



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

(11) BR 112015017748-4 B1



(22) Data do Depósito: 28/01/2014

(45) Data de Concessão: 15/03/2022

(54) Título: PREENCHIMENTO DE RUÍDO NA CODIFICAÇÃO DE ÁUDIO DE TRANSFORMADA PERCEPTUAL

(51) Int.Cl.: G10L 19/028.

(30) Prioridade Unionista: 29/01/2013 US 61/758,209.

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

(72) Inventor(es): DISCH, SASCHA; GAYER, MARC; HELMRICH, CHRISTIAN; MARKOVIC, GORAN; LUIS VALERO, MARIA.

(86) Pedido PCT: PCT EP2014051631 de 28/01/2014

(87) Publicação PCT: WO 2014/118176 de 07/08/2014

(85) Data do Início da Fase Nacional: 24/07/2015

(57) Resumo: PREENCHIMENTO DE RUÍDO NA CODIFICAÇÃO DE ÁUDIO DE TRANSFORMADA PERCEPTUAL O preenchimento de ruído em codecs de áudio de transformada perceptual é melhorado realizando o preenchimento de ruído com uma inclinação espectralmente global ao invés de em forma espectralmente plana.

"PREENCHIMENTO DE RUÍDO NA CODIFICAÇÃO DE ÁUDIO DE TRANSFORMADA PERCEPTUAL".

DESCRIÇÃO

[0001] O presente pedido refere-se ao preenchimento de ruído na codificação de áudio de transformada perceptual.

[0002] Na codificação de transformada é geralmente reconhecido (comparar [1], [2], [3]) que as partes de quantização de um espectro para zero leva a uma degradação perceptual. Estas partes quantizadas a zero são chamadas furos de espectro. Uma solução para este problema apresentado em [1], [2], [3] e [4] é substituir as linhas espectrais quantizadas em zero por ruído. Às vezes, a inserção de ruído é evitada abaixo de uma certa frequência. A frequência inicial para preenchimento de ruído é fixada, mas diferente entre a técnica anterior conhecida.

[0003] Às vezes, FDNS (Formação de Ruído de Domínio de Frequência | *Frequency Domain Noise Shaping*) é utilizada para formar o espectro (incluindo o ruído inserido) e para o controle do ruído de quantização, como em USAC (comparar [4]). FDNS é realizada utilizando a resposta de magnitude do filtro LPC. Os coeficientes do filtro LPC são calculados utilizando o sinal de entrada pré-enfatizado.

[0004] Foi observado em [1] que adicionar o ruído na proximidade imediata de um componente tonal leva a uma degradação e certamente, apenas como em [5] apenas testes longos de zeros são preenchidos com ruído para evitar ocultar os valores quantizados não zero pelo ruído envoltório injetado.

[0005] Em [3] observa-se que há um problema de um compromisso entre a granularidade do preenchimento de ruído e o tamanho da informação adicional necessária. Em [1], [2], [3] e [5] um parâmetro do preenchimento de ruído por espectro completo é transmitido. O ruído inserido é espectralmente formado utilizando LPC como em [2] ou utilizando fatores de escala como em [3]. É descrito em [3] como adaptar os fatores de escala em um preenchimento de ruído com um preenchimento de nível de ruído para todo o espectro. Em [3], os fatores de escala para as bandas que são completamente quantizadas a zero são modificados para evitar os furos espectrais e ter um nível de ruído correto.

[0006] Embora as soluções em [1] e [5] evitem uma degradação de componentes tonais pelo fato de que elas sugerem o não preenchimento de pequenos furos de espectro, há ainda uma necessidade de melhorar mais a qualidade de um sinal de áudio codificado utilizando o preenchimento de ruído, especialmente em taxas de bit muito baixas.

[0007] Há outros problemas além dos discutidos, que resultam dos conceitos de preenchimento de ruído conhecidos até o momento, de acordo com que o ruído é preenchido no espectro em uma forma espectralmente plana.

[0008] Seria favorável ter um conceito melhorado de preenchimento de ruído em mãos que aumenta a qualidade do áudio alcançável resultante do espectro preenchido com ruído, pelo menos, em conexão com a codificação de áudio de transformada perceptual.

[0009] Certamente, é um objeto da presente invenção

fornecer um conceito para preenchimento de ruído na codificação de áudio de transformada perceptual com características melhoradas.

[0010] Este objeto é obtido pelo assunto das reivindicações independentes aqui anexas, caracterizado pelos aspectos vantajosos do presente pedido serem submetidos às reivindicações dependentes.

[0011] É uma constatação básica do presente pedido que o preenchimento de ruído nos codecs de áudio de transformada perceptual pode ser melhorado pela realização do preenchimento de ruído com uma inclinação espectralmente global, ao invés de uma forma espectralmente plana. Por exemplo, a inclinação espectralmente global pode ter um declive negativo, ou seja, exibir uma redução das frequências baixas a altas, a fim de, pelo menos, reverter parcialmente a inclinação espectral causada submetendo o espectro preenchido com ruído à função de ponderação perceptual espectral. Um declive positivo pode ser imaginável também, por exemplo, nos casos onde o espectro codificado exibe um caráter do tipo passa-alta. Em particular, as funções de ponderação perceptual espectral tipicamente tendem a exibir um aumento de frequências baixas a altas. Certamente, o ruído preenchido ao espectro de codificadores de áudio de transformada perceptual em uma forma espectralmente plana, terminaria em um patamar de ruído inclinado no espectro finalmente reconstruído. Os inventores do presente pedido, entretanto, perceberam que esta inclinação no espectro finalmente reconstruído afeta negativamente a qualidade de áudio, pois

leva aos furos espectrais remanescentes nas partes preenchidas pelo ruído do espectro. Certamente, inserir o ruído com uma inclinação espectralmente global de modo que o nível de ruído reduz das frequências baixas a altas, pelo menos, parcialmente compensa tal inclinação espectral causada pela formação subsequente do espectro preenchido com ruído utilizando a função de ponderação perceptual espectral, assim melhorando a qualidade de áudio. Dependendo das circunstâncias, um declive positivo pode ser preferido, conforme observado acima.

[0012] De acordo com uma aplicação, o declive da inclinação espectralmente global é variado receptivo a uma sinalização no fluxo de dados no qual o espectro é codificado. A sinalização pode, por exemplo, explicitamente sinalizar o declive e pode ser adaptada, no lado de decodificação, a quantidade de inclinação espectral causada pela função de ponderação perceptual espectral. Por exemplo, a quantidade de inclinação espectral causada pela função de ponderação perceptual espectral pode resultar de uma pré-ênfase cujo o sinal de áudio é submetido à aplicação prévia da análise LPC nele.

[0013] De acordo com uma aplicação, o preenchimento de ruído de um espectro de um sinal de áudio é melhorado na qualidade com relação ao espectro preenchido com ruído ainda assim de modo que a reprodução do ruído preenchido sinal de áudio seja menos irritante, realizando o preenchimento de ruído em uma forma dependente de uma tonalidade do sinal de áudio.

[0014] De acordo com uma aplicação do presente pedido, uma parte zero espectral contínua do espectro do sinal de áudio é preenchida com ruído espectralmente formado utilizando uma função assumindo um máximo em uma parte interna da parte zero espectral contínua, e tendo bordas exteriormente descendentes de um declive absoluto que depende negativamente da tonalidade, ou seja, o declive reduz com o aumento da tonalidade. De forma adicional ou alternativa, a função utilizada para preenchimento assume um máximo em uma parte interna da parte zero espectral contínua e tem bordas exteriormente descendentes, uma largura espectral que depende positivamente da tonalidade, ou seja, a largura espectral aumenta com o aumento da tonalidade. Ainda assim, de forma adicional ou alternativa, uma função constante ou unimodal pode ser utilizada para preenchimento, um integral que - normalizado em um integral de 1 -sobre quadrantes externos da parte zero espectral contínua depende negativamente da tonalidade, ou seja, o integral reduz com o aumento da tonalidade. Por todas estas medições, o preenchimento de ruído tende a ser menos prejudicial para as partes tonais do sinal de áudio, entretanto, sendo, todavia, efetivo para as partes não tonais do sinal de áudio em termo de redução de furos de espectro. Em outras palavras, sempre que o sinal de áudio tem um conteúdo tonal, o ruído preenchido no espectro do sinal de áudio deixa os picos tonais do espectro não afetados mantendo distância suficiente entre eles, em que, entretanto, o caráter não tonal de fases temporais do sinal de áudio com o conteúdo tonal como não tonal é, todavia,

atendido pelo preenchimento de ruído.

[0015] De acordo com uma aplicação do presente pedido, partes zero espectrais contínuas do espectro de sinal de áudio são identificadas e as partes zero identificadas são preenchidas com o ruído espectralmente formado com funções, de modo que, para cada parte zero espectral contínua a respectiva função é definida dependente de uma respectiva largura da parte zero espectral contínua e uma tonalidade do sinal de áudio. Para facilitar a implementação, a dependência pode ser obtida por uma visualização em uma tabela de visualização de funções, ou as funções podem ser calculadas analiticamente utilizando uma fórmula matemática dependendo da largura da parte zero espectral contínua e da tonalidade do sinal de áudio. Em qualquer caso, o esforço para realizar a dependência é relativamente menor comparado às vantagens resultantes da dependência. Em particular, a dependência pode ser de modo que a respectiva função é definida dependente da largura da parte zero espectral contínua de modo que a função seja confinada à respectiva parte zero espectral contínua e dependente da tonalidade do sinal de áudio de modo que, para uma tonalidade mais alta do sinal de áudio, uma massa da função se torna mais compacta na parte interna da respectiva parte zero espectral contínua e distanciada das respectivas bordas da parte zero espectral contínua.

[0016] De acordo com outra aplicação, o ruído espectralmente formado e preenchido nas partes zero espectrais contínuas é geralmente escalado utilizando um nível de preenchimento de ruído espectralmente global. Em

particular, o ruído é escalado de modo que um integral sobre o ruído nas partes zero espectrais contínuas ou um integral sobre as funções das partes zero espectrais contínuas corresponde a, por exemplo, é igual a, um nível de preenchimento de ruído global. Vantajosamente, um nível de preenchimento de ruído global é codificado dentro dos codecs de áudio existentes de modo que nenhuma sintaxe adicional tenha que ser fornecida para tais codecs de áudio. Isso é, o nível de preenchimento de ruído global pode ser explicitamente sinalizado no fluxo de dados no qual o sinal de áudio é codificado com baixo esforço. Efetivamente, as funções com o qual o ruído da parte contínua com zero é espectralmente formado podem ser escaladas de modo que um integral sobre o ruído cujas todas as partes zero espectrais contínuas são preenchidas corresponde ao nível de preenchimento de ruído global.

[0017] De acordo com uma aplicação do presente pedido, a tonalidade é derivada de um parâmetro de codificação utilizando o sinal de áudio que é codificado. Por esta medida, nenhuma informação adicional precisa ser transmitida dentro um codec de áudio existente. De acordo com as aplicações específicas, os parâmetros de codificação é um indicador ou ganho LTP (Previsão de Longa Duração | *Long-Term Prediction*), um indicador de capacitação ou ganho TNS (Formação de Ruído Temporal | *Temporal Noise Shaping*) e/ou um indicador de capacitação de remodelação do espectro.

[0018] De acordo com outra aplicação, o desempenho do preenchimento de ruído é confinado em uma parte espectral

de alta frequência, em que uma posição inicial de baixa frequência da parte espectral de alta frequência e definida correspondente a uma sinalização explícita em um fluxo de dados e na qual o sinal de áudio é codificado. Por esta medida, um ajuste adaptativo do sinal do limite inferior da parte espectral de alta frequência no qual o preenchimento de ruído é realizado, é viável. Por esta medida, por sua vez, a qualidade de áudio resultante do preenchimento de ruído pode ser elevada. A informação adicional necessária, por sua vez, causada pela sinalização explícita, é comparavelmente pequena.

[0019] O preenchimento de ruído pode ser utilizado no lado da codificação de áudio e/ou decodificação de áudio. Quando utilizado no lado da codificação de áudio, o espectro preenchido com ruído pode ser utilizado para finalidades de análise-por-síntese.

[0020] De acordo com uma aplicação, um codificador determina o nível de escala de ruído global considerando a dependência da tonalidade.

[0021] Aplicações preferidas do presente pedido são descritas abaixo com relação às figuras, entre elas:

[0022] A Figura 1a mostra um diagrama em blocos de um codificador de áudio de transformada perceptual de acordo com uma aplicação;

[0023] A Figura 1b mostra um diagrama em blocos de um decodificador de áudio de transformada perceptual de acordo com uma aplicação;

[0024] A Figura 1c mostra um diagrama esquemático

que ilustra uma possível forma de atingir a inclinação espectralmente global introduzida ao ruído preenchido de acordo com uma aplicação;

[0025] A Figura 2a mostra, em uma forma alinhada pelo tempo, um em cima do outro, de cima para baixo, um fragmento do tempo fora de um sinal de áudio, seu espectrograma utilizando uma variação espectro-temporal em "escala de cinza" esquematicamente indicada da energia espectral e a tonalidade do sinal de áudio, para finalidades de ilustração;

[0026] A Figura 2b mostra um diagrama em blocos de um aparelho de preenchimento de ruído de acordo com uma aplicação;

[0027] A Figura 3 mostra um esquema de um espectro a ser submetido ao preenchimento de ruído e uma função utilizada para espectralmente formar o ruído utilizado para preencher uma parte zero espectral contínua deste espectro de acordo com uma aplicação;

[0028] A Figura 4 mostra um esquema de um espectro a ser submetido ao preenchimento de ruído e uma função utilizada para espectralmente formar o ruído utilizado para preencher uma parte zero espectral contínua deste espectro de acordo com outra aplicação;

[0029] A Figura 5 mostra um esquema de um espectro a ser submetido ao preenchimento de ruído e uma função utilizada para espectralmente formar o ruído utilizado para preencher uma parte zero espectral contínua deste espectro de acordo com outra aplicação;

[0030] A Figura 6 mostra um diagrama em blocos do preenchedor de ruído da Figura 2 de acordo com uma aplicação;

[0031] A Figura 7 mostra esquematicamente uma possível relação entre a tonalidade do sinal de áudio determinada, por um lado, e as possíveis funções disponíveis para espectralmente formar uma parte zero espectral contínua, por outro lado, de acordo com uma aplicação;

[0032] A Figura 8 mostra esquematicamente um espectro para ser ruído preenchido adicionalmente mostrando as funções utilizadas para espectralmente formar o ruído para preenchimento das partes zero espectrais contínuas a fim de ilustrar como escalar o nível de ruído de acordo com uma aplicação;

[0033] A Figura 9 mostra um diagrama em blocos de um codificador que pode ser utilizado dentro de um codec de áudio que adota o conceito do preenchimento de ruído descrito com relação às figuras de 1 a 8;

[0034] A Figura 10 mostra esquematicamente um espectro quantizado para ser ruído preenchido como codificado pelo codificador da figura 9 com a informação adicional transmitida, a saber, os fatores de escala e nível de ruído global, de acordo com uma aplicação;

[0035] A Figura 11 mostra um diagrama em blocos de um decodificador encaixando ao codificador da figura 9 e incluindo um aparelho de preenchimento de ruído de acordo com a figura 2;

[0036] A Figura 12 mostra um esquema de um espectrograma com dados associados da informação adicional de

acordo com uma variante de uma implementação do codificador e do decodificador das figuras 9 e 11;

[0037] A Figura 13 mostra um codificador de áudio de transformada preditiva linear que pode ser incluído em um codec de áudio utilizando o conceito do preenchimento de ruído das figuras de 1 a 8 de acordo com uma aplicação;

[0038] A Figura 14 mostra um diagrama em blocos de um decodificador encaixando ao codificador da figura 13;

[0039] A Figura 15 mostra exemplos de fragmentos fora de um espectro para ser ruído preenchido;

[0040] A Figura 16 mostra um exemplo explícito para uma função para formar o ruído preenchido em uma certa parte zero espectral contínua do espectro para ser ruído preenchido de acordo com uma aplicação;

[0041] As Figuras 17a-d mostram vários exemplos para funções para espectralmente formar o ruído preenchido nas partes zero espectrais contínuas para diferentes larguras das partes zero e diferentes larguras de transição utilizadas para diferentes tonalidades.

[0042] Sempre que na seguinte descrição das figuras, os sinais de referência iguais forem utilizados para os elementos mostrados nestas figuras, a descrição apresentada com relação a um elemento em uma figura deve ser interpretada como transferível no elemento em outra figura que foi referenciada utilizando o mesmo sinal de referência. Com isso, uma descrição extensiva e repetitiva é evitada tanto quanto possível, concentrando, assim, a descrição das várias aplicações nas diferenças entre si ao invés de descrever

todas as aplicações novamente desde o início, de novo e de novo.

[0043] A figura 1a mostra um codificador de áudio de transformada perceptual de acordo com uma aplicação do presente pedido, e a figura 1b mostra um decodificador de áudio de transformada perceptual de acordo com uma aplicação do presente pedido, ambos encaixando-se para formar um codec de áudio de transformada perceptual.

[0044] Conforme mostrado na figura 1a, o codificador de áudio de transformada perceptual compreende um ponderador de espectro 1 configurado para espectralmente ponderar um espectro original do sinal de áudio recebido pelo ponderador de espectro 1 de acordo com um inverso de uma função de ponderação perceptual de ponderação espectral determinada pelo ponderador de espectro 1 em uma forma predeterminada na qual os exemplos são mostrados a seguir. O ponderador espectral 1 obtém, por esta medida, um espectro perceptualmente ponderado, que é então submetido à quantização em uma forma espectralmente uniforme, ou seja, em uma forma igual para as linhas espectrais, em um quantizador 2 do codificador de áudio de transformada perceptual. O resultado emitido pelo quantizador uniforme 2 é um espectro quantizado 34 que finalmente é codificado em um fluxo de dados emitido pelo codificador de áudio de transformada perceptual.

[0045] A fim de controlar o preenchimento de ruído a ser realizado no lado de decodificação para melhorar o espectro 34, com relação à definição do nível do ruído, um

computador do nível de ruído 3 do codificador de áudio de transformada perceptual pode opcionalmente estar presente calculando um parâmetro do nível de ruído pela medição de um nível do espectro perceptualmente ponderado 4 em partes 5 colocadas nas partes zero 40 do espectro quantizado 34. O parâmetro do nível de ruído então calculado pode ser ainda codificado no fluxo de dados previamente mencionado para chegar no decodificador.

[0046] O decodificador de áudio de transformada perceptual é mostrado na figura 1b. O mesmo compreende um aparelho de preenchimento de ruído 30 configurado para realizar o preenchimento de ruído no espectro de entrada 34 do sinal de áudio, conforme codificado no fluxo de dados gerado pelo codificador da figura 1a, pelo preenchimento do espectro 34 com ruído exibindo uma inclinação espectralmente global de modo que o nível de ruído reduz das frequências baixas a altas para obter um espectro preenchido com ruído 36. Um ruído formador de ruído de domínio de frequência do decodificador de áudio de transformada perceptual, indicado utilizando o sinal de referência 6, é configurado para submeter o espectro preenchido com ruído à formação espectral utilizando a função de ponderação perceptual espectral obtida do lado de decodificação através do fluxo de dados em uma forma descrita pelos exemplos específicos mais abaixo. Este espectro emitido pelo formador de ruído de domínio de frequência 6 pode ser encaminhado a um transformador inverso 7 a fim de reconstruir o sinal de áudio no domínio de tempo e semelhante, dentro do codificador de áudio de transformada

perceptual, um transformador 8 pode preceder o ponderador de espectro 1 a fim de fornecer o ponderador de espectro 1 com o espectro do sinal de áudio.

