR) 1101. Esta unidade de eSBR 1101 recebe o fluxo de bit de áudio codificado ou o sinal codificado a partir do codificador 1000

e usa os métodos destacados no presente documento para a geração de um componente de frequência alta ou banda alta do sinal, o qual é fundido com o componente de frequência baixa ou banda baixa decodificada para a produção de um sinal decodificado. Em particular, ele pode compreender a unidade de transposição destacada no contexto das figuras 4, 5 e 6. A unidade de eSBR 1101 pode usar uma informação sobre o componente de frequência alta provido pelo codificador 1000 através do fluxo de bit, de modo a se realizar a reconstrução de frequência alta. Essa informação pode ser a envoltória espectral do componente de frequência alta original para a geração dos sinais de sub-banda de síntese e, finalmente, do componente de frequência alta do sinal decodificado, bem como a ordem de transposição usada, a quantidade de superamostragem de domínio de frequência necessária ou os ganhos empregados.

[000113] Mais ainda, as figuras 10 e 11 ilustram possíveis componentes adicionais de um codificador / decodificador de USAC, tais como:

- uma ferramenta de demultiplexador de carga útil de fluxo de bit, a qual separa a carga útil de fluxo de bit nas partes para cada ferramenta, e provê a cada uma das ferramentas a informação de carga útil de fluxo de bit relacionada àquela ferramenta;
- uma ferramenta de decodificação sem ruído de fator de escala, a qual toma uma informação a partir do demultiplexador de carga útil de fluxo de bit, analisa gramaticalmente aquela informação, e decodifica os fatores de escala de Huffman e de DPCM;
- uma ferramenta de decodificação sem ruído espectral, a qual toma uma informação a partir do demultiplexador de carga útil de fluxo de bit, analisa gramaticalmente aquela informação, decodifica os dados codificados de forma aritmética e reconstrói os espectros quantificados;

- uma ferramenta de quantificador inverso, a qual toma os valores quantificados para os espectros, e converte os valores inteiros nos espectros reconstruídos não escalonados; este quantificador é preferencialmente um quantificador de compressão e expansão, cujo fator de compressão e expansão depende do modo de codificação de núcleo escolhido;

- uma ferramenta de preenchimento de ruído, a qual é usada para o preenchimento de espaços espectrais nos espectros decodificados, os quais ocorrem quando os valores espectrais são quantificados para zero, por exemplo, devido a uma forte restrição na demanda de bit no codificador;

- uma ferramenta de reescalonamento, a qual converte a representação de número inteiro dos fatores de escala nos valores reais, e multiplica os espectros quantificados inversamente não escalonados pelos fatores de escala relevantes;

- uma ferramenta de M/S, conforme descrito na ISO / IEC 14496-3;

- uma ferramenta de conformação de ruído temporal (TNS), conforme descrito na ISO / IEC 14496-3;

- uma ferramenta de comutação de banco / bloco de filtro, a qual aplica o inverso do mapeamento de frequência que foi realizado no codificador; uma transformada de cosseno discreta modificada inversa (IMDCT) preferencialmente é usada para a ferramenta de banco de filtro;

- uma ferramenta de comutação de banco / bloco de filtro de distorção no tempo, a qual substitui a ferramenta de comutação de banco / bloco de filtro normal quando o modo de distorção de tempo for habilitado; o banco de filtro preferencialmente é o mesmo (IMDCT) que para o banco de filtro normal, adicionalmente as amostras em janelas de domínio de tempo são mapeadas a partir do domínio de tempo

distorcido para o domínio de tempo linear por uma reamostragem variando no tempo;

- uma ferramenta de MPEG Surround (MPEGS), a qual produz múltiplos sinais a partir de um ou mais sinais de entrada pela aplicação de um procedimento de upmix sofisticado para o(s) sinal(is) de entrada controlado(s) por parâmetros espaciais apropriados; no contexto de USAC, o MPEGS preferencialmente é usado para a codificação de um sinal de canal múltiplo, pela transmissão de uma informação de lado paramétrico ao longo de um sinal de downmix transmitido;

- uma ferramenta classificadora de sinal, a qual analisa o sinal de entrada original e gera a partir dele uma informação de controle a qual dispara a seleção dos diferentes modos de codificação; a análise do sinal de entrada tipicamente é dependente de implementação, e tentará escolher o modo de codificação de núcleo ótimo para um dado quadro de sinal de entrada; a saída do classificador de sinal opcionalmente também pode ser usada para influenciar o comportamento de outras ferramentas, por exemplo, MPEG Surround, SBR melhorada, banco de filtro distorcido no tempo e outras;

- uma ferramenta de filtro de LPC, a qual produz um sinal de domínio de tempo a partir de um sinal de domínio de excitação pela filtração do sinal de excitação reconstruído através de um filtro de síntese de predição linear; e

- uma ferramenta de ACELP, a qual provê uma forma de se representar eficientemente um sinal de excitação de domínio de tempo pela combinação de um preditor de longo prazo (palavra de código adaptativa) com uma sequência tipo de pulso (palavra de código de inovação).

[000114] A figura 12 ilustra uma modalidade das unidades de eSBR mostradas nas figuras 10 e 11. A unidade de eSBR 1200 será descrita

a seguir no contexto de um decodificador, onde a entrada para a unidade de eSBR 1200 é o componente de frequência baixa, também conhecida como banda baixa, de um sinal.

[000115] Na figura 12, o componente de frequência baixa 1213 é alimentado em um banco de filtro de QMF, de modo a se gerarem bandas de frequência de QMF. Estas bandas de frequência de QMF não são para serem confundidas com as sub-bandas de análise destacadas neste documento. As bandas de frequência de QMF são usadas para fins de manipulação e fusão das componentes de frequência baixa e alta do sinal no domínio de frequência, em vez de no domínio de tempo. O componente de frequência baixa 1214 é alimentado para a unidade de transposição 1204, a qual corresponde aos sistemas para reconstrução de frequência alta destacados no presente documento. A unidade de transposição 1204 gera um componente de frequência alta 1212, também conhecido como banda alta, do sinal, o qual é transformado no domínio de frequência por um banco de filtro de QMF 1203. Ambas o componente de frequência baixa transformada por QMF e o componente de frequência alta transformada por QMF são alimentados em uma unidade de manipulação e de fusão 1205. Esta unidade 1205 pode realizar um ajuste de envoltória do componente de frequência alta e combina o componente de frequência alta e o componente de frequência baixa ajustados. O sinal de saída combinado é retransformado no domínio de tempo por um banco de filtro de QMF inversa 1201.

[000116] Tipicamente, o banco de filtro de QMF 1202 compreende 32 bandas de frequência de QMF. Nesses casos, o componente de frequência baixa 1213 tem uma largura de banda de $f_s/4$, onde $f_s/2$ é a frequência de amostragem do sinal 1213. O componente de frequência alta 1212 tipicamente tem uma largura de banda de $f_s/2$ e é filtrado através do banco de filtro 1203 compreendendo 64 bandas de

frequência de QMF.

[000117] No presente documento, um método para transposição harmônica foi destacado. Este método de transposição harmônica é particularmente bem adequado para a transposição de sinais transientes. Ele compreende a combinação de superamostragem de domínio de frequência com transposição harmônica usando codificadores de voz. A operação de transposição depende da combinação de janela de análise, passo de janela de análise, tamanho de transformada, janela de síntese, passo de janela de síntese, bem como dos ajustes de fase do sinal analisado. Através do uso deste método, efeitos indesejados, tais como pré- e pós-ecos, podem ser evitados. Mais ainda, o método não faz uso de medidas de análise de sinal, tal como detecção de transiente, as quais tipicamente introduzem distorções de sinal devido a descontinuidades no processamento de sinal. Além disso, o método proposto apenas tem complexidade computacional reduzida. O método de transposição harmônica de acordo com a invenção pode ser adicionalmente melhorado por uma seleção apropriada de janelas de análise / síntese, valores de ganho e/ou alinhamento de tempo.

REIVINDICAÇÕES

1. Sistema para gerar um sinal de áudio de saída a partir de um sinal de áudio de entrada (312) usando um fator de transposição T , compreendendo:

uma unidade de janela de análise (602) para aplicar uma janela de análise (311) de comprimento L_a , desse modo se extraíndo um quadro do sinal de áudio de entrada (312);

uma unidade de transformação de análise (603) de ordem M (301), para transformar as amostras em M coeficientes complexos;

uma unidade de processamento não linear (604) para alterar a fase dos coeficientes complexos pelo uso do fator de transposição T ;

uma unidade de transformação de síntese (605) de ordem M , para transformar os coeficientes alterados em M amostras alteradas;
e

uma unidade de janela de síntese (606) para aplicar uma janela de síntese (321) de comprimento L_s às M amostras alteradas, desse modo gerando um quadro do sinal de áudio de saída;

caracterizado pelo fato de que M é baseado no fator de transposição T , e compreendendo ainda

uma unidade de passo de análise (601), para deslocar a janela de análise por um passo de análise de S_a amostras ao longo do sinal de áudio de entrada, desse modo se gerando uma sucessão de quadros do sinal de áudio de entrada;

uma unidade de passo de síntese (607), para deslocar quadros sucessivos do sinal de áudio de saída por um passo de síntese de S_s amostras; e

uma unidade superposição – adição (608), para sobrepor e adicionar quadros deslocados sucessivos dos sinais de saída, desse modo gerando o sinal de áudio de saída.

2. Sistema, de acordo com a reivindicação 1, **caracterizado**

pelo fato de que ainda compreende uma unidade de contração (609), para aumentar a taxa de amostragem do sinal de áudio de saída pela ordem de transposição T; e/ou

para reduzir a amostragem do sinal de áudio de saída pela ordem de transposição T, enquanto mantém a taxa de amostragem não modificada;

desse modo produzindo um sinal de áudio de saída transposto.

3. Sistema, de acordo com a reivindicação 2, **caracterizado pelo fato de que** ainda compreende:

uma segunda unidade de processamento não linear (604), para alterar a fase dos coeficientes complexos pelo uso de um segundo fator de transposição T_2 , desse modo se produzindo um quadro de um segundo sinal de áudio de saída;

uma segunda unidade de passo de síntese (607), para deslocar quadros sucessivos do segundo sinal de áudio de saída por um segundo passo de síntese, desse modo se gerando o segundo sinal de áudio de saída na unidade de superposição – adição (608);

uma segunda unidade de contração (609), para usar a segunda ordem de transposição T_2 , desse modo produzindo um segundo sinal de áudio de saída transposto; e

uma unidade de combinação (502), para fundir os primeiro e segundo sinais de áudio de saída transpostos.

4. Sistema, de acordo com a reivindicação 3, **caracterizado pelo fato de que**:

a unidade de combinação (502) atribui pesos aos primeiro e segundo sinais de áudio de saída transpostos, antes da fusão; e

a atribuição de peso é realizada de modo que a energia ou a energia por largura de banda dos primeiro e segundo sinais de áudio de saída transpostos corresponde à energia ou energia por largura de

banda do sinal de áudio de entrada, respectivamente.