[0047] A significância do espectro de preenchimento 34 com ruído 9 que exibe uma inclinação espectralmente global é o seguinte: posterior, quando o espectro preenchido com ruído 36 é submetido à formação espectral pelo formador de ruído de domínio de frequência 6, o espectro 36 será submetido a uma função de ponderação inclinada. Por exemplo, o espectro será amplificado nas frequências altas quando comparado a uma ponderação das baixas frequências. Isso é, o nível de espectro 36 será elevado em frequências mais baixas com relação a frequências mais baixas. Isso causa uma inclinação espectralmente global com declive positivo nas partes planas espectralmente originais do espectro 36. Certamente, se o ruído 9 fosse preenchido ao espectro 36 para preencher as partes zero 40 deste, em uma forma espectralmente plana, então o espectro emitido por FDNS 6 mostraria dentro destas partes 40 um patamar de ruído que tende a aumentar, por exemplo, das frequências baixas a altas. Isso é, ao examinar todo o espectro ou, pelo menos, a largura de banda da parte do espectro, onde o preenchimento de ruído é realizado, um veria que o ruído dentro das partes 40 tem uma tendência ou função de regressão linear com declive positivo ou declive negativo. Como aparelho de preenchimento de ruído 30, entretanto, preenche o espectro 34 com ruído exibindo uma inclinação espectralmente global de declive positivo ou declive negativo, indicado como α na

figura 1b, e sendo inclinado na direção oposta comparado à inclinação causada por FDNS 9, a inclinação espectral causada por FDNS 6 é compensada e o patamar de ruído então introduzido ao espectro finalmente reconstruído na saída de FDNS 6 é plana ou, pelo menos, plana, assim aumentando a qualidade de áudio deixando menos furos de ruído profundos.

[0048] "Inclinação espectralmente global" deve denotar que o ruído 9 preenchido no espectro 34 tem um nível que tende a reduzir (ou aumentar) a partir de frequências baixas a altas. Por exemplo, ao colocar uma linha de regressão linear através da máxima local do ruído 9 como preenchido, por exemplo, mutualmente espectralmente distanciado, em partes zero espectrais contínuas 40, a linha de regressão linear resultante tem declive negativo (ou positivo) declive α .

[0049] Embora não seja obrigatório, o computador do nível de ruído do codificador de áudio de transformada perceptual pode explicar a forma inclinada do ruído de preenchimento ao espectro 34 pela medição do nível do espectro perceptualmente ponderado 4 nas partes 5 em uma forma ponderada com uma inclinação espectralmente global tendo, por exemplo, um declive positivo no caso de α ser negativo e declive negativo se α for positivo. O declive aplicado pelo computador do nível de ruído, que é indicado como β na figura 1a, não tem que ser o mesmo que o aplicado no lado de decodificação desde que o valor absoluto deste seja referido, mas de acordo com uma aplicação, isso este pode ser o caso. Fazer isso, o computador do nível de ruído 3

é capaz de adaptar o nível do ruído 9 inserido no lado de decodificação mais precisamente ao nível de ruído que aproxima o sinal original em uma forma melhor e pela largura de banda espectral total.

[0050] Mais tarde, será descrito que pode ser viável controlar uma variação de um declive da inclinação espectralmente global α através da sinalização explícita no fluxo de dados ou através da sinalização implícita em que, por exemplo, o aparelho de preenchimento de ruído 30 deduz o declive, por exemplo, da própria função de ponderação perceptual espectral ou de uma comutação do comprimento da janela de transformada. Por dedução, por exemplo, o declive pode ser adaptado ao comprimento da janela.

[0051] Há diferentes formas viáveis nas quais o aparelho de preenchimento de ruído 30 faz com que o ruído 9 exiba a inclinação espectralmente global. A figura 1c, por exemplo, ilustra que o aparelho de preenchimento de ruído 30 realiza uma multiplicação por linha espectral 11 entre um sinal de ruído intermediário 13, representando um estado intermediário no processo de preenchimento de ruído e uma função uniformemente decrescente (ou crescente) 15, ou seja, uma função que reduz (ou aumenta) uniformemente espectralmente por todo o espectro ou, pelo menos, a parte onde o preenchimento de ruído é realizado, para obter o ruído 9. Conforme ilustrado na figura 1c, o sinal de ruído intermediário 13 já pode ser espectralmente formado. Detalhes com relação a isso pertencem às aplicações específicas descritas mais abaixo, de acordo com o preenchimento de ruído

que também é realizado dependente da tonalidade. A formação espectral, entretanto, também pode ser deixada ou pode ser realizada após a multiplicação 11. O sinal do parâmetro do nível de ruído e o fluxo de dados podem ser utilizados para definir o nível do sinal de ruído intermediário 13, mas de modo alternativo o sinal de ruído intermediário pode ser gerado utilizando um nível padrão, aplicando o parâmetro do nível de ruído em escala para escalar a linha do espectro após a multiplicação 11. A função uniformemente decrescente 15 pode, conforme ilustrado na figura 1c, ser uma função linear, uma função linear por peças, uma função polinomial ou qualquer outra função.

[0052] Conforme será descrito em mais detalhes abaixo, seria viável definir de forma adaptativa a parte de todo o espectro dentro do preenchimento de ruído que é realizado pelo aparelho de preenchimento de ruído 30.

[0053] Em conexão com as aplicações descritas ainda abaixo, de acordo com as partes zero espectrais contínuas no espectro 34, ou seja, os furos de espectro, são preenchidos em uma forma dependente não plana específica e de tonalidade, será explicado que há ainda alternativas para a multiplicação 11 ilustrada na figura 1c a fim de provocar a inclinação espectralmente global discutida até o momento.

[0054] A seguinte descrição procede com as aplicações específicas para realizar o preenchimento de ruído. A seguir, diferentes aplicações são apresentadas para vários codecs de áudio, onde o preenchimento de ruído pode ser embutido, com especificações que aplicam em conexão com

um respectivo codec de áudio apresentado. Observa-se que o preenchimento de ruído descrito a seguir pode, em qualquer caso, ser realizado no lado de decodificação. Dependendo do codificador, entretanto, o preenchimento de ruído conforme descrito a seguir também pode ser realizado no lado de decodificação como, por exemplo, por razões de análise por síntese. Um caso intermediário de acordo com a forma modificada de preenchimento de ruído de acordo com as aplicações descritas abaixo muda parcialmente meramente a forma que o codificador opera como, por exemplo, a fim de determinar um nível de preenchimento de ruído espectralmente global, é ainda descrito abaixo.

[0055] A figura 2a mostra, para finalidades de ilustração, um sinal de áudio 10, ou seja, o curso temporal de suas amostras de áudio, por exemplo, o espectrograma alinhado pelo tempo 12 do sinal de áudio foi derivado do sinal de áudio 10, pelo menos, *inter alias*, através de uma transformação adequada como a transformação revestida ilustrada em 14 exemplar para as duas janelas de transformada consecutiva 16 e os espectros associados 18 que, assim, representa uma parte do espectrograma 12 em um período correspondente a uma metade da janela de transformada 16 associada, por exemplo. Exemplos para o espectrograma 12 e como o mesmo é derivado são apresentados mais abaixo. Em qualquer caso, o espectrograma 12 foi submetido a algum tipo de quantização e, assim, tem partes zero onde os valores espectrais nos quais o espectrograma 12 é espectro-temporalmente amostrado são continuamente zero. A

transformada revestida 14 pode, por exemplo, ser uma transformada criticamente amostrada como uma MDCT. As janelas de transformada 16 podem ter uma sobreposição de 50% um ao outro, mas diferentes aplicações são viáveis também. Ainda, a resolução espectro-temporal na qual o espectrograma 12 é amostrado nos valores espectrais pode variar no tempo. Em outras palavras, a distância temporal entre os espectros consecutivos 18 do espectrograma 12 pode variar no tempo e o mesmo se aplica à resolução espectral de cada espectro 18. Em particular, a variação no tempo na medida em que a distância temporal entre os espectros consecutivos 18 é referida, pode ser inversa à variação da resolução espectral dos espectros. A quantização utiliza, por exemplo, um tamanho da etapa de quantização adaptativo por sinal espectralmente variante, variando, por exemplo, de acordo com um envelope espectral LPC do sinal de áudio descrito pelos coeficientes LP sinalizados no fluxo de dados no qual os valores espectrais quantizados do espectrograma 12 com os espectros 18 para ser o ruído preenchido é codificado, ou de acordo com os fatores de escala determinados, por sua vez, de acordo com um modelo psicoacústico e sinalizados no fluxo de dados.

[0056] Além disso, em uma forma alinhada pelo tempo a figura 2a mostra uma característica do sinal de áudio 10 e sua variação temporal, a saber, a tonalidade do sinal de áudio. De forma geral falando, a "tonalidade" indica uma medição que descreve o quão condensado a energia do sinal de áudio está em um certo ponto no tempo no respectivo espectro 18 associado com este ponto no tempo. Se a energia for muito

espalhada, como nas fases temporais com ruído do sinal de áudio 10, então a tonalidade é baixa. Mas se a energia for substancialmente condensada a um ou mais picos espectrais, então a tonalidade é alta.

[0057] A figura 2b mostra um aparelho de preenchimento de ruído 30 configurado para realizar o preenchimento de ruído em um espectro de um sinal de áudio de acordo com uma aplicação do presente pedido. Conforme será descrito em mais detalhes abaixo, o aparelho é configurado para realizar o preenchimento de ruído dependente de uma tonalidade do sinal de áudio.

[0058] O aparelho da figura 2b compreende um preenchedor de ruído 32 e um determinador de tonalidade 34, que é opcional.

[0059] O preenchimento de ruído real é realizado pelo preenchedor de ruído 32. O preenchedor de ruído 32 recebe o espectro no qual o preenchimento de ruído deve ser aplicado. Este espectro é ilustrado na figura 2b como espectro esparso 34. O espectro esparso 34 pode ser um espectro 18 fora do espectrograma 12. Os espectros 18 entram no preenchedor de ruído 32 sequencialmente. O preenchedor de ruído 32 submete o espectro 34 ao preenchimento de ruído e emite o "espectro preenchido" 36. O preenchedor de ruído 32 realiza o preenchimento de ruído dependente de uma tonalidade do sinal de áudio, como a tonalidade 20 na figura 2a. Dependendo da circunstância, a tonalidade pode não estar diretamente disponível. Por exemplo, os codecs de áudio existentes não fornecem uma sinalização explícita da

tonalidade do sinal de áudio no fluxo de dados, de modo que se o aparelho 30 for instalado no lado de decodificação, não seria viável reconstruir a tonalidade sem um alto grau de falsa estimativa. Por exemplo, o espectro 34 pode ser, devido a sua escassez e/ou devido a sua quantização variante adaptativa por sinal, nenhuma base ótima para uma tonalidade estimativa.

[0060] Certamente, é a tarefa do determinador de tonalidade 34 fornecer o preenchedor de ruído 32 com uma estimativa da tonalidade com base no outro toque de tonalidade 38 como será descrito em mais detalhes abaixo. De acordo com as aplicações descritas posteriormente, o toque de tonalidade 38 pode estar disponível nos lados de decodificação e de codificação de qualquer forma, em forma de um respectivo parâmetro de codificação conduzido dentro do fluxo de dados do codec de áudio dentro do qual o aparelho 30 é, por exemplo, utilizado. Na figura 1b, o aparelho 30 é empregado no lado de decodificação, mas de modo alternativo, o aparelho 30 poderia também ser empregado no lado de codificação, como em um circuito de retorno de previsão do codificador da figura 1a se presente.

[0061] A figura 3 mostra um exemplo para o espectro esparso 34, ou seja, um espectro quantizado tendo partes contínuas 40 e 42 consistindo em testes de valores espectrais espectralmente próximos do espectro 34, sendo quantizados a zero. As partes contínuas 40 e 42 são, então, espectralmente separadas ou distanciadas entre si através de pelo menos uma linha espectral não quantizada a zero no espectro 34.

[0062] A dependência da tonalidade do preenchimento de ruído geralmente descrita acima com relação à figura 2b pode ser implementada como segue. A figura 3 mostra uma parte temporal 44 incluindo uma parte zero espectral contínua 40, exagerada em 46. O preenchedor de ruído 32 é configurado para preencher esta parte zero espectral contínua 40 em uma forma dependente da tonalidade do sinal de áudio no tempo a qual o espectro 34 pertence. Em particular, o preenchedor de ruído 32 preenche a parte zero espectral contínua com ruído espectralmente formado utilizando uma função assumindo um máximo em uma parte interna da parte zero espectral contínua e tendo bordas exteriormente descendentes, um declive absoluto que depende negativamente da tonalidade. A figura 3 de forma exemplar mostra as duas funções 48 para duas tonalidades diferentes. Ambas as funções são "unimodal", ou seja, assumem um máximo absoluto na parte interna da parte zero espectral contínua 40 e têm meramente um máximo local que pode ser uma estagnação ou uma única frequência espectral. Aqui, o máximo local é assumido pelas funções 48 e 50 continuamente sobre um intervalo estendido 52, ou seja, uma estagnação, disposta no centro da parte zero 40. O domínio das funções 48 e 50 é a parte zero 40. O intervalo central 52 cobre meramente uma parte central da parte zero 40 e é acompanhada por uma parte da borda 54 em um lado com frequência mais alta do intervalo 52 e uma parte da borda com frequência mais baixa 56 em um lado com frequência mais baixa do intervalo 52. Dentro da parte da borda 54, as funções 48 e 52 tem uma borda de descida 58 e dentro da parte da borda 56,

uma borda de subida 60. Um declive absoluto pode ser atribuído a cada borda 58 e 60, respectivamente, como o declive médio dentro da parte da borda 54 e 56, respectivamente. Isso é, o declive atribuído à borda de descida 58 pode ser o declive médio da respectiva função 48 e 52, respectivamente, dentro da parte da borda 54 e o declive atribuído à borda de subida 60 pode ser o declive médio da função 48 e 52, respectivamente, dentro da parte da borda 56.

[0063] Como pode ser visto, o valor absoluto do declive das bordas 58 e 60 é mais alto para a função 50 do que para a função 48. O preenchedor de ruído 32 seleciona para preencher a parte zero 40 com a função 50 para tonalidades mais baixas do que as tonalidades nas quais o preenchedor de ruído 32 seleciona para utilizar a função 48 para preenchimento da parte zero 40. Por esta medida, o preenchedor de ruído 32 evita agrupar a periferia imediata dos picos espectrais potencialmente tonais do espectro 34, como, por exemplo, o pico 62. Quanto menor o declive absoluto das bordas 58 e 60, mais longe o ruído preenchido na parte zero 40 ficará das partes não zero do espectro 34 ao redor da parte zero 40.

[0064] O preenchedor de ruído 32 pode, por exemplo, escolher selecionar a função 48 no case da tonalidade do sinal de áudio ser τ_2 , e a função 50 no caso da tonalidade do sinal de áudio ser τ_1 , mas a descrição trazida mais abaixo revelará que o preenchedor de ruído 32 pode discriminar mais do que dois estados diferentes de tonalidade do sinal de áudio, ou seja, pode suportar mais do que duas funções 48, 50

diferentes para preenchimento de uma certa parte zero espectral contínua e escolher entre estas dependente da tonalidade através de um mapeamento subjetivo das tonalidades às funções.

[0065] Como uma observação menor, observa-se que a construção das funções 48 e 50 de acordo com as mesmas tem uma estagnação na parte interna intervalo 52, acompanhadas pelas bordas 58 e 60 para resultar em funções unimodais, é meramente um exemplo. De modo alternativo, as funções em forma de sino podem ser utilizadas, por exemplo, de acordo com uma alternativa. O intervalo 52 pode, de modo alternativo, ser definido como o intervalo entre o qual a função é mais alta do que 95% de seu valor máximo.

[0066] A figura 4 mostra uma alternativa para a variação da função utilizada para espectralmente formar o ruído com o qual uma determinada parte zero espectral contínua 40 é preenchida pelo preenchedor de ruído 32, na tonalidade. De acordo com a figura 4, a variação pertence à largura espectral das partes da borda 54 e 56 e das bordas exteriormente descendentes 58 e 60, respectivamente. Conforme mostrado na figura 4, de acordo com o exemplo da figura 4, o declive das bordas' 58 e 60 pode ainda ser independente, ou seja, não alterado de acordo com, da tonalidade. Em particular, de acordo com o exemplo da figura 4, o preenchedor de ruído 32 define a função utilizando o ruído para preenchimento da parte zero 40 que é espectralmente formado de modo que a largura espectral das bordas exteriormente descendentes 58 e 60 dependa positivamente da

tonalidade, ou seja, para tonalidades mais altas, a função 48 é utilizada para a largura espectral das bordas exteriormente descendentes 58 e 60 que é maior, e para as tonalidades mais baixas, a função 50 é utilizada para a largura espectral das bordas exteriormente descendentes 58 e 60 que é menor.

[0067] A figura 4 mostra outro exemplo de uma variação de uma função utilizada pelo preenchedor de ruído 32 para espectralmente formar o ruído com a parte zero espectral contínua 40 que é preenchida: aqui, a característica da função que varia com a tonalidade é integral sobre os quadrantes externos da parte zero 40. Quanto mais alta a tonalidade, maior o intervalo. Antes de determinar o intervalo, o intervalo geral da função sobre a parte zero completa 40 é equalizado/normalizado como para 1.

[0068] A fim de explicar isso, veja a figura 5. A parte zero espectral contínua 40 é mostrada para ser dividida em quatro quadrantes com tamanho igual a, b, c, d, entre os quais os quadrantes a e d são quadrantes externos. Como pode ser visto, ambas as funções 50 e 48 têm seu centro de massa na parte interna, aqui de forma exemplar no meio da parte zero 40, mas ambas delas se estendem dos quadrantes internos b, c aos quadrantes externos a e d. A parte de sobreposição das funções 48 e 50, sobrepondo os quadrantes externos a e d, respectivamente, é mostrada simplesmente sombreada.

[0069] Na figura 5, ambas as funções têm o mesmo integral sobre toda a parte zero 40, ou seja, sobre todos os quatro quadrantes a, b, c, d. O integral é, por exemplo, normalizado em 1.

[0070] Nesta situação, o integral da função 50 sobre os quadrantes a, d é maior do que o integral da função 48 sobre quadrantes a, d e certamente, o preenchedor de ruído 32 utiliza a função 50 para tonalidades mais altas e a função 48 para as tonalidades mais baixas, ou seja, o integral sobre os quadrantes externos das funções normalizadas 50 e 48 depende negativamente da tonalidade.

[0071] Para finalidades de ilustração, no caso da figura 5 ambas as funções 48 e 50 foram, de forma exemplar, mostradas como funções constantes ou binárias. A função 50, por exemplo, é uma função assumindo um valor constante sobre todo o domínio, ou seja, toda a parte zero 40 e a função 48 é uma função binária sendo zero nas bordas externas da parte zero 40 e assumindo um valor constante não zero entre elas. Deve estar claro que, de forma geral falando, as funções 50 e 48 de acordo com o exemplo da figura 5 podem ser qualquer função constante ou unimodal como as correspondentes àquelas mostradas nas figuras 3 e 4. Para ser ainda mais preciso, pelo menos uma pode ser unimodal e pelo menos uma constante (por peças) e potencial qualquer uma da unimodal ou constante.