5. Sistema para gerar um sinal de áudio de saída a partir de um sinal de áudio de entrada (312) usando um fator de transposição T , **caracterizado pelo fato de que** compreende:

uma unidade de janela de análise (602) que aplica uma janela de análise (311) de comprimento L , desse modo se extraíndo um quadro do sinal de áudio de entrada (312);

uma unidade de transformação de análise (603) de ordem M (301), que transforma as amostras em M coeficientes complexos;

uma unidade de processamento não linear (604) que altera a fase dos coeficientes complexos pelo uso do fator de transposição T ;

uma unidade de transformação de síntese (605) de ordem M , que transforma os coeficientes alterados em M amostras alteradas;

e

uma unidade de janela de síntese (606) que aplica uma janela de síntese (321) de comprimento L às M amostras alteradas, desse modo gerando um quadro do sinal de áudio de saída;

em que a janela de análise (311) e a janela de síntese (321) são diferentes de cada outra e biortogonais com respeito a cada outra;

e

em que a transformada z da janela de análise (311) tem zeros duplos no círculo unitário.

6. Método para transpor um sinal de áudio de entrada (312) por um fator de transposição T , compreendendo as etapas de:

extrair um quadro de amostras do sinal de áudio de entrada (312) usando uma janela de análise (311) de comprimento L_a ;

transformar o quadro do sinal de áudio de entrada do domínio de tempo para o domínio de frequência produzindo M coeficientes complexos;

alterar a fase dos coeficientes complexos com o fator de

transposição T;

transformar os M coeficientes complexos alterados no domínio de frequência produzindo M amostras alteradas; e

gerar um quadro de um sinal de áudio de saída usando uma janela de síntese (321) de comprimento L_s ;

caracterizado pelo fato de que M é baseado no fator de transposição T, e compreendendo ainda

deslocar a janela de análise por um passo de análise de S_a amostras ao longo do sinal de áudio de entrada, deste modo produzindo uma sucessão de quadro do sinal de áudio de entrada;

deslocar quadros sucessivos do sinal de áudio de saída por um passo de síntese de S_s amostras; e

superpor e adicionar quadros deslocados sucessivos dos sinais de áudio de saída, gerando deste modo o sinal de áudio de saída.

7. Método, de acordo com a reivindicação 6, **caracterizado pelo fato de que** ainda compreende as etapas de:

alterar a fase dos coeficientes complexos pelo uso de um segundo fator de transposição T_2 , desse modo gerando um quadro de um segundo sinal de áudio de saída; e

deslocar quadros sucessivos do segundo sinal de áudio de saída por um segundo passo de síntese, desse modo gerando um segundo sinal de áudio de saída pela superposição e adição dos quadros deslocados do segundo sinal de áudio de saída.

8. Método para transpor um sinal de áudio de entrada (312) por um fator de transposição T, **caracterizado pelo fato de que** compreende as etapas de:

extrair um quadro de amostras do sinal de áudio de entrada (312) usando-se uma janela de análise (311) de comprimento L;

transformar o quadro do sinal de áudio de entrada do domínio de tempo para o domínio de frequência produzindo M

coeficientes complexos;

alterar a fase dos coeficientes complexos com o fator de transposição T;

transformar os M coeficientes complexos alterados no domínio de frequência produzindo M amostras alteradas; e

gerar um quadro de um sinal de áudio de saída usando uma janela de síntese (321) de comprimento L;

em que a janela de análise (311) e a janela de síntese (321) são diferentes de cada outra e são biortogonais com respeito a cada outra; e em que a transformada z da janela de análise (311) tem zeros duplos no círculo de unidades.

9. Meio de armazenamento, **caracterizado pelo fato de que** compreende um processador para realizar das etapas do método conforme definido em qualquer uma das reivindicações 6 a 8, quando realizadas em um dispositivo de computação.

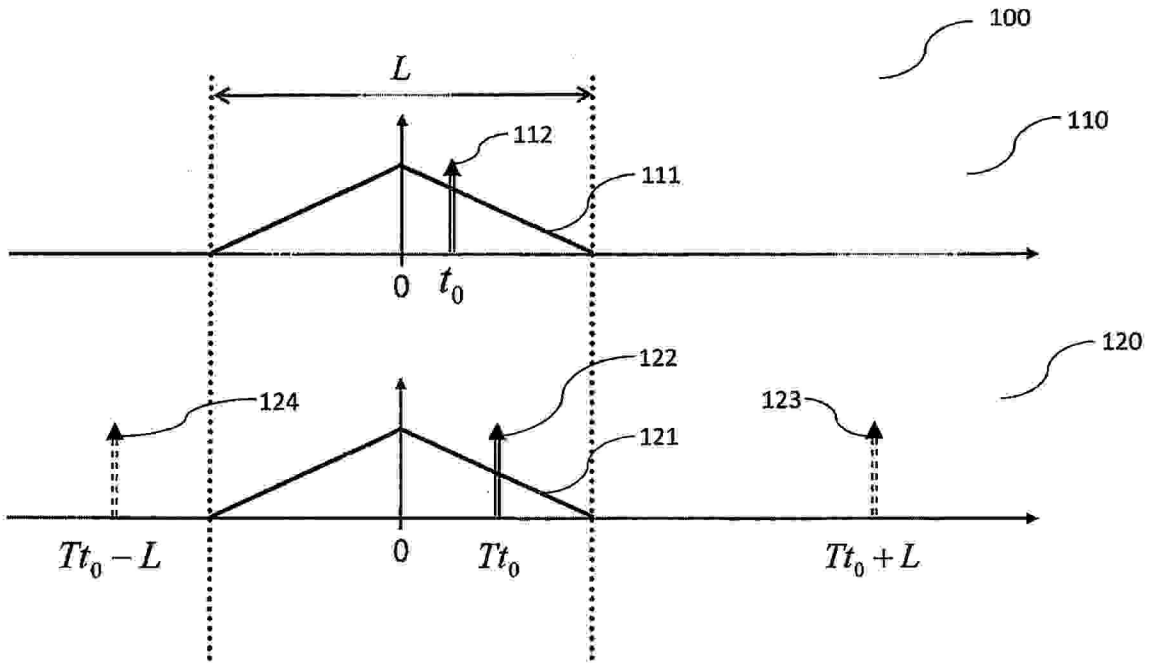


Fig. 1

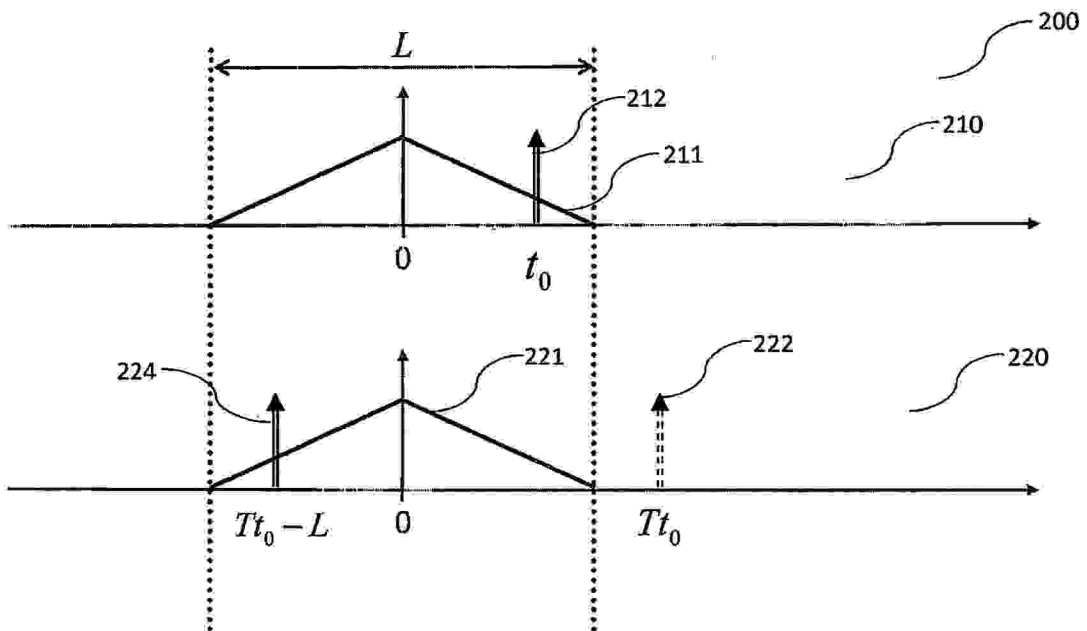


Fig. 2

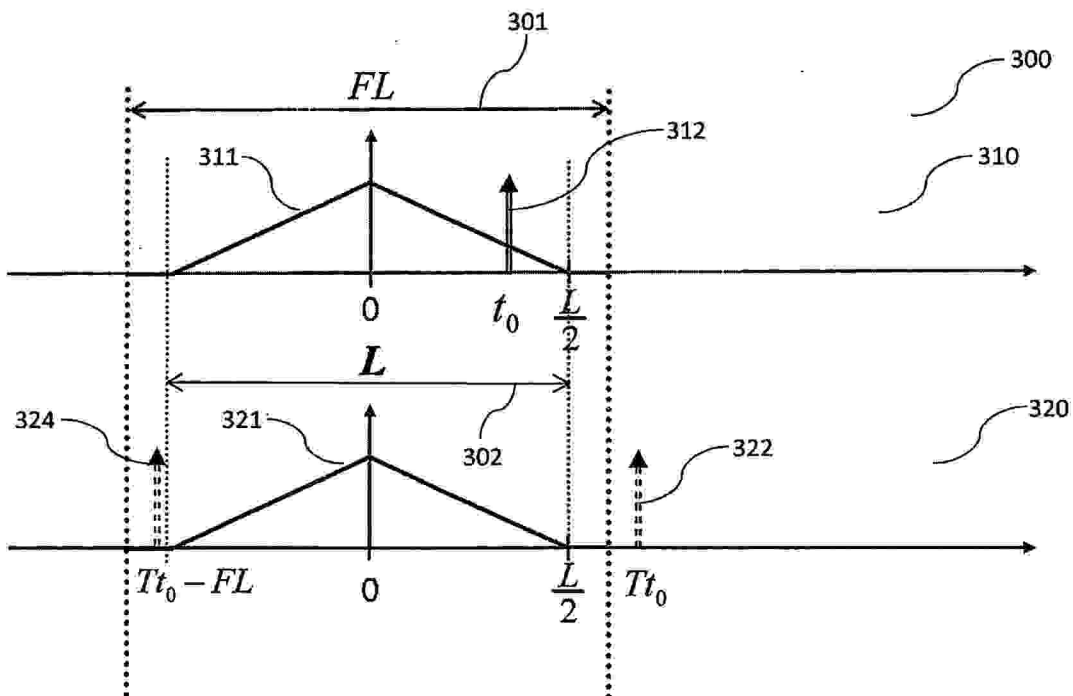


Fig. 3