[0072] Embora o tipo de variação de funções 48 e 50 dependendo da tonalidade varie, todos os exemplos das figuras de 3 a 5 têm em comum que, para a tonalidade crescente, o grau de dispersão imediata ao redor dos picos tonais no espectro 34 é reduzido ou evitado de modo que a qualidade do preenchimento de ruído seja elevada visto que o preenchimento de ruído não afeta negativamente as fases tonais do sinal de

áudio e, todavia, resulta em uma aproximação agradável de fases não tonais do sinal de áudio.

[0073] Até agora, a descrição das figuras de 3 a 5 focou no preenchimento de uma parte zero espectral contínua. De acordo com a aplicação da figura 6, o aparelho da figura 2b é configurado para identificar as partes zero espectrais contínuas do sinal de áudio e aplicar o preenchimento de ruído nas partes zero espectrais contínuas então identificadas. Em particular, a figura 6 mostra o preenchedor de ruído 32 da figura 2b em mais detalhes como compreendendo um identificador da parte zero 70 e um preenchedor da parte zero 72. O identificador da parte zero busca no espectro 34 por partes zero espectrais contínuas como 40 e 42 na figura 3. Conforme já descrito acima, as partes zero espectrais contínuas podem ser definidas como testes de valores espectrais que foram quantizados a zero. O identificador da parte zero 70 pode ser configurado para confinar a identificação em uma parte espectral de alta frequência do sinal de áudio espectro começando, ou seja, acima de, alguma frequência inicial. Certamente, o aparelho pode ser configurado para confinar o desempenho do preenchimento de ruído em tal parte espectral de alta frequência. A frequência inicial acima do identificador da parte zero 70 que realiza a identificação das partes zero espectrais contínuas, e acima do aparelho que é configurado para confinar o desempenho do preenchimento de ruído, pode ser fixada ou pode variar. Por exemplo, a sinalização explícita em um fluxo de dados do sinal de áudio na qual o sinal de áudio é codificado através

de seu espectro pode ser utilizada para sinalizar a frequência inicial a ser utilizada.

[0074] O preenchedor da parte zero 72 é configurado para preencher as partes zero espectrais contínuas identificadas pelo identificador 70 com ruído espectralmente formado de acordo com uma função conforme descrito acima com relação à figura 3, 4 ou 5. Certamente, o preenchedor da parte zero 72 preenche as partes zero espectrais contínuas identificadas pelo identificador 70 com funções definidas dependentes de uma respectiva largura da parte zero espectral contínua, como o número de valores espectrais que foram quantizados a zero do teste de valores espectrais quantizados por zero da respectiva parte zero espectral contínua e da tonalidade do sinal de áudio.

[0075] Em particular, o preenchimento individual de cada parte zero espectral contínua identificada pelo identificador 70 pode ser realizado pelo preenchedor 72 como segue: a função é definida dependente da largura da parte zero espectral contínua de modo que a função seja confinada à respectiva parte zero espectral contínua, ou seja, o domínio da função coincide com a largura da parte zero espectral contínua. A definição da função é ainda dependente da tonalidade do sinal de áudio, a saber na forma descrita acima com relação às figuras de 3 a 5, de modo que se a tonalidade do sinal de áudio aumentar, a massa da função se torna mais compacta na parte interna da respectiva parte zero contínua e distanciada das respectivas bordas da parte zero espectral contínua. Utilizando esta função, um estado preliminarmente

preenchido da parte zero espectral contínua de acordo com cada um dos valores espectrais que é definido como um valor aleatório, pseudoaleatório ou de patch/copiado, é espectralmente formado, a saber pela multiplicação da função com os valores espectrais preliminares.

[0076] Já foi descrito acima que a dependência do preenchimento de ruído da tonalidade pode discriminar entre mais do que apenas duas tonalidades diferentes como 3, 4 ou ainda mais do que 4. A figura 7, por exemplo, mostra o domínio de possíveis tonalidades, ou seja, o intervalo de possíveis valores de intertonalidade, conforme determinado pelo determinador 34 no sinal de referência 74. Em 76, a figura 7 de forma exemplar mostra o conjunto de possíveis funções utilizado para espectralmente formar o ruído com o qual as partes zero espectrais contínuas podem ser preenchidas. O conjunto 76, conforme ilustrado na figura 7, é um conjunto de instanciações de função discreta mutualmente distinguindo entre si pela largura espectral ou comprimento de domínio e/ou formação, ou seja, compacidade e distância das bordas externas. Em 78, a figura 7 ainda mostra o domínio de possíveis larguras da parte zero. Enquanto o intervalo 78 é um intervalo de valores discretos que variam de uma largura mínima a uma largura máxima, os valores de tonalidade emitidos pelo determinador 34 para medir a tonalidade do sinal de áudio podem ser tanto com valor de número inteiro ou de algum outro tipo, como valores do ponto de flutuação. O mapeamento do par de intervalos 74 e 78 ao conjunto de possíveis funções 76 pode ser realizado pela visualização da

tabela ou utilizando uma função matemática. Por exemplo, para uma certa parte zero espectral contínua identificada pelo identificador 70, o preenchedor da parte zero 72 pode utilizar a largura da respectiva parte zero espectral contínua e a tonalidade atual conforme determinado pelo determinador 34 para visualização em uma tabela em que uma função do conjunto 76 é definida, por exemplo, como uma sequência de valores de função, o comprimento da sequência que coincide com a largura da parte zero espectral contínua. De modo alternativo, o preenchedor da parte zero 72 visualiza os parâmetros da função e preenche estes parâmetros da função em uma função predeterminada para derivar a função a ser utilizada para espectralmente formar o ruído a ser preenchido na respectiva parte zero espectral contínua. Em outra alternativa, o preenchedor da parte zero 72 pode inserir diretamente a respectiva largura da parte zero espectral contínua e a tonalidade atual em uma fórmula matemática a fim de chegar nos parâmetros da função a fim de criar a respectiva função de acordo com o parâmetro da função matematicamente calculado.

[0077] Até agora, a descrição de certas aplicações do presente pedido focou na formação da função utilizada para espectralmente formar o ruído com o qual certas partes zero espectrais contínuas são preenchidas. É vantajoso, entretanto, controlar todo o nível de ruído adicionado a um certo espectro para ser o ruído preenchido para resultar em uma reconstrução agradável, ou para ainda controlar a introdução do nível de ruído espectralmente.

[0078] A figura 8 mostra um espectro para ser ruído preenchido, onde as partes não quantizadas a zero e certamente, não submetidas ao preenchimento de ruído, são indicadas com sombreado ondulado, em que as três partes zero espectrais contínuas 90, 92 e 94 são mostradas em um estado pré-preenchido sendo ilustrado pelas partes zero tendo inscritas nelas a função selecionada para formação espectral do ruído preenchido nestas partes 90-94, utilizando uma escala sem preocupação.

[0079] De acordo com uma aplicação, o conjunto de funções 48, 50 disponível para espectralmente formar o ruído a ser preenchido nas partes 90-94, tem uma escala predefinida que é conhecida pelo codificador e decodificador. Um fator de escala espectralmente global é sinalizado explicitamente dentro do fluxo de dados no qual o sinal de áudio, ou seja, a parte não quantizada do espectro, é codificada. Este fator indica, por exemplo, RMS ou outra medição para um nível de ruído, ou seja, valores da linha espectral aleatória ou pseudoaleatória, com as partes 90-94 que são predefinidas no lado de decodificação sendo, então, espectralmente formadas utilizando as funções selecionadas dependentemente da tonalidade 48, 50 como elas são. Como para a forma que o fator de escala do ruído global poderia ser determinado no lado do decodificador é descrita mais abaixo. Por exemplo, deixamos A ser o conjunto de índices i das linhas espectrais, onde o espectro é quantizado a zero e que pertence a qualquer uma das partes 90-94, e deixamos N denotar o fator de escala do ruído global. Os valores do espectro devem ser denotados

como x_i . Adicionalmente, " $random(N)$ [aleatório(N)]" deve denotar uma função que fornece um valor aleatório de um nível correspondente ao nível " N " e $left(i)$ [esquerdo(i)] deve ser uma função indicando, para qualquer valor espectral quantizado a zero no índice i , o índice do valor quantizado a zero na extremidade de baixa frequência da parte zero a qual i pertence, e $F_i(j)$ com $j=0$ a $J_i - 1$ deve denotar a função 48 ou 50 atribuída, dependendo da tonalidade, a parte zero 90-94 começando no índice i , com J_i indicando a largura desta parte zero. Então, as partes 90-94 são preenchidas de acordo com $x_i = F_{left(i)}(i - left(i)) \cdot random(N)$.

[0080] Adicionalmente, o preenchimento de ruído nas partes 90-94, pode ser controlado de modo que o nível de ruído reduza das frequências baixas a altas. Isso pode ser feito pela formação espectral do ruído com o qual as partes são pré-definidas, ou espectralmente formando a disposição das funções 48,50 de acordo com uma função de transferência do filtro passa-baixa. Isso pode compensar uma inclinação espectral causada quando a reescalada/dequantização do espectro preenchido devido a, por exemplo, uma pré-ênfase utilizada ao determinar o curso espectral do tamanho da etapa de quantização. Certamente, o declive da redução ou a função de transferência do filtro passa-baixa pode ser controlado de acordo com um grau de pré-ênfase aplicado. Aplicar a nomenclatura utilizada acima, as partes 90-94 podem ser preenchidas de acordo com $x_i = F_{left(i)}(i - left(i)) \cdot random(N) \cdot LPF(i)$ com $LPF(i)$ denotando a função de transferência do filtro de baixa frequência que pode ser linear. Dependendo

das circunstâncias, a função LPF que corresponde à função 15 pode ter um declive positivo e LPF mudado para ler HPF certamente.

[0081] Ao invés de utilizar uma escala fixa das funções selecionadas dependendo da tonalidade e da largura da parte zero, a correção espectral de inclinação agora descrita pode ser diretamente explicada utilizando a posição espectral da respectiva parte zero contínua ainda como um índice na visualização ou, caso contrário, determinando 80 a função para ser utilizada para formação espectral do ruído com o qual a respectiva parte zero espectral contínua deve ser preenchida. Por exemplo, um valor médio da função ou sua pré-escala utilizada para espectralmente formar o ruído a ser preenchido em uma certa parte zero 90-94 pode depender da posição espectral da parte zero 90-94 de modo que, sobre toda a largura de banda do espectro, as funções utilizadas para as partes zero espectrais contínuas 90-94 são pré-escaladas para simular uma função de transferência do filtro passa-baixa para compensar qualquer função de transferência de pré-ênfase passa-alta utilizada para derivar as partes do espectro quantizadas a não zero.

[0082] Finalmente, observa-se que enquanto a figura 8 de forma exemplar referida à aplicação utilizando o preenchimento de ruído espectralmente formado das partes zero espectrais contínuas, o mesmo pode ser, de modo alternativo, modificado para se referir às aplicações sem utilizar o preenchimento de ruído formado espectral, mas preenchendo as partes zero espectrais contínuas em uma forma espectralmente

plana, por exemplo. Assim, as partes 90-94 então seriam preenchidas de acordo com $x_i = \text{LPF}(i) \cdot \text{random}(N)$.

[0083] Tendo descrito as aplicações para realizar o preenchimento de ruído, nas seguintes aplicações os codecs de áudio são apresentados onde o preenchimento de ruído descrito acima pode ser vantajosamente embutido. As figuras 9 e 10, por exemplo, mostram um par de um codificador e um decodificador, respectivamente, juntos implementando um codec de áudio perceptual com base na transformada do tipo que forma a base, por exemplo, de AAC (Codificação Avançada de Áudio | *Advanced Audio Coding*). O codificador 100 mostrado na figura 9 submete o sinal de áudio original 102 a uma transformada em um transformador 104. A transformação realizada pelo transformador 104 é, por exemplo, uma transformada revestida que corresponde a uma transformação 14 da figura 1: ela decompõe espectralmente o sinal de áudio original de entrada 102 submetendo as janelas de transformada de sobreposição mutualmente consecutivas do sinal de áudio original em uma sequência de espectros 18 juntos compondo o espectrograma 12. Conforme denotado acima, a patch da janela de intertransformada que define a resolução temporal do espectrograma 12 pode variar no tempo, apenas como o comprimento temporal das janelas de transformada pode definir a resolução espectral de cada espectro 18. O codificador 100 ainda compreende um modelador perceptual 106 que deriva do sinal de áudio original, com base na versão do domínio de tempo que entra no transformador 104 ou a versão decomposta espectralmente emitida pelo transformador 104, um limite de

mascamamento perceptual que define uma curva espectral abaixo da qual o ruído de quantização pode ser oculto de modo que o mesmo não é perceptível.

[0084] A representação por linha espectral do sinal de áudio, ou seja, o espectrograma 12, e o limite de mascaramento entram no quantizador 108 que é responsável por quantizar as amostras espectrais do espectrograma 12 utilizando um tamanho da etapa de quantização espectralmente variante que depende do limite de mascaramento: quanto maior o limite de mascaramento, menor o tamanho da etapa de quantização. Em particular, o quantizador 108 informa o lado de decodificação da variação do tamanho da etapa de quantização na forma dos chamados fatores de escala que, em forma da relação agora descrita entre o tamanho da etapa de quantização, por um lado, e o limite de mascaramento perceptual, por outro, representam um tipo de representação do próprio limite de mascaramento perceptual. A fim de encontrar um bom compromisso entre a quantidade de informação adicional a ser gasta para transmitir os fatores de escala ao lado de decodificação, e a granularidade para adaptar o ruído de quantização ao limite de mascaramento perceptual, o quantizador 108 define/varia os fatores de escala em uma resolução espectro-temporal que é menor do que, ou mais grosseira do que, a resolução espectro-temporal na qual os níveis espectrais quantizados descrevem a representação por linha espectral do espectrograma do sinal de áudio 12. Por exemplo, o quantizador 108 subdivide cada espectro em bandas do fator de escala 110 como bandas de Bark, e transmite um

fator de escala por banda do fator de escala 110. Desde que a resolução temporal seja referida, a mesma também ser menor desde que a transmissão dos fatores de escala seja referida, comparado aos níveis espectrais dos valores espectrais do espectrograma 12.

[0085] Ambos os níveis espectrais dos valores espectrais do espectrograma 12, bem como os fatores de escala 112 são transmitidos ao lado de decodificação. Entretanto, a fim de melhorar a qualidade de áudio, o codificador 100 transmite dentro do fluxo de dados ainda em um nível de ruído global que sinaliza ao lado de decodificação do nível de ruído até as partes quantizadas a zero da representação 12 serem preenchidas com ruído antes da reescala, ou dequantização, do espectro aplicando os fatores de escala 112. Isso é mostrado na figura 10. A figura 10 mostra, utilizando sombreado ondulado, o espectro ainda não reescalado do sinal de áudio como 18 na figura 9. Ela tem partes zero espectrais contínuas 40a, 40b, 40c e 40d. O nível de ruído global 114 que também pode ser transmitido no fluxo de dados para cada espectro 18, indica ao decodificador o nível no qual estas partes zero de 40a a 40d devem ser preenchidas com ruído antes de submeter este espectro preenchido à reescala ou requantização utilizando os fatores de escala 112.

[0086] Conforme já denotado acima, o preenchimento de ruído no qual o nível de ruído global 114 se refere, pode ser submetido a uma restrição em que este tipo de preenchimento de ruído meramente refere-se a frequências

acima de alguma frequência inicial que é indicada na figura 10 meramente para finalidades de ilustração como f_{start} .

[0087] A figura 10 ainda ilustra outra característica específica, que pode ser implementada no codificador 100: pois podem haver espectros 18 compreendendo bandas do fator de escala 110 onde todos os valores espectrais dentro das respectivas bandas do fator de escala foram quantizadas a zero, o fator de escala 112 associado com esta banda do fator de escala é realmente supérfluo. Certamente, o quantizador 100 utiliza este fator de escala para preencher individualmente a banda do fator de escala com ruído além do ruído preenchido na banda do fator de escala utilizando o nível de ruído global 114, ou em outros termos, a fim de escalar o ruído atribuído à respectiva banda do fator de escala receptiva ao nível de ruído global 114. Veja, por exemplo, a figura 10. A figura 10 mostra uma subdivisão exemplar do espectro 18 em bandas do fator de escala de 110a a 110h. A banda do fator de escala 110e é uma banda do fator de escala, os valores espectrais que foram todos quantizados a zero. Certamente, o fator de escala associado 112 é "livre" e é utilizado para determinar 114 o nível do ruído no qual esta banda do fator de escala é preenchida completamente. As outras bandas do fator de escala que compreendem os valores espectrais quantizados a níveis não zero, têm fatores de escala associados com eles que são utilizados para reescalar os valores espectrais do espectro 18 que não foi quantizado a zero, incluindo o ruído utilizando as partes zero de 40a a 40d que foram preenchidas, cuja escala é indicada utilizando

a seta 116, de forma representativa.

[0088] O codificador 100 da figura 9 já pode considerar que dentro do lado de decodificação o preenchimento de ruído utilizando o nível de ruído global 114 será realizado utilizando o preenchimento de ruído das aplicações descritas acima, por exemplo, utilizando uma dependência da tonalidade e/ou impondo uma inclinação espectralmente global no ruído e/ou variando a frequência inicial do preenchimento de ruído e assim por diante.

[0089] Desde que a dependência da tonalidade é referida, o codificador 100 pode determinar o nível de ruído global 114 e inserir o mesmo ao fluxo de dados, associando às partes zero de 40a a 40d da função para espectralmente formar o ruído para preenchimento da respectiva parte zero. Em particular, o codificador pode utilizar estas funções a fim de ponderar o original, ou seja, os valores espectrais do sinal de áudio ponderados, mas ainda não quantizados, nestas partes de 40a a 40d a fim de determinar o nível de ruído global 114. Assim, o nível de ruído global 114 determinado e transmitido dentro do fluxo de dados, leva a um preenchimento de ruído no lado de decodificação que recupera mais proximamente o espectro original do sinal de áudio.

[0090] O codificador 100 pode, dependendo do conteúdo do sinal de áudio, decidir se utilizar algumas opções de codificação que, por sua vez, podem ser utilizadas como toques de tonalidade como o toque de tonalidade 38 mostrado na figura 2 para permitir que o lado de decodificação defina corretamente a função para

espectralmente formar o ruído utilizado para preencher as partes de 40a a 40d. Por exemplo, o codificador 100 pode utilizar a previsão temporal a fim de prever um espectro 18 de um espectro prévio utilizando um chamado parâmetro de ganho da previsão de longa duração. Em outras palavras, o ganho de previsão de longa duração pode definir o grau no qual a previsão temporal é utilizada ou não. Certamente, o ganho de previsão de longa duração, ou ganho LTP, é um parâmetro que pode ser utilizado como um toque de tonalidade, pois quanto mais alto o ganho LTP, mais alta a tonalidade do sinal de áudio provavelmente será. Assim, o determinador de tonalidade 34 da figura 2, por exemplo, pode definir a tonalidade de acordo com uma dependência monótona positiva do ganho LTP. Ao invés disso, ou além disso, um ganho LTP, o fluxo de dados pode compreender uma sinalização do indicador de capacitação LTP ligando/desligando a LTP, assim ainda revelando um toque com valor binário referente à tonalidade, por exemplo.