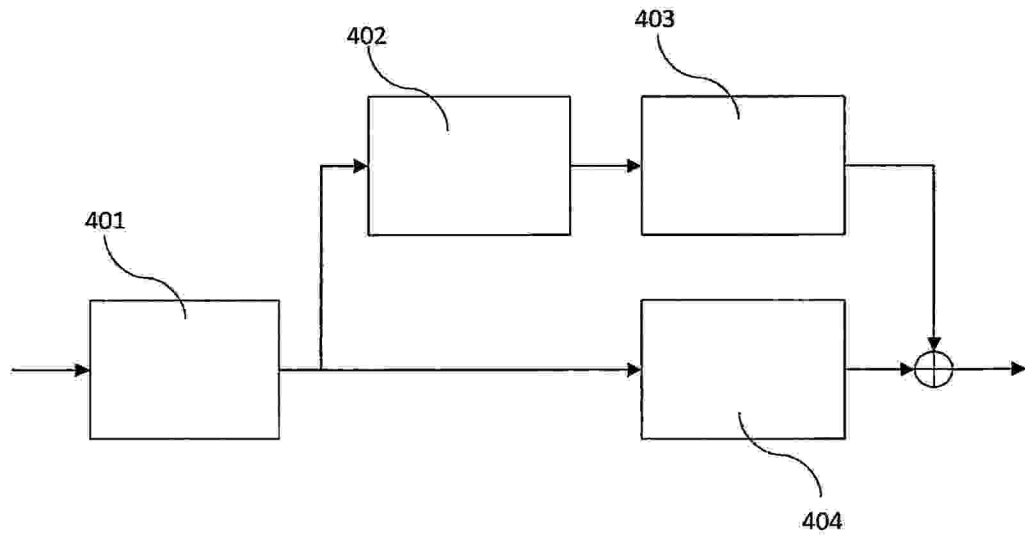


Fig. 4

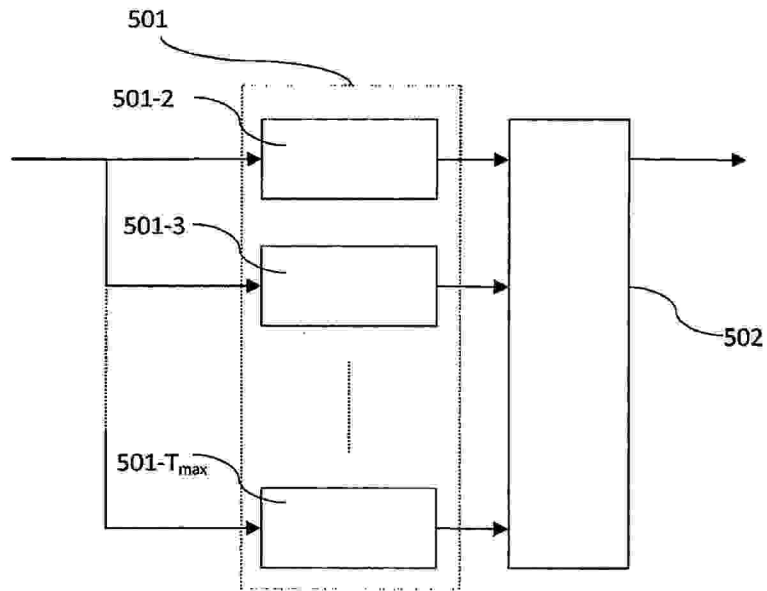


Fig. 5

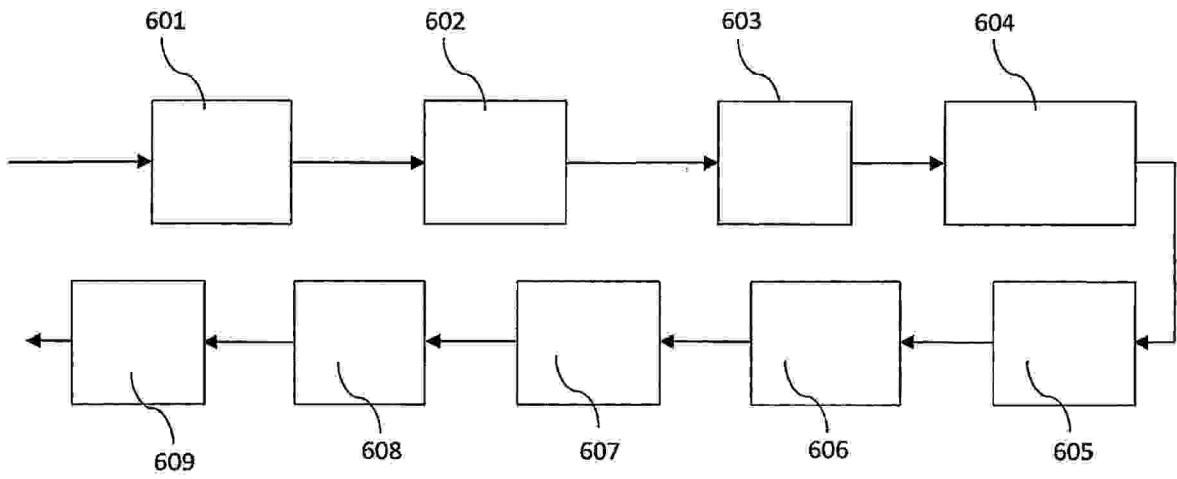


Fig. 6

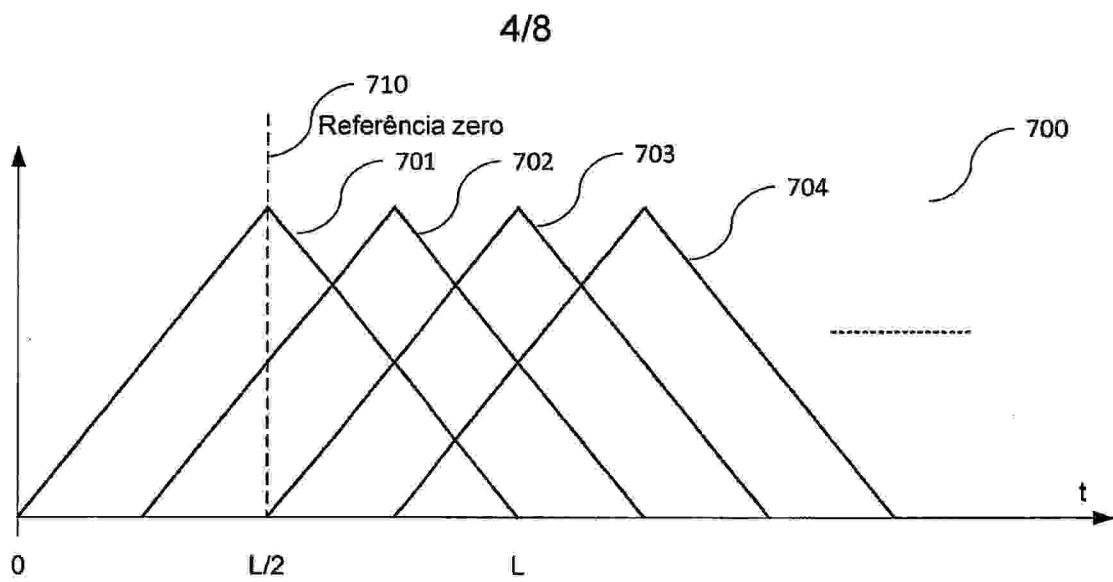
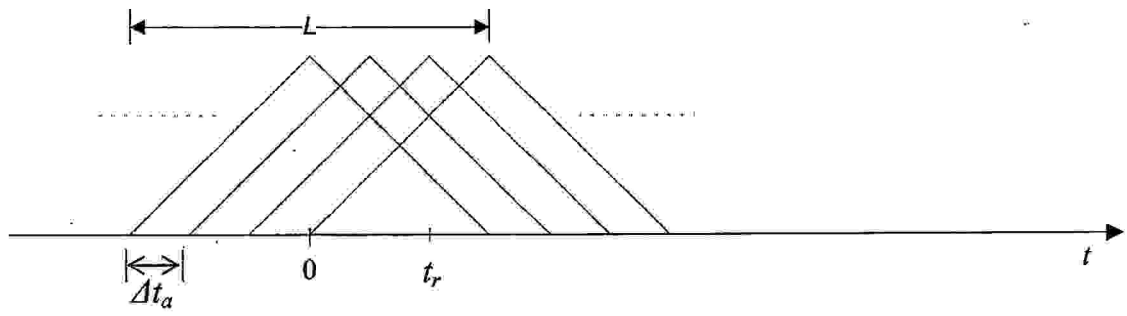
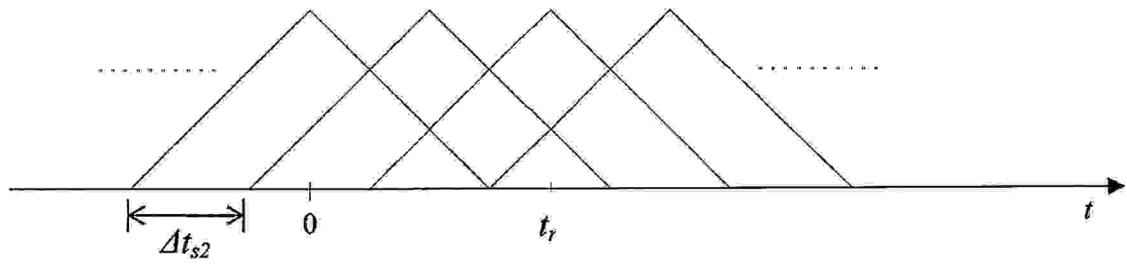


Fig. 7

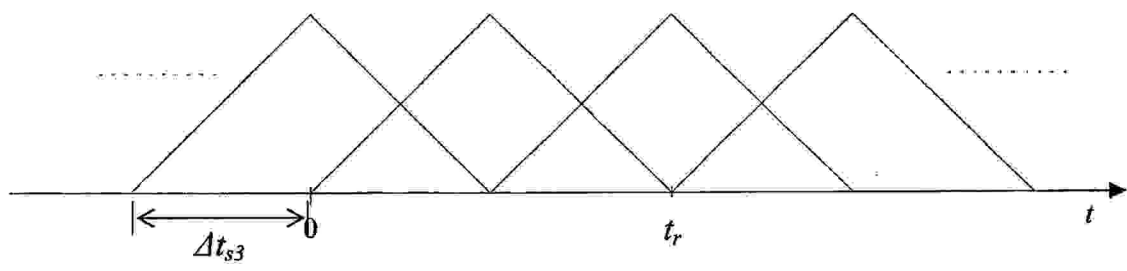
5/8



(a)



(b)



(c)