[0091] De forma adicional ou alternativa, o codificador 100 pode suportar a formação do ruído temporal. Isso é, em uma base por espectro 18, por exemplo, o codificador 100 pode escolher submeter o espectro 18 à formação do ruído temporal com indicação dessa decisão em forma de um indicador de capacitação da formação do ruído temporal ao decodificador. O indicador de capacitação TNS indica se os níveis espectrais do espectro 18 formam a previsão residual de uma previsão espectral, ou seja, ao longo da direção de frequência determinada, a previsão linear

do espectro ou se o espectro não for previsto por LP. Se TNS for sinalizado para ser habilitado, o fluxo de dados adicionalmente compreende os coeficientes da previsão linear para espectralmente prever linearmente o espectro de modo que o decodificador possa recuperar o espectro utilizando estes coeficientes da previsão linear aplicando o mesmo no espectro antes ou após a reescala ou dequantização. O indicador de capacitação TNS é ainda um toque de tonalidade: se o indicador de capacitação TNS sinalizar que o TNS deve ser ligado, por exemplo, em um transiente, então o sinal de áudio muito provavelmente será tonal, pois o espectro parece ser bem previsível pela previsão linear ao longo do eixo de frequência e, assim, não estacionário. Certamente, a tonalidade pode ser determinada com base no indicador de capacitação TNS de modo que a tonalidade seja mais alta se o indicador de capacitação TNS desabilitar TNS e for menor se o indicador de capacitação TNS sinalizar a capacitação de TNS. Ao invés de, ou além disso de um indicador de capacitação TNS, pode ser possível derivar dos coeficientes do filtro TNS um ganho TNS indicando um grau no qual TNS pode ser utilizado para prever o espectro, assim, ainda revelando um toque com valor maior do que dois referentes à tonalidade.

[0092] Outros parâmetros de codificação também podem ser codificados dentro do fluxo de dados pelo codificador 100. Por exemplo, um indicador de capacitação de remodelação espectral pode sinalizar uma opção de codificação de acordo com o espectro 18 que é codificado pela remodelação dos níveis espectrais, ou seja, dos valores espectrais

quantizados, espectralmente, de forma adicional, transmitindo dentro do fluxo de dados a prescrição da remodelação de modo que o decodificador possas remodular, ou reescalar, os níveis espectrais para recuperar o espectro 18. Se o indicador de capacitação de remodelação do espectro for habilitado, ou seja, a remodulação do espectro for aplicada, isso indica que o sinal de áudio provavelmente será tonal, pois a remodulação tende a ser mais efetiva por taxa/distorção ao comprimir o fluxo de dados se houver muitos picos tonais dentro do espectro. Certamente, de forma adicional ou alternativa, o indicador de capacitação de remodelação do espectro pode ser utilizado como um toque tonal e a tonalidade utilizada para preenchimento de ruído pode ser definida como maior no caso do indicador de capacitação de remodelação do espectro sendo habilitado, e inferior se o indicador de capacitação de disposição do espectro for desabilitado.

[0093] Para fins de conclusã⁹, e ainda com referência à figura 2b, observa-se que o número de diferentes funções par espectralmente formar uma parte zero de 40a a 40d, ou seja, o número de tonalidades diferentes discriminadas para definição da função para espectralmente formar, pode, por exemplo, ser maior do que quatro, ou ainda maior do que oito, pelo menos, para as larguras das partes zero espectrais contínuas acima de uma largura mínima predeterminada.

[0094] Desde que o conceito para impor uma inclinação espectralmente global no ruído e considerando o mesmo ao calcular o parâmetro do nível de ruído no lado de

decodificação seja referido, o codificador 100 pode determinar o nível de ruído global 114 e inserir o mesmo ao fluxo de dados, pela ponderação das partes dos valores espectrais do sinal de áudio ainda não quantizados, mas com o inverso dos valores espectrais do sinal de áudio ponderados da função de ponderação perceptual, espectralmente colocados nas partes zero de 40a a 40d, com uma função espectralmente se estendendo, pelo menos, sobre toda a largura de banda do preenchimento de ruído da parte do espectro e tendo um declive do sinal oposto com relação à função 15 utilizada no lado de decodificação para preenchimento de ruído, por exemplo, e medição do nível com base nos valores não quantizados não ponderados.

[0095] A figura 11 mostra um decodificador encaixando ao codificador da figura 9. O decodificador da figura 11 é geralmente indicado utilizando o sinal de referência 130 e compreende um preenchedor de ruído 30 correspondente às aplicações descritas acima, um dequantizador 132 e um transformador inverso 134. O preenchedor de ruído 30 recebe a sequência de espectros 18 dentro do espectrograma 12, ou seja, a representação por linha espectral incluindo os valores espectrais quantizados, e, opcionalmente, os toques de tonalidade do fluxo de dados como um ou vários dos parâmetros de codificação discutidos acima. O preenchedor de ruído 30 então preenche as partes zero espectrais contínuas de 40a a 40d com ruído, conforme descrito acima, como utilizando a dependência da tonalidade descrita acima e/ou impondo uma inclinação espectralmente

global no ruído, e utilizando o nível de ruído global 114 para escalar o nível de ruído conforme descrito acima. Assim preenchidos, estes espectros atingem o dequantizador 132, que, por sua vez, dequantiza ou reescala o espectro preenchido com ruído utilizando os fatores de escala 112. O transformador inverso 134, por sua vez, submete o espectro dequantizado a uma transformação inversa para recuperar o sinal de áudio. Conforme descrito acima, a transformação inversa 134 pode compreender, ainda, um processo de adição por sobreposição, a fim de atingir o cancelamento da distorção do domínio de tempo causado no caso da transformação utilizada pelo transformador 104 sendo uma transformada revestida criticamente amostrada como uma MDCT, neste caso a transformação inversa aplicada pelo transformador inverso 134 seria uma IMDCT (MDCT inversa).

[0096] Conforme já descrito com relação às figuras 9 e 10, o dequantizador 132 aplica os fatores de escala ao espectro pré-preenchido. Isso é, os valores espectrais dentro das bandas do fator de escala não completamente quantizados a zero são escalados utilizando o fator de escala independentemente do valor espectral que representa um valor espectral não zero ou um ruído que foi espectralmente formado pelo preenchedor de ruído 30 conforme descrito acima. As bandas espectrais quantizadas completamente a zero têm fatores de escala associados com elas, que são completamente livres para controlar o preenchimento de ruído e o preenchedor de ruído 30 pode tanto utilizar este fator de escala para individualmente escalar o ruído com o qual a

banda do fator de escala foi preenchida em forma do preenchimento de ruído do preenchedor de ruído 30 das partes zero espectrais contínuas, ou o preenchedor de ruído 30 pode utilizar o fator de escala para adicionalmente preencher, ou seja, adicionar o ruído adicional desde que estas bandas espectrais quantizadas a zero sejam referidas.

[0097] Observa-se que o ruído cujo preenchedor de ruído 30 espectralmente forma na maneira dependente da tonalidade descrita acima e/ou submete a uma inclinação espectralmente global em uma forma descrita acima, pode resultar de uma fonte de ruído pseudoaleatório, ou pode ser derivado do preenchedor de ruído 30 com base na cópia ou interligação espectral de outras áreas do mesmo espectro ou espectros relacionados, como um espectro alinhado pelo tempo de outro canal, ou um espectro temporariamente precedente. A interligação uniforme do mesmo espectro pode ser viável, como cópia de áreas da frequência inferior do espectro 18 (cópia espectral). Independentemente da forma que o preenchedor de ruído 30 deriva o ruído, o preenchedor 30 espectralmente forma o ruído para preenchimento em partes zero espectrais contínuas de 40a a 40d na forma dependente da tonalidade descrita acima e/ou submete o mesmo a uma inclinação espectralmente global em uma forma descrita acima.

[0098] Para fins de conclusão apenas, é mostrado na figura 12 que as aplicações do codificador 100 e do decodificador 130 das figuras 9 e 11 podem ser variadas em que a justaposição entre os fatores de escala, por um lado, e os níveis de ruído específico por fator de escala são

diferentemente implementados. De acordo com o exemplo da figura 12, o codificador transmite dentro do fluxo de dados a informação de um envelope do ruído, espectro-temporalmente amostrado em uma resolução mais grossa do que a resolução espectral por linhas do espectrograma 12, como, por exemplo, na mesma resolução espectro-temporal que os fatores de escala 112, além dos fatores de escala 112. Esta informação do envelope do ruído é indicada utilizando o sinal de referência 140 na figura 12. Por esta medida, para bandas do fator de escala não completamente quantizadas a zero dois valores existem: um fator de escala para reescala ou dequantização dos valores espectrais não zero dentro desta respectiva banda do fator de escala, bem como um nível de ruído 140 para a banda do fator de escala individual escalando o ruído nível dos valores espectrais quantizados por zero dentro desta banda do fator de escala. Este conceito é, às vezes, chamado IGF (Preenchimento de Lacuna Inteligente | *Intelligent Gap Filling*).

[0099] Ainda aqui, o preenchedor de ruído 30 pode aplicar o preenchimento dependente da tonalidade das partes zero espectrais contínuas de 40a a 40d de forma exemplar conforme mostrado na figura 12.

[0100] De acordo com os exemplos do codec de áudio descritos acima com relação às figuras de 9 a 12, a formação espectral do ruído de quantização foi realizada pela transmissão de uma informação referente ao limite de mascaramento perceptual utilizando uma representação espectro-temporal na forma de fatores de escala. As figuras

13 e 14 mostram um par de codificador e decodificador onde ainda as aplicações do preenchimento de ruído descritas com relação às figuras de 1 a 8 podem ser utilizadas, mas onde o ruído de quantização é espectralmente formado de acordo com uma descrição da LP (Previsão Linear | *Linear Prediction*) do espectro do sinal de áudio. Em ambas as aplicações, o espectro a ser ruído preenchido está no domínio ponderado, ou seja, é quantizado utilizando um tamanho da etapa espectralmente constante no domínio ponderado ou domínio perceptualmente ponderado.

[0101] A figura 13 mostra um codificador 150 que compreende um transformador 152, um quantizador 154, um pré-enfatizador 156, um analisador de LPC 158 e um conversor de LPC em linha espectral 160. O pré-enfatizador 156 é opcional. O pré-enfatizador 156 submete o sinal de áudio de entrada 12 a uma pré-ênfase, a saber uma filtragem passa-alta com uma função de transferência do filtro passa-alta superficial utilizando, por exemplo, um filtro FIR ou IIR. Um filtro passa-alta de primeira ordem pode, por exemplo, ser utilizado para o pré-enfatizador 156 como $H(z) = 1 - \alpha z^{-1}$ com a definição de α , por exemplo, a quantidade ou resistência da pré-ênfase na linha com a qual, de acordo com uma das aplicações, a inclinação espectralmente global na qual o ruído para ser preenchido ao espectro é submetido, é variada. Uma possível definição de α poderia ser 0,68. A pré-ênfase causada pelo pré-enfatizador 156 é para mudar a energia dos valores espectrais quantizados transmitidos pelo codificador 150, de uma frequência alta para baixa, assim considerando as

leis psicoacústica de acordo com a percepção humana que é mais alta na região de baixa frequência do que na região de alta frequência. Se o sinal de áudio for ou não pré-enfatizado, o analisador de LPC 158 realiza uma análise LPC no sinal de áudio de entrada 12 para linearmente prever o sinal de áudio ou, para ser mais preciso, estimar o envelope espectral. O analisador de LPC 158 determina em unidades de tempo de, por exemplo, subestruturas que consistem em um número de amostras de áudio do sinal de áudio 12, os coeficientes da previsão linear e transmite os mesmos conforme mostrado em 162 no lado de decodificação dentro do fluxo de dados. O analisador de LPC 158 determina, por exemplo, os coeficientes da previsão linear utilizando a autocorrelação nas janelas de análise e utilizando, por exemplo, um algoritmo Levinson-Durbin. Os coeficientes da previsão linear podem ser transmitidos no fluxo de dados em uma versão quantizada e/ou transformada como na forma de pares de linha espectral ou semelhantes. Em qualquer caso, o analisador de LPC 158 encaminha ao conversor de LPC em linha espectral 160 os coeficientes da previsão linear como ainda disponível no lado de decodificação através do fluxo de dados e o conversor 160 converte os coeficientes da previsão linear em uma curva espectral utilizada pelo quantizador 154 para espectralmente variar/definir o tamanho da etapa de quantização. Em particular, o transformador 152 submete o sinal de áudio de entrada 12 a uma transformação como na mesma forma que o transformador 104 faz. Assim, o transformador 152 emite uma sequência de espectros e o

quantizador 154 pode, por exemplo, dividir cada espectro pela curva espectral obtida do conversor 160 utilizando, então, um tamanho da etapa de quantização espectralmente constante para todo o espectro. O espectrograma de uma sequência de espectros emitida pelo quantizador 154 é mostrado em 164 na figura 13 e compreende ainda algumas partes zero espectrais contínuas que podem ser preenchidas no lado de decodificação. Um parâmetro global do nível de ruído pode ser transmitido dentro do fluxo de dados pelo codificador 150.

[0102] A figura 14 mostra um decodificador encaixando ao codificador da figura 13. O decodificador da figura 14 é geralmente indicado utilizando o sinal de referência 170 e compreende um preenchedor de ruído 30, um conversor de LPC em linha espectral 172, um dequantizador 174 e um transformador inverso 176. O preenchedor de ruído 30 recebe os espectros quantizados 164, realiza o preenchimento de ruído nas partes zero espectrais contínuas conforme descrito acima, e então encaminha o espectrograma preenchido ao dequantizador 174. O dequantizador 174 recebe do conversor de LPC para linha espectral 172 uma curva espectral a ser utilizada pelo dequantizador 174 para formar novamente o espectro preenchido ou, em outras palavras, para dequantizá-lo. Este processo é às vezes chamado FDNS (*Formação do Ruído do Domínio de Frequência* | *Frequency Domain Noise Shaping*). O conversor de LPC em linha espectral 172 deriva a curva espectral com base na informação de LPC 162 no fluxo de dados. O espectro dequantizado, ou espectro reformado, emitido pelo dequantizador 174 é submetido a uma

transformação inversa pelo transformador inverso 176 a fim de recuperar o sinal de áudio. Novamente, a sequência de espectros reformados pode ser submetida pelo transformador inverso 176 a uma transformação inversa seguida por um processo de adição por sobreposição a fim de realizar o cancelamento de distorção de domínio de tempo entre as retransformações consecutivas no caso da transformação do transformador 152 ser uma transformada revestida criticamente amostrada como MDCT.

[0103] Em forma de linhas pontilhadas nas figuras 13 e 14, é mostrado que a pré-ênfase aplicada pelo pré-enfatizador 156 pode variar no tempo, com uma variação sendo sinalizada dentro do fluxo de dados. O preenchedor de ruído 30 pode, neste caso, considerar a pré-ênfase ao realizar o preenchimento de ruído conforme descrito acima com relação à figura 8. Em particular, a pré-ênfase causa uma inclinação espectral no espectro quantizado emitido pelo quantizador 154 em que os valores espectrais quantizados, ou seja, os níveis espectrais, tendem a reduzir das frequências mais baixas às frequências mais altas, ou seja, eles mostram uma inclinação espectral. Esta inclinação espectral pode ser compensada, ou melhor simulada ou adaptada, pelo preenchedor de ruído 30 na forma descrita acima. Se sinalizado no fluxo de dados, o grau de pré-ênfase sinalizado pode ser utilizado para realizar a inclinação adaptativa do ruído preenchido em uma forma dependente do grau de pré-ênfase. Isso é, o grau de pré-ênfase sinalizado no fluxo de dados pode ser utilizado pelo decodificador para definir o grau de inclinação espectral

imposto no ruído preenchido ao espectro pelo preenchedor de ruído 30.

[0104] Até agora, várias aplicações foram descritas e, a seguir, os exemplos de implementação específica são apresentados. Os detalhes trazidos com relação a estes exemplos, devem ser entendidos como sendo individualmente transferíveis nas aplicações acima para ainda especificar os mesmos. Antes disso, entretanto, deve-se observar que todas as aplicações descritas acima podem ser utilizadas tanto na codificação de áudio quanto na codificação de voz. Elas geralmente se referem à codificação de transformada e utilizam um conceito adaptativo do sinal para substituir os zeros introduzidos no processo de quantização pelo ruído espectralmente formado utilizando quantidade muito pequena de informação adicional. Nas aplicações descritas acima, a observação foi explicada que os furos espectrais às vezes aparecem logo abaixo de uma frequência inicial do preenchimento de ruído se qualquer frequência inicial for utilizada e que estes furos espectrais são às vezes perceptualmente desagradáveis. As aplicações acima utilizando uma sinalização explícita da frequência inicial permitem a remoção dos furos que trazem a degradação, mas permitem evitar a inserção do ruído em baixas frequências em qualquer lugar que a inserção de ruído introduziria as distorções.

[0105] Ainda, algumas das aplicações descritas acima utilizam um preenchimento do ruído controlado por pré-ênfase a fim de compensar a inclinação espectral causada pela pré-ênfase. Estas aplicações consideraram a observação que se o

filtro LPC fosse calculado em um sinal de pré-ênfase, meramente aplicando uma magnitude média ou global ou energia do ruído médio a ser inserido faria com que a formação do ruído introduzisse uma inclinação espectral no ruído inserido, pois a FDNS no lado de decodificação submeteria o ruído espectralmente plano inserido em uma formação espectral ainda mostrando a inclinação espectral da pré-ênfase. Certamente, as últimas aplicações realizaram o preenchimento de ruído em tal forma que a inclinação espectral da pré-ênfase é considerada e compensada.

[0106] Assim, em outras palavras, as figuras 11 e 14 mostraram um decodificador de áudio de transformada perceptual. Ele compreende um preenchedor de ruído 30 configurado para realizar o preenchimento de ruído em um espectro 18 de um sinal de áudio. O desempenho pode ser feito dependente da tonalidade conforme descrito acima. O desempenho pode ser feito pelo preenchimento do espectro com ruído exibindo uma inclinação espectralmente global para obter um espectro preenchido por ruído, conforme descrito acima. "Inclinação espectralmente global" deve, por exemplo, significar que a inclinação manifesta, por exemplo, em um envelope que envolve o ruído por todas as partes 40 a serem preenchidas com ruído, que é inclinado, ou seja, tem um declive não zero. "Envelope" é, por exemplo, definido ser uma curva de regressão espectral como uma função linear ou outro polinomial de ordem dois ou três, por exemplo, levando através da máxima local do ruído preenchido à parte 40 que é autocontínua, mas espectralmente distanciada. "Reduzir de

frequências baixas a altas" significa que esta inclinação tem um declive negativo, e "aumentar das frequências baixas a altas" significa que esta inclinação tem um declive positivo. Ambos os aspectos do desempenho podem aplicar simultaneamente ou meramente um deles.

[0107] Ainda, o decodificador de áudio de transformada perceptual compreende um formador de ruído de domínio de frequência 6 na forma de dequantizador 132, 174, configurado para submeter o espectro preenchido por ruído à formação espectral utilizando uma função de ponderação perceptual espectral. No caso da figura 11, o formador de ruído de domínio de frequência 132 é configurado para determinar a função de ponderação perceptual espectral da informação de coeficiente de previsão linear 162 sinalizada no fluxo de dados na qual o espectro é codificado. No caso da figura 14, o formador de ruído de domínio de frequência 174 é configurado para determinar a função de ponderação perceptual espectral dos fatores de escala 112 referentes às bandas do fator de escala 110, sinalizadas no fluxo de dados. Conforme descrito com referência à figura 8 e ilustrado com relação à figura 11, o preenchedor de ruído 34 pode ser configurado para variar um declive da inclinação espectralmente global receptivo a uma sinalização explícita no fluxo de dados, ou deduzir o mesmo de uma parte do fluxo de dados, que sinaliza a função de ponderação perceptual espectral como pela avaliação do envelope espectral LPC ou dos fatores de escala, ou pela dedução do mesmo a partir do espectro quantizado e transmitido 18.