Fig. 8

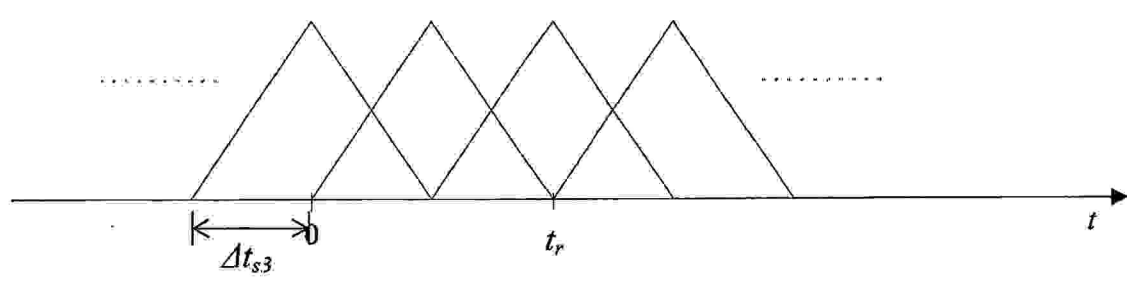


Fig. 9

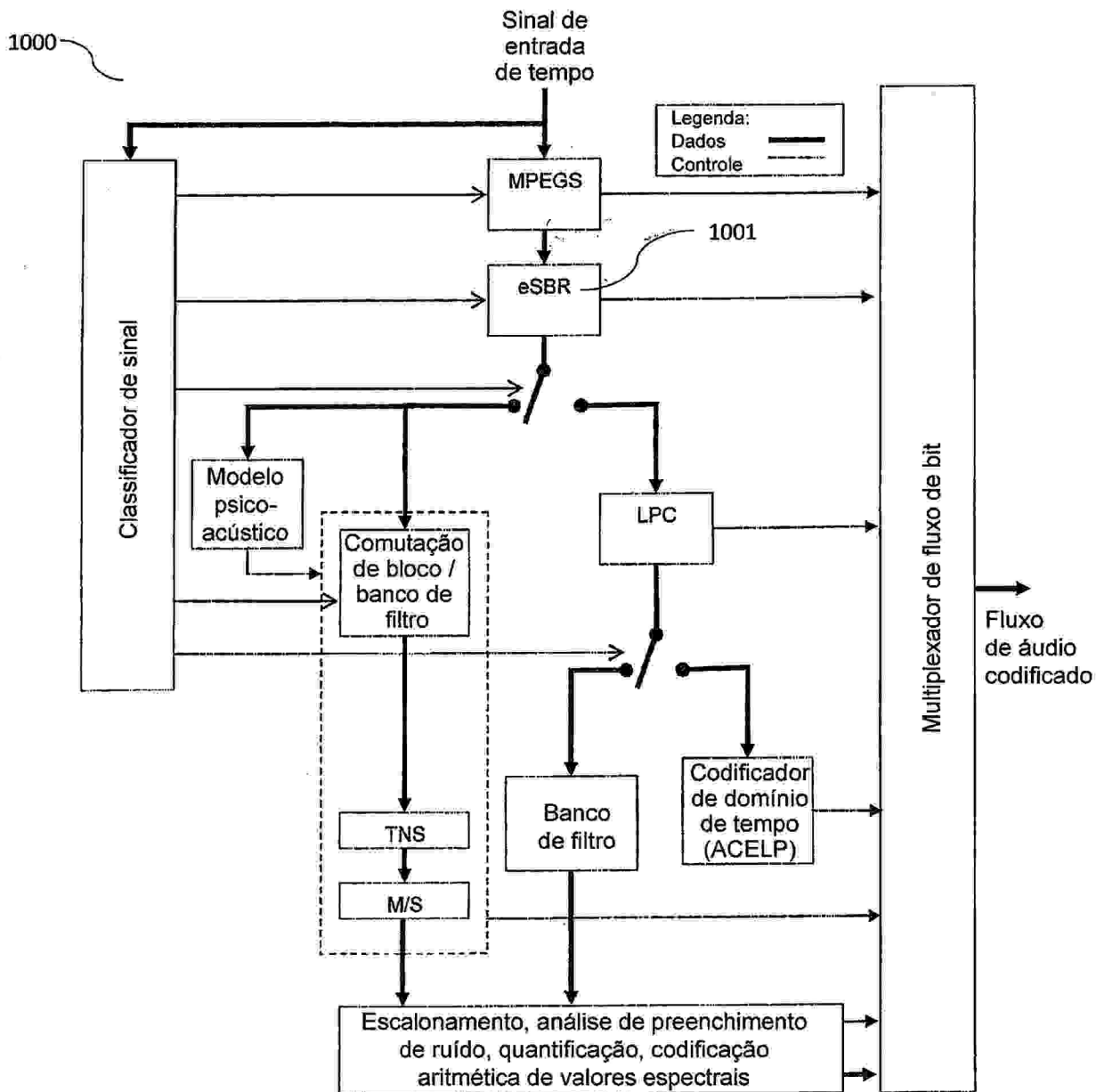