[0108] Ainda, o decodificador de áudio de transformada perceptual compreende um transformador inverso 134, 176 configurado para transformar inversamente o espectro preenchido por ruído, espectralmente formado pelo formador de ruído de domínio de frequência, para obter uma transformada inversa, e submeter a transformada inversa a um processo de adição por sobreposição.

[0109] De forma correspondente, as figuras 13 e 9 mostraram exemplos para um codificador de áudio de transformada perceptual configurado para realizar uma ponderação do espectro 1 e quantização 2 ambas implementadas nos módulos do quantizador 108, 154 mostrados nas figuras 9 e 13. A ponderação do espectro 1 espectralmente pondera um espectro original do sinal de áudio de acordo com um inverso de uma função de ponderação perceptual espectral para obter um espectro perceptualmente ponderado e a quantização 2 quantiza o espectro perceptualmente ponderado em uma forma espectralmente uniforme para obter um espectro quantizado. O codificador de áudio de transformada perceptual ainda realiza um cálculo computacional do nível de ruído 3 dentro dos módulos de quantização 108, 154, por exemplo, calculando um parâmetro do nível de ruído pela medição de um nível do espectro perceptualmente ponderado colocalizado nas partes zero do espectro quantizado em uma forma ponderada com uma inclinação espectralmente global aumentando das frequências baixas a altas. De acordo com a figura 13, o codificador de áudio de transformada perceptual compreende um analisador LPC 158 configurado para determinar a informação de coeficiente

de previsão linear 162 representando um envelope do espectro original do sinal de áudio LPC, em que o ponderador espectral 154 é configurado para determinar a função de ponderação perceptual espectral para seguir o envelope espectral LPC. Conforme descrito, o analisador LPC 158 pode ser configurado para determinar a informação de coeficiente de previsão linear 162 pela realização da análise LP em uma versão do sinal de áudio, submeter a um filtro de pré-ênfase 156. Conforme descrito acima com relação à figura 13, o filtro de pré-ênfase 156 pode ser configurado para filtro passa-alta do sinal de áudio com uma quantidade de pré-ênfase variante para obter a versão do sinal de áudio, submeter a um filtro de pré-ênfase, em que o cálculo computacional do nível de ruído pode ser configurado para definir uma quantidade da inclinação espectralmente global dependendo da quantidade de pré-ênfase. Explicitamente a sinalização da quantidade da inclinação espectralmente global ou da quantidade de pré-ênfase no fluxo de dados pode ser utilizada. No caso da figura 9, o codificador de áudio de transformada perceptual compreende uma determinação do fator de escala, controlada através de um modelo perceptual 106, que determina os fatores de escala 112 referentes às bandas do fator de escala 110 para seguir um limite de mascaramento. Esta determinação é implementada no módulo de quantização 108, por exemplo, que ainda age como o ponderador espectral configurado para determinar a função de ponderação perceptual espectral para seguir os fatores de escala.

[0110] Todas as aplicações descritas acima têm em

comum os furos de espectro que são evitados e ainda a ocultação das linhas não quantizadas a zero tonais que é evitada. Na forma descrita acima, a energia em partes com ruído de um sinal pode ser preservada e a adição do ruído que mascarou os componentes tonais é evitada em uma forma descrita acima.

[0111] Nas implementações específicas descritas abaixo, a parte da informação adicional para realizar o preenchimento dependente da tonalidade de ruído não adiciona qualquer coisa à informação adicional do codec existente onde o preenchimento de ruído é utilizado. Toda a informação do fluxo de dados que é utilizada para a reconstrução do espectro, independentemente do ruído, também pode ser utilizada para a formação do preenchimento de ruído.

[0112] De acordo com um exemplo da implementação, o preenchimento de ruído no preenchedor de ruído 30 é realizado como segue. Todas as linhas espectrais acima de um índice inicial de preenchimento do ruído que são quantizadas a zero são substituídas por um valor não zero. Isso é feito, por exemplo, em uma forma aleatória ou pseudoaleatória com função de densidade da probabilidade espectralmente constante ou utilizando a interligação de outras localizações espectrais do espectrograma (fontes). Veja, por exemplo, a figura 15. A figura 15 mostra dois exemplos para um espectro a ser submetido a um preenchimento de ruído como o espectro 34 ou os espectros 18 no espectrograma 12 emitidos pelo quantizador 108 ou os espectros 164 emitidos pelo quantizador 154. O índice inicial de preenchimento do ruído é um índice da linha

espectral entre $iFreq0$ e $iFreq1$ ($0 < iFreq0 \leq iFreq1$), onde $iFreq0$ e $iFreq1$ são predeterminadas, índices da linha espectral dependentes da taxa de bit e largura de banda. O índice inicial de preenchimento do ruído é igual ao índice $iStart$ ($iFreq0 \leq iStart \leq iFreq1$) de uma linha espectral quantizada a um valor não zero, onde todas as linhas espectrais com índices j ($iStart < j \leq Freq1$) são quantizadas a zero. Diferentes valores para $iStart$, $iFreq0$ ou $iFreq1$ poderiam ainda ser transmitidos no fluxo de dados contínuos para permitir a inserção do ruído de baixa frequência em certos sinais (por exemplo, ruído ambiental).

[0113] O ruído inserido é formado nas seguintes etapas:

1. No domínio residual ou domínio ponderado.
A formação no domínio residual ou domínio ponderado foi extensivamente descrita acima com relação às figuras 1-14.

2. Formação espectral utilizando um LPC ou FDNS (formação no domínio de transformada utilizando a resposta de magnitude de LPC) foi descrita com relação às figuras 13 e 14. O espectro ainda pode ser formado utilizando fatores de escala (como em AAC) ou utilizando qualquer outro método de formação espectral para formação do espectro completo conforme descrito com relação às figuras 9-12.

3. Formação opcional utilizando TNS (Formação de Ruído Temporal) utilizando um número menor de bits, foi descrita brevemente com relação às figuras 9-12.

[0114] Apenas a informação adicional necessária para o preenchimento de ruído tem o nível, que é transmitido

utilizando 3 bits, por exemplo.

[0115] Ao utilizar FDNS não há necessidade de adaptar a um preenchimento de ruído específico e forma o ruído sobre o espectro completo utilizando número menor de bits do que dos fatores de escala.

[0116] Uma inclinação espectral pode ser introduzida no ruído inserido para neutralizar a inclinação espectral da pré-ênfase na formação de ruído perceptual com base em LPC. Visto que a pré-ênfase representa um filtro passa-alta suave aplicado ao sinal de entrada, a compensação da inclinação pode contrariar isso multiplicando o equivalente da função de transferência de um filtro passa-baixa sutil no espectro inserido no ruído. A inclinação espectral desta operação passa-baixa é dependente do fator de pré-ênfase e, preferivelmente, taxa de bit e largura de banda. Isso foi discutido com referência à figura 8.

[0117] Para cada furo espectral, constituído de 1 ou mais linhas espectrais quantizadas em zero consecutivas, o ruído inserido pode ser formado conforme descrito na figura 16. O preenchimento de nível de ruído pode ser encontrado no codificador e transmitido no fluxo contínuo de bits. Não há preenchimento de ruído nas linhas espectrais não quantizadas a zero e ele aumenta na área de transição até o total preenchimento de ruído. Na área do preenchimento de ruído total, o preenchimento de nível de ruído é igual ao nível transmitido no fluxo contínuo de bits, por exemplo. Isso evita inserir o alto nível de ruído na proximidade imediata das linhas espectrais não quantizadas a zero que poderia

potencialmente mascarar ou distorcer os componentes tonais. Entretanto, todas as linhas quantizadas a zero são substituídas por um ruído, deixando nenhum furo de espectro.

[0118] A largura de transição é dependente da tonalidade do sinal de entrada. A tonalidade é obtida para cada período de tempo. Nas figuras 17a-d a formação do preenchimento de ruído é, de forma exemplar, descrita para diferentes tamanhos do furo e larguras de transição.

[0119] A medição do espectro da tonalidade pode ser baseada na informação disponível no fluxo de dados contínuos:

- ganho LTP
- indicador de remodulação do espectro habilitado (veja [6])
- indicador de TNS

[0120] A largura de transição é proporcional à tonalidade - pequena para ruído como os sinais, grande para sinais muito tonais.

[0121] Em uma aplicação, a largura de transição é proporcional ao ganho de LTP se o ganho de LTP > 0 . Se o ganho de LTP for igual a 0 e a remodulação do espectro for habilitada, então a largura de transição para o ganho de LTP médio é utilizado. Se TNS for habilitada, então não há área de transição, mas o preenchimento de ruído total deveria ser aplicado a todas as linhas espectrais quantizadas em zero. Se o ganho de LTP for igual a 0 e a TNS e a remodulação do espectro são desabilitadas, uma largura mínima de transição é utilizada.

[0122] Se não há informação de tonalidade no fluxo

de dados contínuos uma medição de tonalidade pode ser calculada no sinal decodificado sem o preenchimento de ruído. Se não há informação de TNS, uma medição de nivelamento temporal pode ser calculada no sinal decodificado. Se, entretanto, a informação de TNS está disponível, tal medição de nivelamento pode ser derivada dos coeficientes do filtro TNS diretamente, por exemplo, pelo cálculo computacional do ganho de previsão do filtro.

[0123] No codificador, o preenchimento de nível de ruído pode ser calculado preferivelmente considerando a largura de transição. Várias formas para determinar o preenchimento de nível de ruído do espectro quantizado são possíveis. A mais simples é somar a energia (quadrado) de todas as linhas do espectro de entrada normalizado na região do preenchimento de ruído (ou seja, acima de i_{start}) que foram quantizadas a zero, então para dividir esta soma pelo número destas linhas para obter a energia média por linha, e para finalmente calcular um nível de ruído quantizado a partir da raiz quadrada da energia da linha média. Dessa forma, o nível de ruído é efetivamente derivado de RMS dos componentes espectrais quantizados a zero. Deixamos, por exemplo, A ser o conjunto de índices i das linhas espectrais onde o espectro foi quantizado a zero e que pertence a qualquer uma das partes zero, por exemplo, está acima da frequência inicial, e deixamos N denotar o fator de escala do ruído global. Os valores do espectro ainda como não quantizado devem ser denotados y_i . Adicionalmente, $\text{left}(i)$ deve ser uma função, indicando para qualquer valor espectral

quantizado a zero no índice i , o índice do valor quantizado a zero na extremidade de baixa frequência da parte zero a qual i pertence, e $F_i(j)$ com $j=0$ a $J_i - 1$ devem denotar a função atribuída a, dependendo da tonalidade, a parte zero começando no índice i , com J_i indicando a largura desta parte zero. Então, N pode ser determinado por $N = \sqrt{\sum_{i \in A} y_i^2 / \text{cardinality}(A)}$.

[0124] Na aplicação preferida, os tamanhos do furo individual bem como a largura de transição são considerados. Para esta finalidade, os testes das linhas quantizadas a zero consecutivas são agrupadas em regiões do furo. Cada linha espectral de entrada normalizada em uma região do furo, ou seja, cada valor espectral do sinal original em uma posição espectral dentro de qualquer parte zero espectral contínua, é então escalado pela função de transição, conforme descrito na seção anterior, e subsequentemente a soma das energias das linhas escaladas é calculada. Como na aplicação simples anterior, o preenchimento de nível de ruído pode então ser calculado de RMS das linhas quantizadas a zero. Aplicando a nomenclatura acima, N pode ser calculado como por $N = \sqrt{\sum_{i \in A} (F_{\text{left}(i)}(i - \text{left}(i)) \cdot y_i)^2 / \text{cardinality}(A)}$.

[0125] Um problema com esta abordagem, entretanto, é que a energia espectral em pequenas regiões do furo (ou seja, regiões com uma largura muito menor que duas vezes a largura de transição) é subestimada visto que no cálculo de RMS, o número de linhas espectrais na soma pela qual a soma da energia é dividida não é mudada. Em outras palavras, quando os espectros quantizados exhibe, na maioria das vezes, regiões

muito pequenas do furo, o preenchimento de nível de ruído resultante será menor do que quando o espectro é escasso e tem apenas poucas regiões longas do furo. Para garantir que em ambos os casos um nível de ruído semelhante é encontrado, é então vantajoso adaptar a contagem por linha utilizada no denominador do cálculo de RMS na largura de transição. De forma mais importante, se um tamanho da região do furo for menor do que duas vezes a largura de transição, o número de linhas espectrais nesta região do furo não é contado como é, ou seja, como um número inteiro de linhas, mas como um número por linha fracionária, que é menor do que o número por linha do número inteiro. Na formula acima referente a N , por exemplo, a "*cardinality(A)* [*cardinalidade(A)*]" seria substituída por um número menor dependendo do número de "pequenas" partes zero.

[0126] Além disso, a compensação da inclinação espectral no preenchimento de ruído devido à codificação perceptual com base em LPC deveria ainda ser considerada durante o cálculo do nível de ruído. Mais especificamente, o inverso da compensação da inclinação do preenchimento de ruído do lado do decodificador é preferivelmente aplicado nas linhas espectrais não quantizados originais que foram quantizados a zero, antes do nível de ruído ser calculado. No contexto da codificação com base em LPC que emprega a pré-ênfase, isso implica que as linhas de frequência mais alta são amplificadas levemente com relação às linhas de frequência mais baixa antes do nível de ruído estimativa. Aplicando a nomenclatura acima, N pode ser calculado como por

$$N = \text{sqrt}(\sum_{i \in A} (F_{\text{left}(i)}(i - \text{left}(i)) \cdot \text{LPF}(i)^{-1} \cdot y_i)^2 / \text{cardinality}(A)) .$$

Conforme mencionado acima, dependendo das circunstâncias, a função LPF que corresponde à função 15 pode ter um declive positivo e LPF mudado para ler HPF corretamente. É brevemente observado que em todas as fórmulas acima utilizando "LPF", definição F_{left} em uma função constante como para ser uma, revelariam uma forma de como aplicar o conceito para submeter o ruído a ser preenchido ao espectro 34 com uma inclinação espectralmente global sem o preenchimento do furo dependente da tonalidade.

[0127] Os possíveis cálculos de N podem ser realizados no codificador como, por exemplo, em 108 ou 154.

[0128] Finalmente, foi observado que quando harmônicas de um sinal estacionário muito tonal foram quantizados a zero, as linhas representando estas harmônicas levam a um nível de ruído relativamente alto ou instável (ou seja, variável no tempo). Esta perturbação pode ser reduzida utilizando no cálculo do nível de ruído a magnitude média das linhas quantizadas a zero ao invés de sua RMS. Enquanto esta abordagem alternativa nem sempre garante que a energia das linhas do ruído preenchido no decodificador reproduz a energia das linhas originais nas regiões do preenchimento de ruído, não garante que os picos espectrais nas regiões do preenchimento de ruído têm apenas a contribuição limitada ao nível de ruído geral, assim reduzindo o risco de sobre-estimativa do nível de ruído.

[0129] Finalmente, observa-se que um codificador pode ainda ser configurado para realizar o preenchimento de

ruído completamente a fim de mantê-lo em linha com o decodificador como, por exemplo, para análise por finalidades de síntese.

[0130] Assim, a aplicação acima, *inter alias*, descreve um método adaptativo do sinal para substituir os zeros introduzidos no processo de quantização pelo ruído espectralmente formado. Uma extensão do preenchimento de ruído para um codificador e um decodificador é descrito atendendo as exigências precisamente mencionadas pela implementação do seguinte:

- Preenchimento do índice inicial de ruído pode ser adaptado ao resultado da quantização do espectro, mas limitado a uma determinada faixa

- Uma inclinação espectral pode ser introduzida ao ruído inserido para neutralizar a inclinação espectral da formação de ruído perceptual

- Todas as linhas quantizadas a zero acima do índice inicial de preenchimento do ruído são substituídas pelo ruído

- Por meios de uma função de transição, o ruído inserido é atenuado próximo às linhas espectrais não quantizadas a zero

- A função de transição é dependente das características instantâneas do sinal de entrada

- A adaptação do índice inicial de preenchimento do ruído, a inclinação espectral e a função de transição podem ser baseadas na informação disponível no decodificador

[0131] Não há necessidade de informação adicional, exceto para um preenchimento de nível de ruído.

[0132] Embora alguns aspectos tenham sido descritos no contexto de um aparelho, é evidente que estes aspectos também representam uma descrição do método correspondente, onde um bloco ou dispositivo corresponde a uma etapa do método ou uma característica de uma etapa do método. De forma análoga, os aspectos descritos no contexto de uma etapa do método também representam uma descrição de um bloco ou item ou característica correspondente de um aparelho correspondente. Algumas ou todas as etapas do método podem ser executadas por (ou utilizando) um aparelho de hardware, como, por exemplo, um microprocessador, um computador programável ou um circuito eletrônico. Em algumas aplicações, uma ou mais etapa(s) mais importante(s) do método pode(m) ser executada(s) por este aparelho.

[0133] Dependendo de certas exigências da implementação, as aplicações da invenção podem ser implementadas em hardware ou em software. A implementação pode ser realizada utilizando um meio de armazenamento digital, por exemplo, um disquete, um DVD, um Blu-Ray, um CD, uma memória ROM, uma PROM, uma EPROM, uma EEPROM ou uma memória flash, tendo sinais de controle eletronicamente legíveis armazenados nela, que cooperam (ou podem cooperar) com um sistema de computador programável de modo que o respectivo método seja realizado. Assim, o meio de armazenamento digital pode ser legível por computador.

[0134] Algumas aplicações de acordo com a invenção

compreendem um transportador de dados tendo sinais de controle eletronicamente legíveis que são capazes de cooperar com um sistema de computador programável, de modo que um dos métodos descritos aqui seja realizado.

[0135] Geralmente, as aplicações da presente invenção podem ser implementadas como um produto do programa de computador com um código de programa, o código de programa sendo operativo para realizar um dos métodos quando o produto do programa de computador é executado em um computador. O código de programa pode, por exemplo, ser armazenado em um transportador legível por máquina.

[0136] Outras aplicações compreendem o programa de computador para realizar um dos métodos descritos aqui, armazenados em um transportador legível por máquina.

[0137] Em outras palavras, uma aplicação do método inventivo é, portanto, um programa de computador tendo um código de programa para realizar um dos métodos descritos aqui, quando o programa de computador é executado em um computador.

[0138] Outra aplicação dos métodos inventivos é, portanto, um transportador de dados (ou um meio de armazenamento digital, ou um meio legível por computador) compreendendo, gravado nele, o programa de computador para realizar um dos métodos descritos aqui. O transportador de dados, o meio de armazenamento digital ou o meio gravado são tipicamente tangíveis e/ou não transitório.

[0139] Outra aplicação do método inventivo é, portanto, um fluxo de dados ou uma sequência de sinais que

representa o programa de computador para realizar um dos métodos descritos aqui. O fluxo de dados ou a sequência de sinais pode, por exemplo, ser configurado para ser transferido através de uma conexão de comunicação de dados, por exemplo, através da Internet.

[0140] Outra aplicação compreende um meio de processamento, por exemplo, um computador, ou um dispositivo de lógica programável, configurado para ou adaptado para realizar um dos métodos descritos aqui.

[0141] Outra aplicação compreende um computador tendo nele o programa de computador instalado para realizar um dos métodos descritos aqui.