Fig. 10

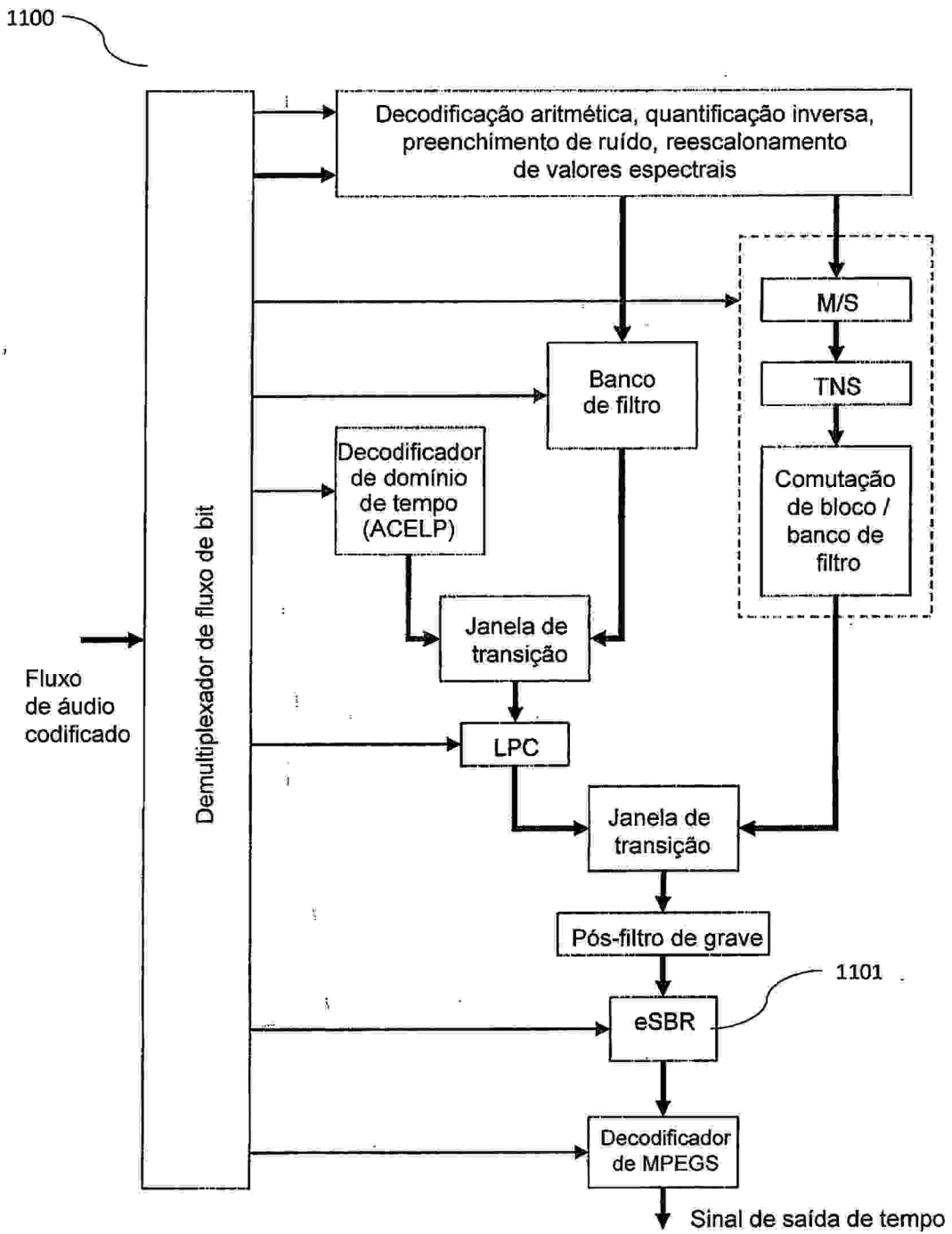


Fig. 11

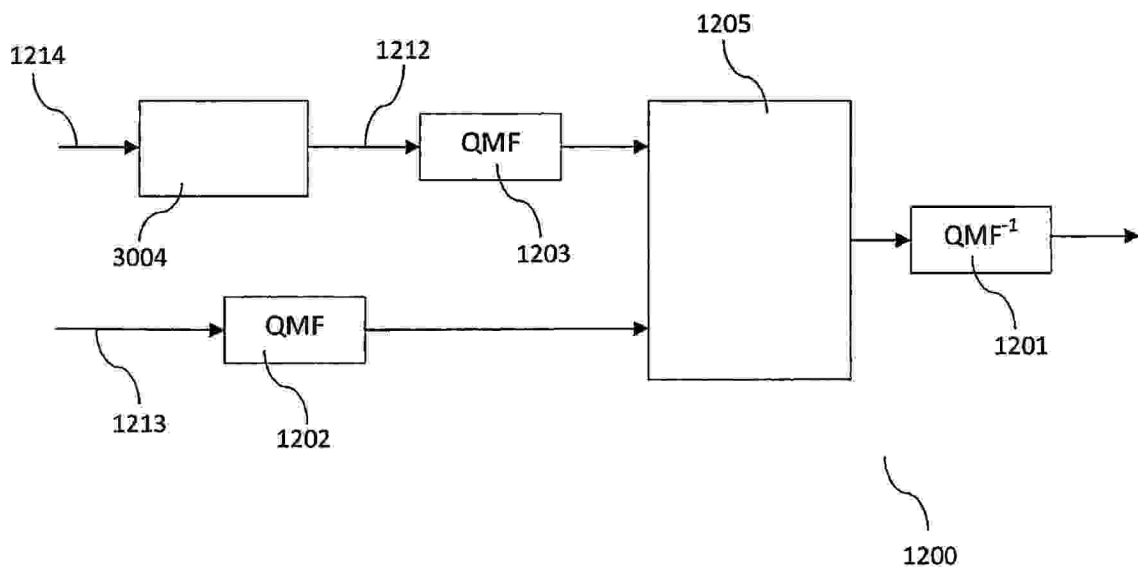


Fig. 12