[0142] Outra aplicação de acordo com a invenção compreende um aparelho ou um sistema configurado para transferir (por exemplo, eletrônica ou opticamente) um programa de computador para realizar um dos métodos descritos aqui a um receptor. O receptor pode, por exemplo, ser um computador, um dispositivo móvel, um dispositivo de memória ou semelhante. O aparelho ou sistema pode, por exemplo, compreender um servidor de arquivo para transferir o programa de computador ao receptor.

[0143] Em algumas aplicações, um dispositivo de lógica programável (por exemplo, um arranjo de portas de campo programáveis) pode ser utilizado para realizar algumas ou todas as funcionalidades dos métodos descritos aqui. Em algumas aplicações, uma matriz de campo de portas programáveis pode cooperar com um microprocessador a fim de realizar um dos métodos descritos aqui. Geralmente, os

métodos são preferivelmente realizados por qualquer aparelho de hardware.

[0144] O aparelho descrito aqui pode ser implementado utilizando um aparelho de hardware, ou utilizando um computador, ou utilizando uma combinação de um aparelho de hardware e um computador.

[0145] Os métodos descritos aqui podem ser realizados utilizando um aparelho de hardware, ou utilizando um computador, ou utilizando uma combinação de um aparelho de hardware e um computador.

[0146] As aplicações descritas acima são meramente ilustrativas para os princípios da presente invenção. Entende-se que as modificações e variações das disposições e os detalhes descritos aqui serão evidentes a outros especialistas na técnica. É intenção da invenção, portanto, ser limitada apenas pelo escopo das reivindicações da patente anexas e não pelos detalhes específicos apresentados em forma de descrição e explicação das aplicações no presente documento.

REFERÊNCIAS

[0147] [1] B. G. G. F. S. G. M. M. H. P. J. H. S. W. G. S. J. H. Nikolaus Rettelbach, "Noise Filler, Noise Filling Parameter Calculator Encoded Audio Signal Representation, Methods and Computer Program". Patent US 2011/0173012 A1.

[0148] [2] *Extended Adaptive Multi-Rate-Wideband (AMR-WB+) codec*, 3GPP TS 26.290 V6.3.0, 2005-2006.

[0149] [3] B. G. G. F. S. G. M. M. H. P. J. H. S. W. G. S. J. H. Nikolaus Rettelbach, "Audio encoder, audio

decoder, methods for encoding and decoding an audio signal, audio stream and computer program". Patent WO 2010/003556 A1.

[0150] [4] M. M. N. R. G. F. J. R. J. L. S. W. S. B. S. D. C. H. R. L. P. G. B. B. J. L. K. K. H. Max Neuendorf, "MPEG Unified Speech and Audio Coding - The ISO/MPEG Standard for High-Efficiency Audio Coding of all Content Types," in *132nd Conversion AES*, Budapest, 2012. Also appears in the *Journal of the AES*, vol. 61, 2013.

[0151] [5] M. M. M. N. a. R. G. Guillaume Fuchs, "MDCT-Based Coder for Highly Adaptive Speech and Audio Coding," in *17th European Signal Processing Conference (EUSIPCO 2009)*, Glasgow, 2009.

[0152] [6] H. Y. K. Y. M. T. Harada Noboru, "Coding Mmethod, Decoding Method, Coding Device, Decoding Device, Program, and Recording Medium". Patent WO 2012/046685 A1

REIVINDICAÇÕES

1. Aparelho configurado para realizar preenchimento de ruído em um espectro (34) de um sinal de áudio de uma maneira dependente de uma tonalidade do sinal de áudio, em que o aparelho é configurado para desquantizar (132; 174) o espectro (34), conforme derivado após o preenchimento de ruído, usando um tamanho de etapa de quantização adaptativo de sinal e espectralmente variável controlado por meio de um envelope espectral de predição linear sinalizado por meio de coeficientes de predição linear (162) em um fluxo de dados no qual o espectro (34) é codificado (164), ou fatores de escala (112) em relação às bandas do fator de escala (110), sinalizadas no fluxo de dados em que o espectro (34) é codificado, caracterizado pelo fato de que o aparelho é configurado para preencher uma porção zero espectral contígua (40) do espectro do sinal de áudio (34) com ruído espectralmente moldado usando uma função (48, 50) assumindo um máximo em um interior (52) da porção zero espectral contígua (40), e tendo bordas descendentes para fora (58, 60), uma inclinação absoluta da qual depende negativamente da tonalidade, ou uma função (48, 50) assumindo um máximo em um interior (52) da porção zero espectral contígua (40) e tendo bordas descendentes para fora (58, 60) uma largura espectral (54, 56) da qual depende positivamente de a tonalidade, ou uma função constante ou unimodal (48, 50) uma integral da qual - normalizada para uma integral de 1 - ao longo dos

quartos externos (a, d) da porção zero espectral contígua (40) depende negativamente da tonalidade, ou um conjunto de funções (80) dependente da largura da porção zero espectral contígua de modo que a função seja confinada à porção zero espectral contígua e dependente da tonalidade do sinal de áudio de modo que, se a tonalidade do sinal de áudio aumentar, a função fica mais compacta no interior da porção zero espectral contígua e distanciada das bordas externas da porção zero espectral contígua.

2. Aparelho, de acordo com a reivindicação 1, caracterizado pelo fato de que o aparelho é configurado para dimensionar o ruído com o qual as porções zero espectrais contíguas são preenchidas usando um nível de ruído global escalar sinalizado no fluxo de dados no qual o espectro é codificado de uma maneira espectralmente global.

3. Aparelho de acordo com a reivindicação 1 ou 2, em que o aparelho está configurado para gerar o ruído com o qual as porções zero espectrais contíguas são preenchidas, usando um processo aleatório ou pseudo-aleatório ou usando patching.

4. Aparelho de acordo com qualquer uma das reivindicações 1 a 3, em que o aparelho é configurado para derivar a tonalidade de um parâmetro de codificação codificado dentro do fluxo de dados.

5. Aparelho, de acordo com a reivindicação 4, caracterizado pelo fato de que o aparelho é configurado de modo que o parâmetro de codificação seja um sinalizador de

ativação LTP (predição de longo prazo) ou TNS (modelagem de ruído temporal) ou ganho e / ou um sinalizador de ativação de rearranjo de espectro, o sinalizador de ativação de rearranjo espectral sinalizando uma opção de codificação de acordo com a qual valores espectrais quantizados são reorganizados espectralmente com a transmissão adicional dentro do fluxo de dados da prescrição de rearranjo.

6. Aparelho, de acordo com qualquer uma das reivindicações anteriores, caracterizado pelo fato de que o aparelho é configurado para confinar o desempenho do preenchimento de ruído em uma porção espectral de alta frequência do espectro do sinal de áudio.

7. Aparelho, de acordo com qualquer uma das reivindicações anteriores, caracterizado pelo fato de que o aparelho é configurado para definir uma posição inicial de baixa frequência da porção espectral de alta frequência correspondente a uma sinalização explícita no fluxo de dados.

8. Aparelho, de acordo com qualquer uma das reivindicações anteriores, caracterizado pelo fato de que o aparelho é configurado para, ao realizar o preenchimento de ruído, preencher porções zero espectrais contíguas (40) do espectro (34) com ruído, cujo nível exibe uma diminuição de baixo para alto frequências, aproximando a função de transferência de um filtro passa-baixa espectral de modo a neutralizar uma inclinação espectral causada por uma pré-ênfase usada para codificar o espectro do sinal de áudio.

9. Aparelho, de acordo com a reivindicação 8, caracterizado pelo fato de que o aparelho é configurado para adaptar uma inclinação da diminuição a um fator de pré-ênfase da pré-ênfase.

10. Aparelho, de acordo com qualquer uma das reivindicações anteriores, caracterizado pelo fato de que o aparelho é configurado para identificar porções zero espectrais contíguas do espectro do sinal de áudio e preencher as porções zero espectrais contíguas com funções definidas dependentes de uma largura de porção zero spectral contígua respectiva de modo que a função é confinada à respectiva porção contígua spectral zero e dependente da tonalidade do sinal de áudio de modo que, se a tonalidade do sinal de áudio aumentar, a função se torne cada vez mais compacta no interior do respectivo zero spectral contíguo porção e distanciada das respectivas bordas da porção zero spectral contígua e, adicionalmente, dependente da respectiva porção zero spectral contígua 's posição spectral de forma que a escala da função dependa da respectiva posição spectral da porção zero spectral contígua.

11. Decodificador de áudio que suporta preenchimento de ruído, compreendendo um aparelho de acordo com qualquer uma das reivindicações anteriores.

12. Decodificador de áudio de transformação perceptual compreendendo um aparelho configurado para realizar preenchimento de ruído em um espectro (34) de um sinal de áudio de acordo com qualquer uma das reivindicações 1 a

10; e um modelador de ruído de domínio de frequência configurado para sujeitar o espectro preenchido com ruído à formação espectral usando uma função de ponderação perceptual espectral.

13. Codificador de áudio que suporta preenchimento de ruído compreendendo um aparelho de acordo com qualquer uma das reivindicações 1 a 10, o codificador sendo configurado para usar um espectro preenchido com ruído pelo aparelho, para análise por síntese.

14. Método que compreende realizar preenchimento de ruído em um espectro (34) de um sinal de áudio de uma maneira dependente de uma tonalidade do sinal de áudio, em que o método compreende desquantizar (132; 174) o espectro (34), conforme derivado após o preenchimento de ruído , usando um tamanho de etapa de quantização adaptável a sinal e espectralmente variável controlado por meio de um envelope espectral de predição linear sinalizado por meio de coeficientes de predição linear (162) em um fluxo de dados no qual o espectro (34) é codificado (164), ou fatores de escala (112) em relação às bandas do fator de escala (110), sinalizadas no fluxo de dados no qual o espectro (34) é codificado,

caracterizado pelo fato de que o método compreende o preenchimento de uma porção zero espectral contígua (40) do espectro do sinal de áudio (34) com ruído em forma espectral usando

uma função (48, 50) assumindo um máximo em um interior (52) da porção zero espectral contígua (40), e tendo

bordas descendentes para fora (58, 60), uma inclinação absoluta da qual depende negativamente da tonalidade, ou uma função (48, 50) assumindo um máximo em um interior (52) da porção zero espectral contígua (40) e tendo bordas descendentes para fora (58, 60) uma largura espectral (54, 56) da qual depende positivamente de a tonalidade, ou uma função constante ou unimodal (48, 50) uma integral da qual - normalizada para uma integral de 1 - sobre os quartos externos (a, d) da porção zero espectral contígua (40) depende negativamente da tonalidade ou um conjunto de funções (80) dependente da largura da porção zero espectral contígua de modo que a função seja confinada à porção zero espectral contígua e dependente da tonalidade do sinal de áudio de modo que, se a tonalidade do sinal de áudio aumentar, a função fica mais compacta no interior da porção zero espectral contígua e distanciada das bordas externas da porção zero espectral contígua.

15. Codificador configurado para executar o ruído de enchimento

em um espectro (34) de um sinal de áudio através do enchimento do espectro com ruído, de modo a obter um espectro de ruído preenchido

por

geração de um sinal de ruído intermediário;

identificar porções zero espectrais contíguas do espectro do sinal de áudio; determinar uma função para cada porção zero espectral contígua dependendo de

a largura da respectiva porção zero espectral contígua, de

modo que a função seja confinada à respectiva porção zero espectral contígua,

a respectiva posição espectral da porção zero espectral contígua de modo que uma escala da função dependa da posição espectral da porção zero espectral contígua respectiva de modo que uma quantidade da escala aumenta ou diminui monotonicamente com o aumento da frequência do espectro da porção zero espectral contígua respectiva posição; e

modelar espectralmente, para cada porção zero espectral contígua, o sinal de ruído intermediário usando a função determinada para a respectiva porção zero espectral contígua de modo que o ruído exiba uma inclinação espectral global com uma inclinação negativa.

16. Codificador de acordo com a reivindicação 1, em que o codificador é um codificador de áudio de transformação perceptual que compreende

um filtro de pré-ênfase;

um analisador LPC configurado para determinar a informação do coeficiente de predição linear (162) realizando a análise LP em uma versão do sinal de áudio, sujeito ao filtro de pré-ênfase, a informação do coeficiente de predição linear (162) representando um envelope espectral LPC de um espectro de a versão pré-enfatizada do sinal de áudio;

um transformador configurado para fornecer um espectro original do sinal de áudio;

um pesador de espectro configurado para ponderar

espectralmente o espectro original de um sinal de áudio de acordo com um inverso de uma função de ponderação perceptual espectral de modo a obter um espectro ponderado perceptualmente, em que o pesador espectral é configurado para determinar a função de ponderação perceptual espectral de modo a seguir o LPC envelope espectral;

um quantizador configurado para quantizar o espectro perceptualmente ponderado de uma maneira igual para linhas espectrais do espectro perceptualmente ponderado, de modo a obter um espectro quantizado, em que o codificador é configurado para codificar o espectro quantizado em um fluxo de dados para ser enviado para uma transformação perceptual decodificador de áudio de acordo com qualquer uma das reivindicações anteriores, a informação do coeficiente de predição linear também sendo sinalizada no fluxo de dados;

um computador de nível de ruído configurado para calcular um parâmetro de nível de ruído medindo um nível do espectro perceptualmente ponderado co-localizado a porções espectrais zero contíguas do espectro quantizado de uma maneira ponderada com uma inclinação espectralmente global tendo uma inclinação positiva.

17. Codificador de áudio de transformação perceptual, de acordo com a reivindicação 2, caracterizado pelo fato de que o filtro de pré-ênfase é configurado para filtrar o sinal de áudio com uma quantidade de pré-ênfase variável, de modo a obter a versão do sinal de áudio, sujeito a um filtro de pré-ênfase , em que o computador de nível de

ruído é configurado para definir uma inclinação da inclinação espectralmente global dependendo da quantidade de pré-ênfase.

18. Codificador de áudio de transformação perceptual de acordo com a reivindicação 3, configurado para codificar explicitamente a quantidade da inclinação espectralmente global ou a quantidade de pré-ênfase no fluxo de dados no qual o espectro quantizado (34) é codificado (164).

19. Codificador de áudio de transformação perceptual, de acordo com a reivindicação 4, caracterizado pelo fato de que compreende um determinante de fator de escala configurado para, controlado por meio de um modelo perceptivo, determinar fatores de escala (112) relacionados às bandas de fator de escala (110) de modo a seguir um limite de mascaramento, em que o pesador espectral configurado para determinar a função de ponderação perceptual espectral de modo a seguir os fatores de escala.

20. Codificador de áudio de transformação perceptual, de acordo com a reivindicação 2, caracterizado pelo fato de que o computador de nível de ruído é ainda configurado para determinar, para cada porção zero espectral contígua, a função (48, 50) de modo que a mesma assume um máximo em um interno (52) do contíguo porção zero espectral (40), e tem bordas descendentes externamente (58, 60) uma inclinação absoluta da qual depende negativamente da tonalidade, o mesmo assume um máximo em uma parte interna (52) da porção

zero espectral contígua (40), e tem bordas descendentes externamente (58, 60) uma largura espectral (54, 56) da qual depende positivamente da tonalidade e / ou mesmo é uma função constante ou unimodal (48, 50) uma integral da qual - normalizada para uma integral de 1 - sobre os quartos externos (a, d) da porção zero espectral contígua (40) depende negativamente da tonalidade.

21. Codificador de áudio de transformação perceptual, de acordo com a reivindicação 6, caracterizado pelo fato de que o computador de nível de ruído é configurado para deduzir a tonalidade de um sinalizador de ativação LTP (predição de longo prazo) ou TNS (modelagem de ruído temporal) ou ganho e / ou um sinalizador de ativação de rearranjo de espectro usado por o codificador de áudio de transformação perceptual para codificar o sinal de áudio, o sinalizador de ativação de rearranjo espectral sinalizando uma opção de codificação de acordo com a qual valores espectrais quantizados são reorganizados espectralmente com transmissão adicional dentro do fluxo de dados da prescrição de rearranjo.

22. Codificador de áudio de transformação perceptual, de acordo com qualquer uma das reivindicações 2 a 7, caracterizado pelo fato de que o preenchimento de ruído é configurado para confinar o preenchimento de ruído em uma porção espectral de alta frequência do espectro do sinal de áudio.

23. Codificador de áudio de transformação perceptual, de acordo com qualquer uma das reivindicações 2 a 8,

caracterizado pelo fato de que o computador de nível de ruído é configurado para restringir a medição a uma porção espectral de alta frequência com sinalização explícita definida uma posição inicial de baixa frequência da mesma em um fluxo de dados no qual o sinal de áudio está codificado.

24. Método para codificação compreendendo realizar preenchimento de ruído

em um espectro (34) de um sinal de áudio, preenchendo o espectro com ruído de modo a obter um espectro preenchido com ruído

gerando um sinal de ruído intermediário;

identificar porções zero espectrais contíguas do espectro do sinal de áudio; determinar uma função para cada porção zero espectral contígua dependendo de

a largura da respectiva porção zero espectral contígua, de modo que a função seja confinada à respectiva porção zero espectral contígua,

a respectiva posição espectral da porção zero espectral contígua de modo que uma escala da função dependa da posição espectral da porção zero espectral contígua respectiva de modo que uma quantidade da escala aumenta ou diminui monotonicamente com o aumento da frequência do espectro da porção zero espectral contígua respectiva posição; e

modelar espectralmente, para cada porção zero espectral contígua, o sinal de ruído intermediário usando a função determinada para a respectiva porção zero espectral

contígua de modo que o ruído exiba uma inclinação espectral global com uma inclinação negativa.

25. Método, de acordo com a reivindicação 10, caracterizado pelo fato de que o método usa codificação de áudio de transformação perceptual e compreende

determinar a informação do coeficiente de predição linear (162) realizando análise LP em uma versão do sinal de áudio, sujeito a um filtro de pré-ênfase, a informação do coeficiente de predição linear (162) representando um envelope espectral LPC de um espectro da versão pré-enfatizada do sinal de áudio;

fornecer um espectro original do sinal de áudio por um transformador;

ponderar espectralmente o espectro original do sinal de áudio de acordo com um inverso de uma função de ponderação perceptual espectral de modo a obter um espectro ponderado perceptualmente, em que a função de ponderação espectral é determinada de modo a seguir o envelope espectral LPC;

quantificar o espectro perceptualmente ponderado de uma maneira igual para linhas espectrais do espectro perceptualmente ponderado de modo a obter um espectro quantizado, em que o espectro quantizado é codificado em um fluxo de dados para ser enviado a um decodificador de áudio de transformação perceptual de acordo com qualquer uma das reivindicações 1 a 14, a informação do coeficiente de predição linear também sendo sinalizada no fluxo de dados;

computar um parâmetro de nível de ruído medindo um nível

do espectro perceptualmente ponderado co-localizado a porções espectrais zero contíguas do espectro quantizado de uma maneira ponderada com uma inclinação espectralmente global tendo uma inclinação positiva.

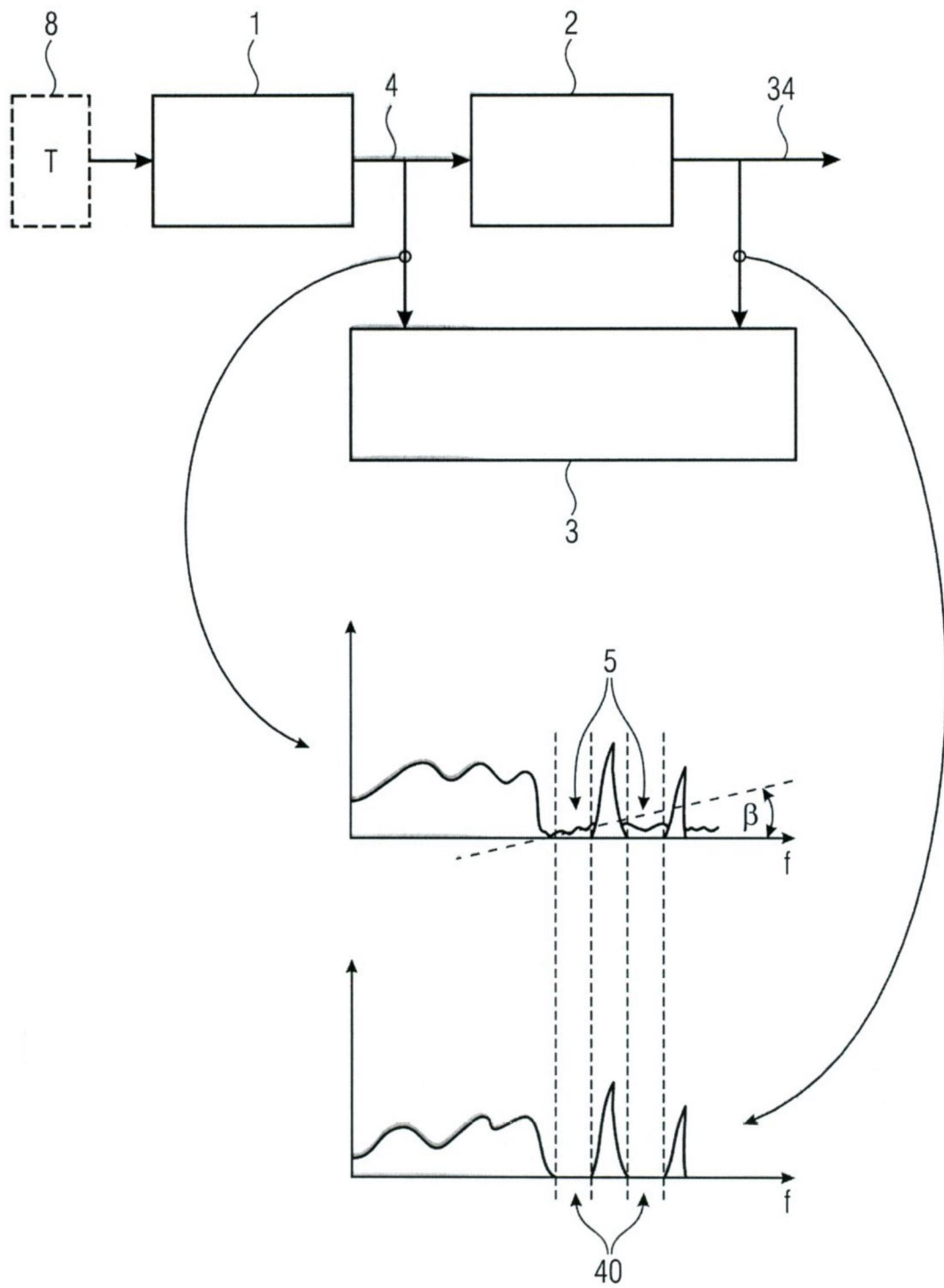


FIG 1A

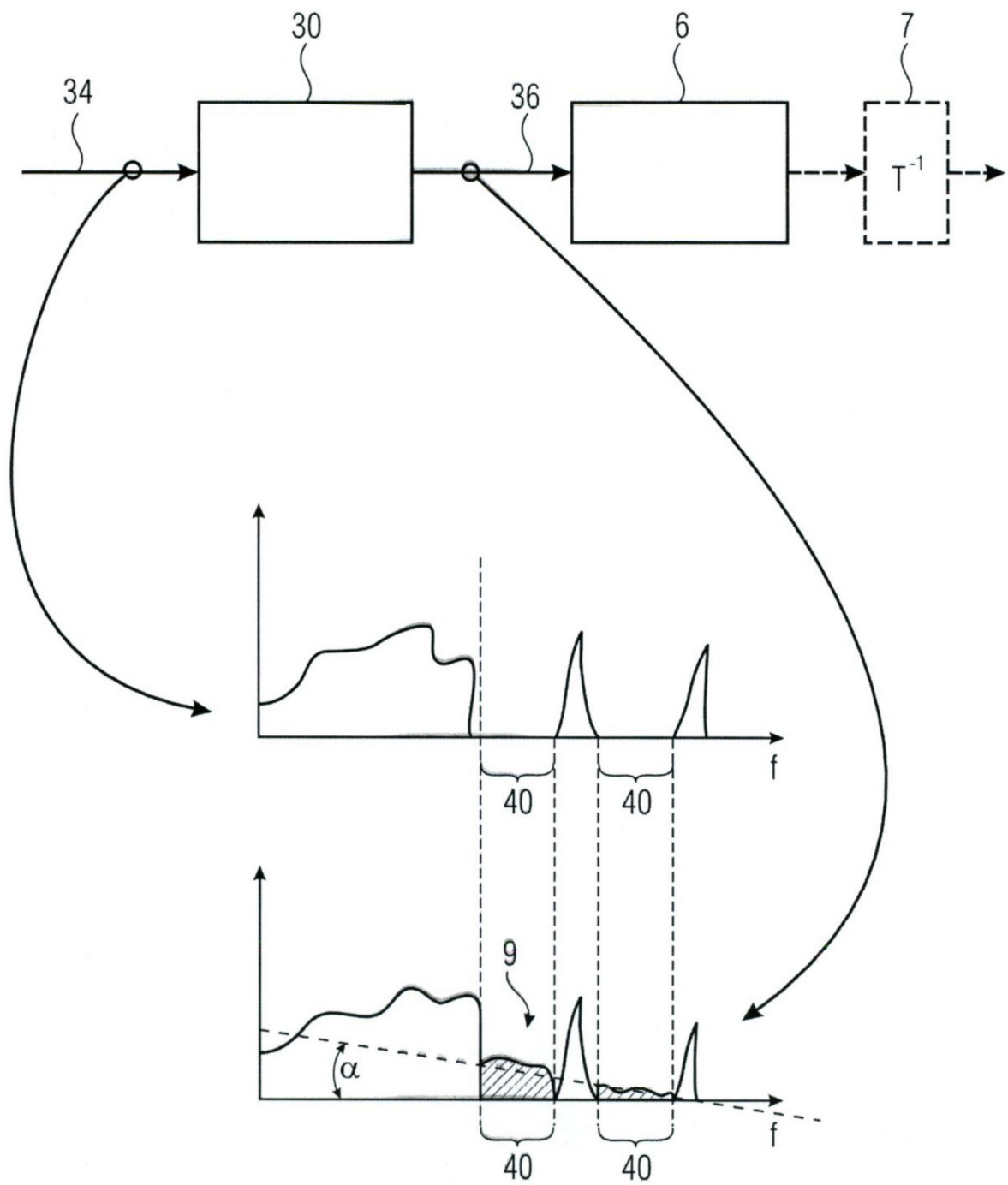
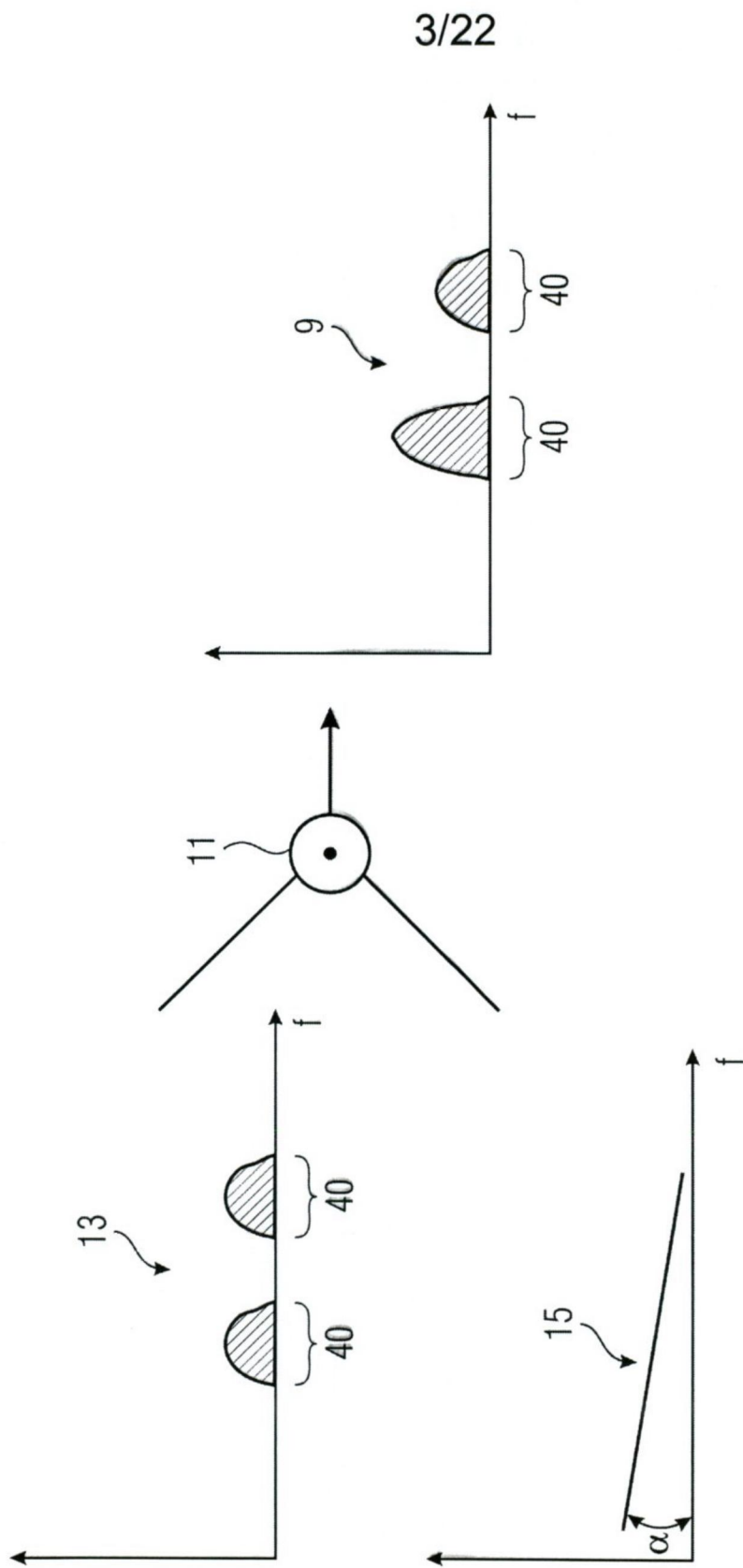


FIG 1B



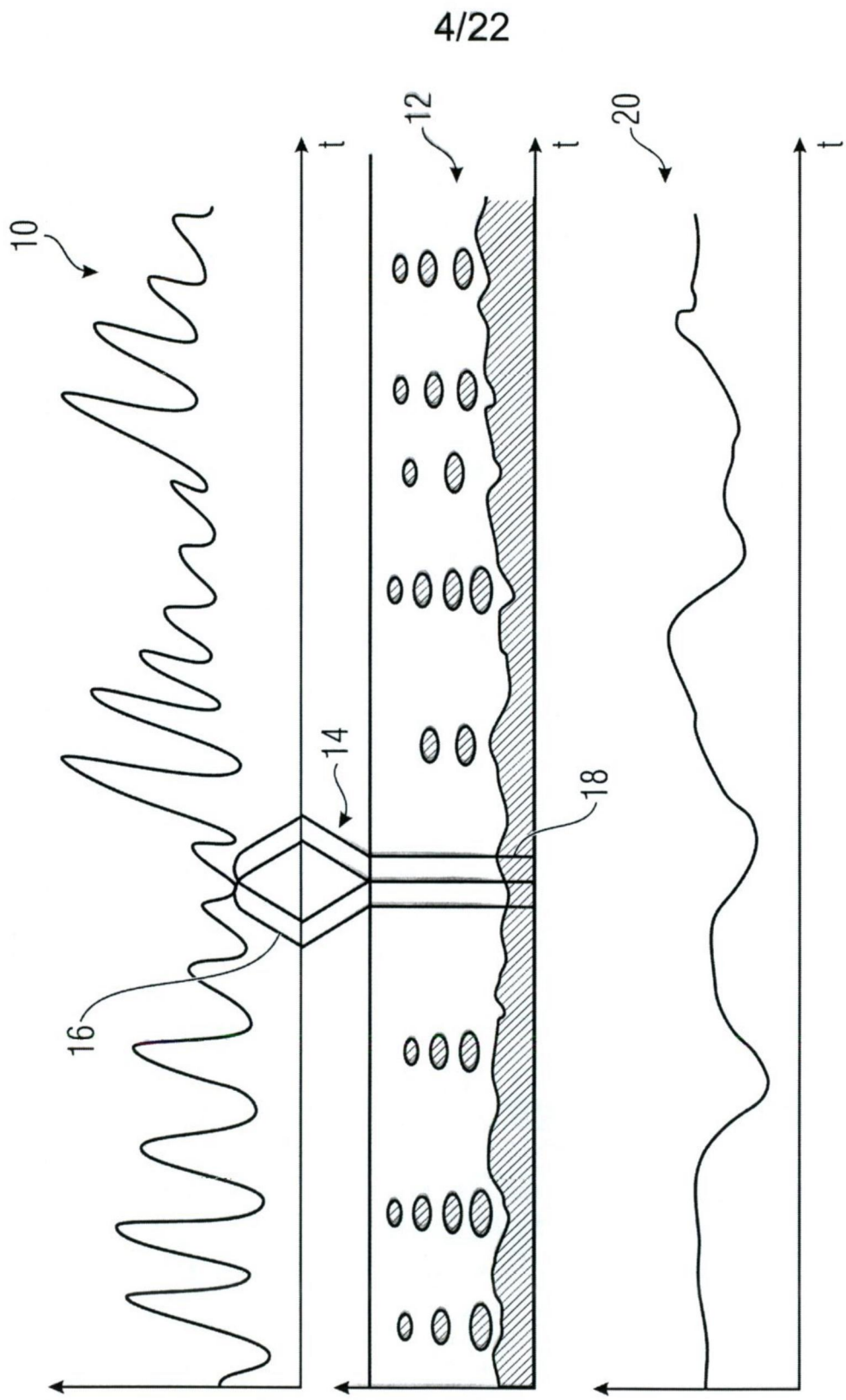


FIG 2A

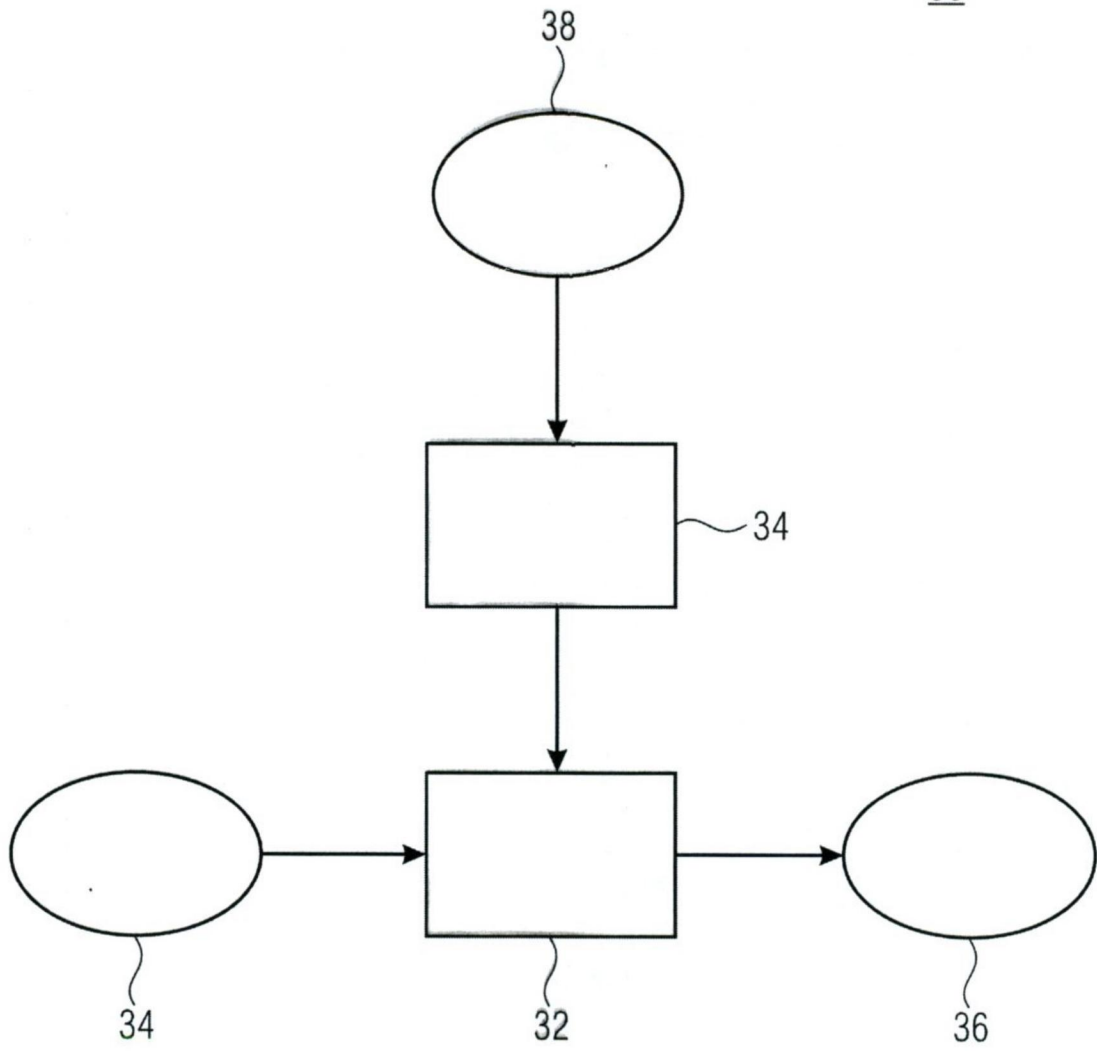


FIG 2B

6/22

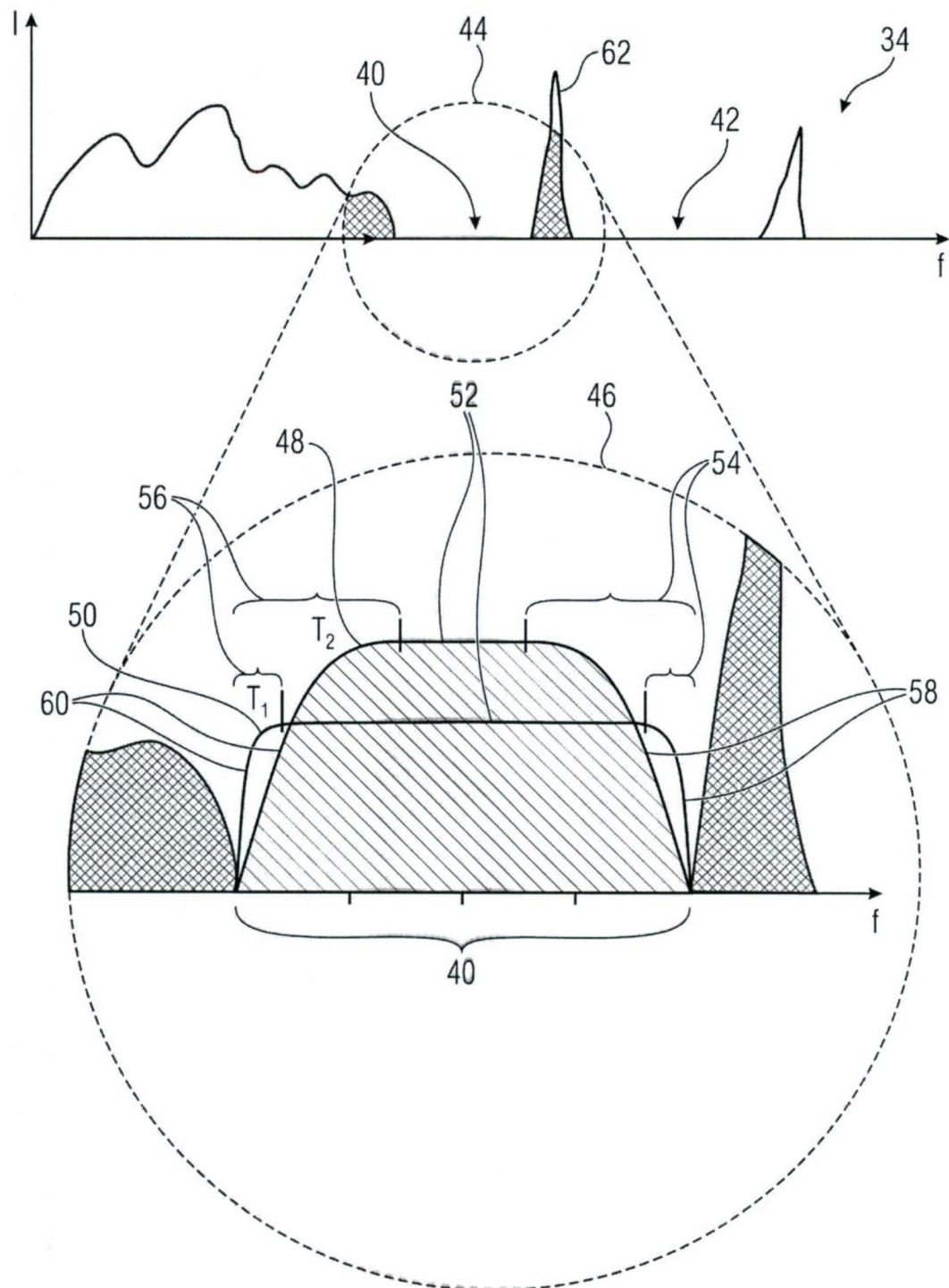


FIG 3

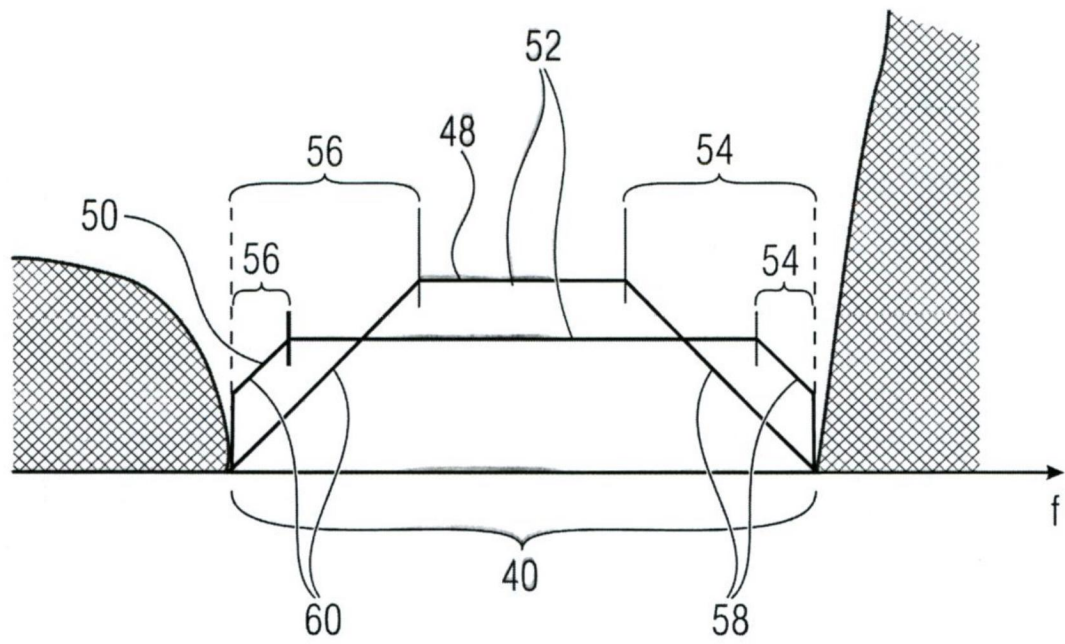


FIG 4

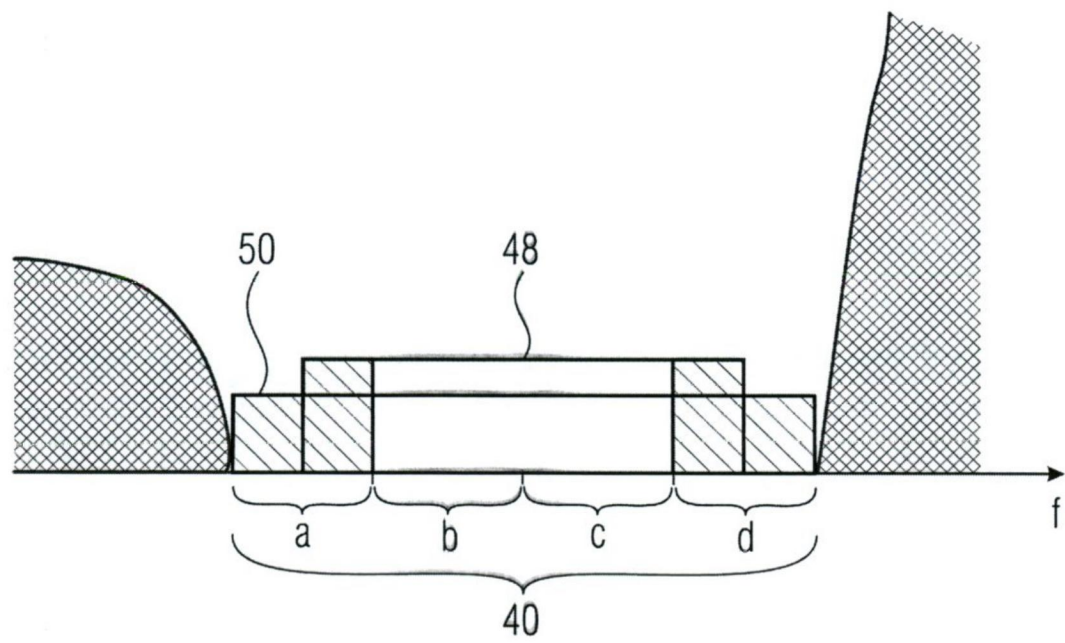


FIG 5

8/22

32

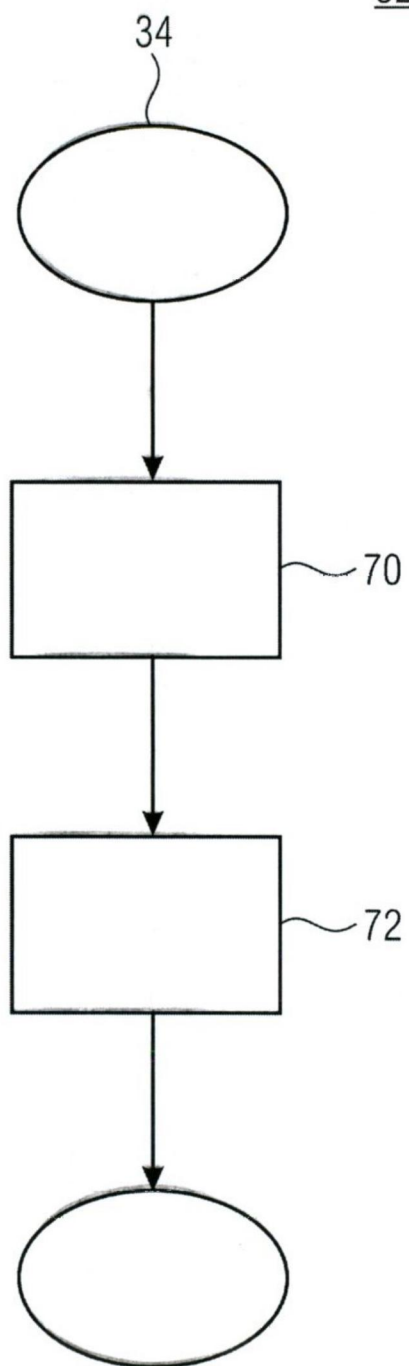


FIG 6

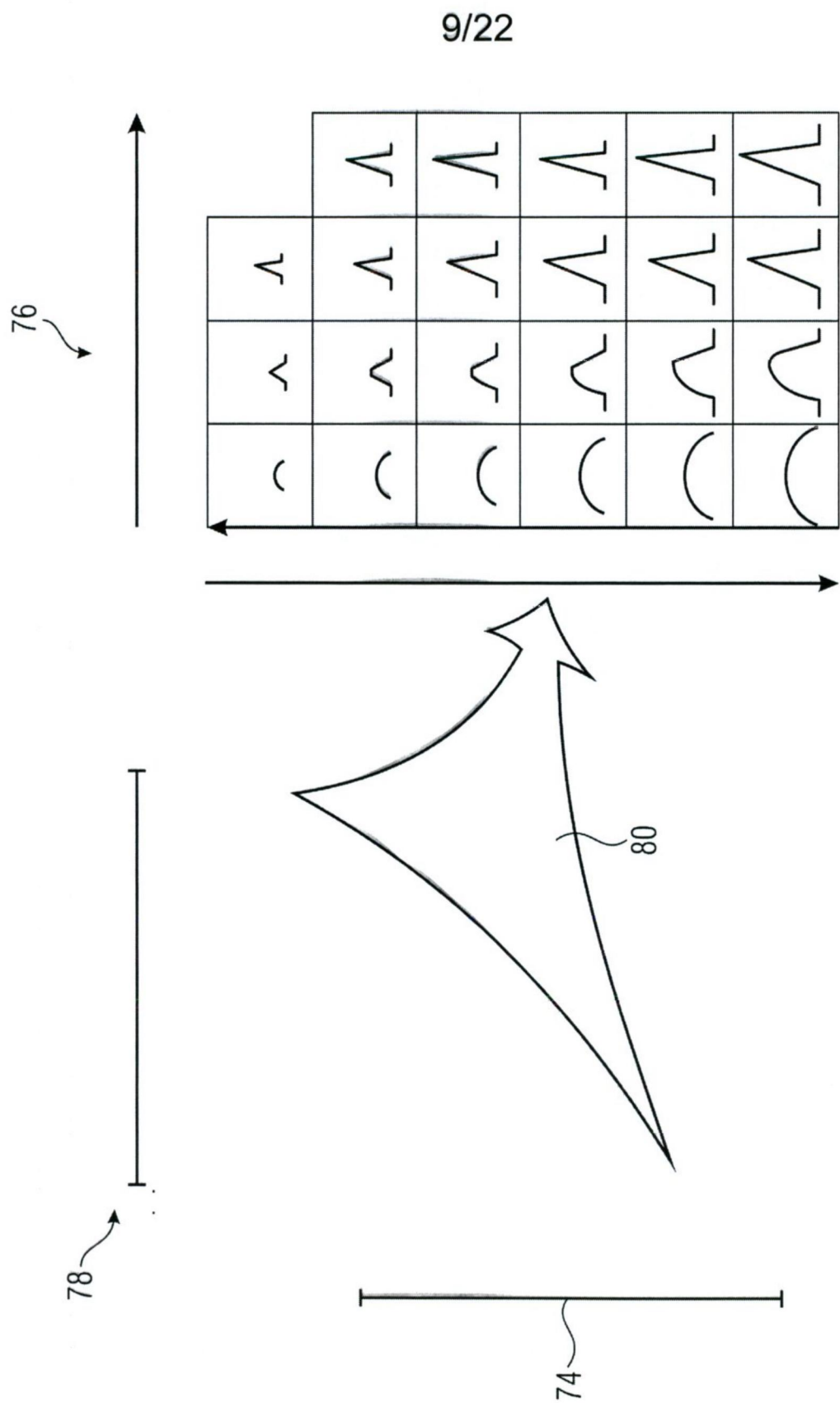


FIG 7

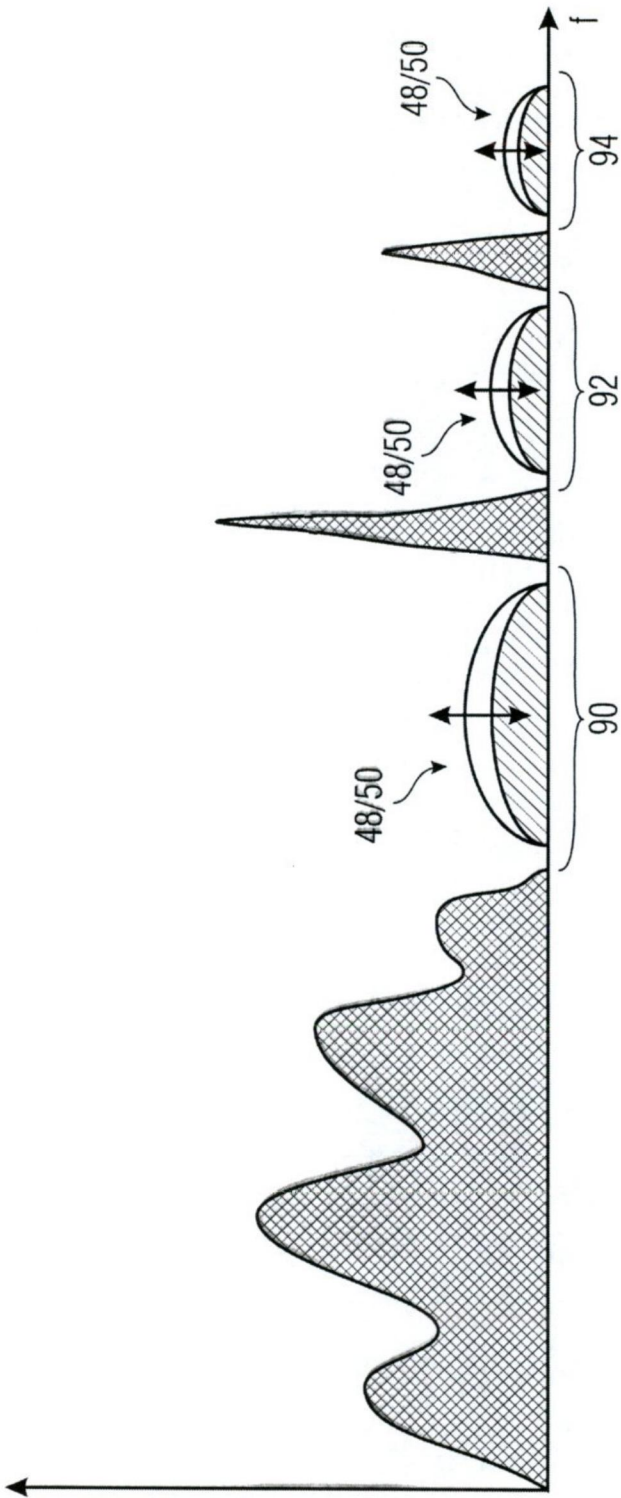


FIG 8

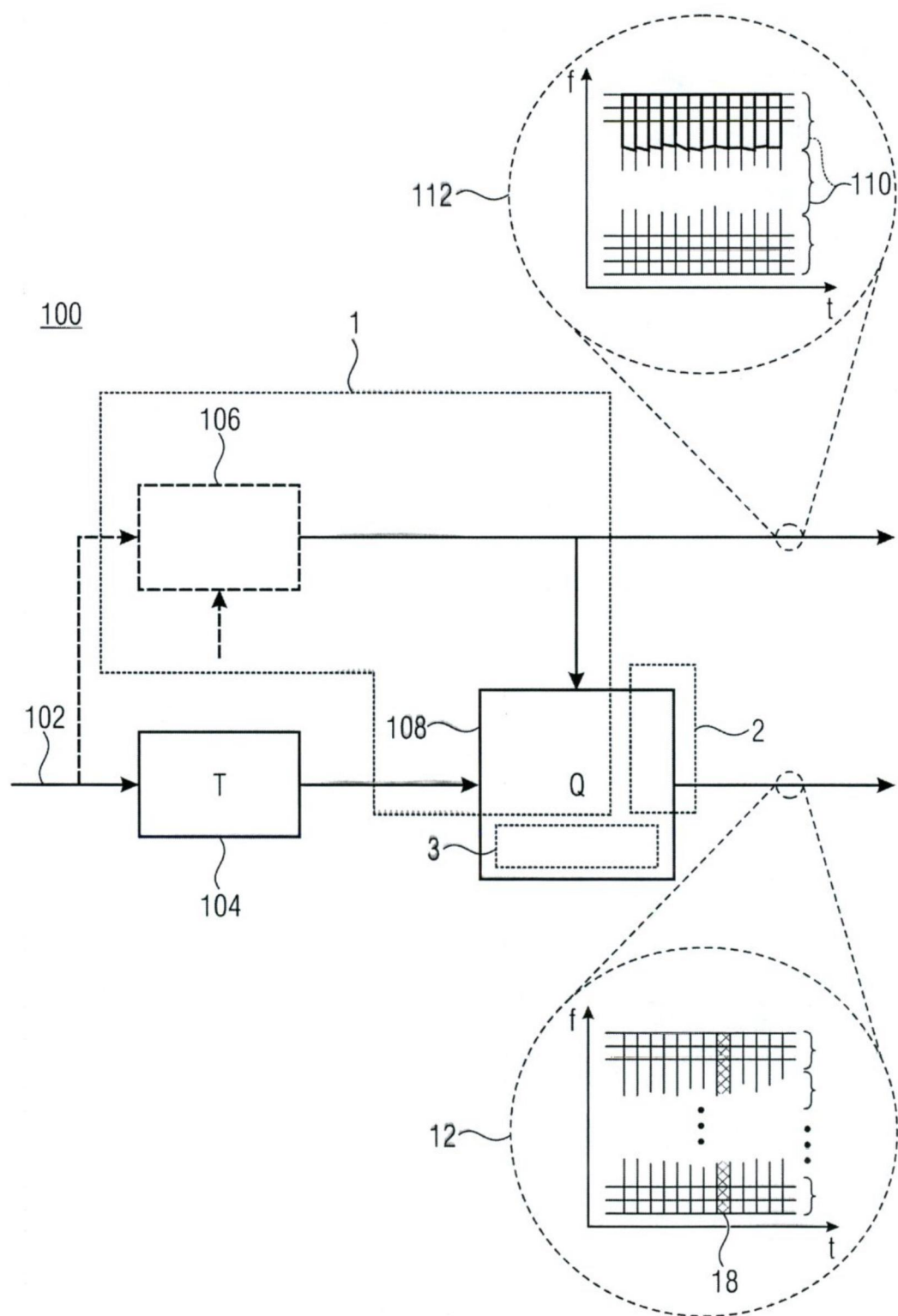


FIG 9

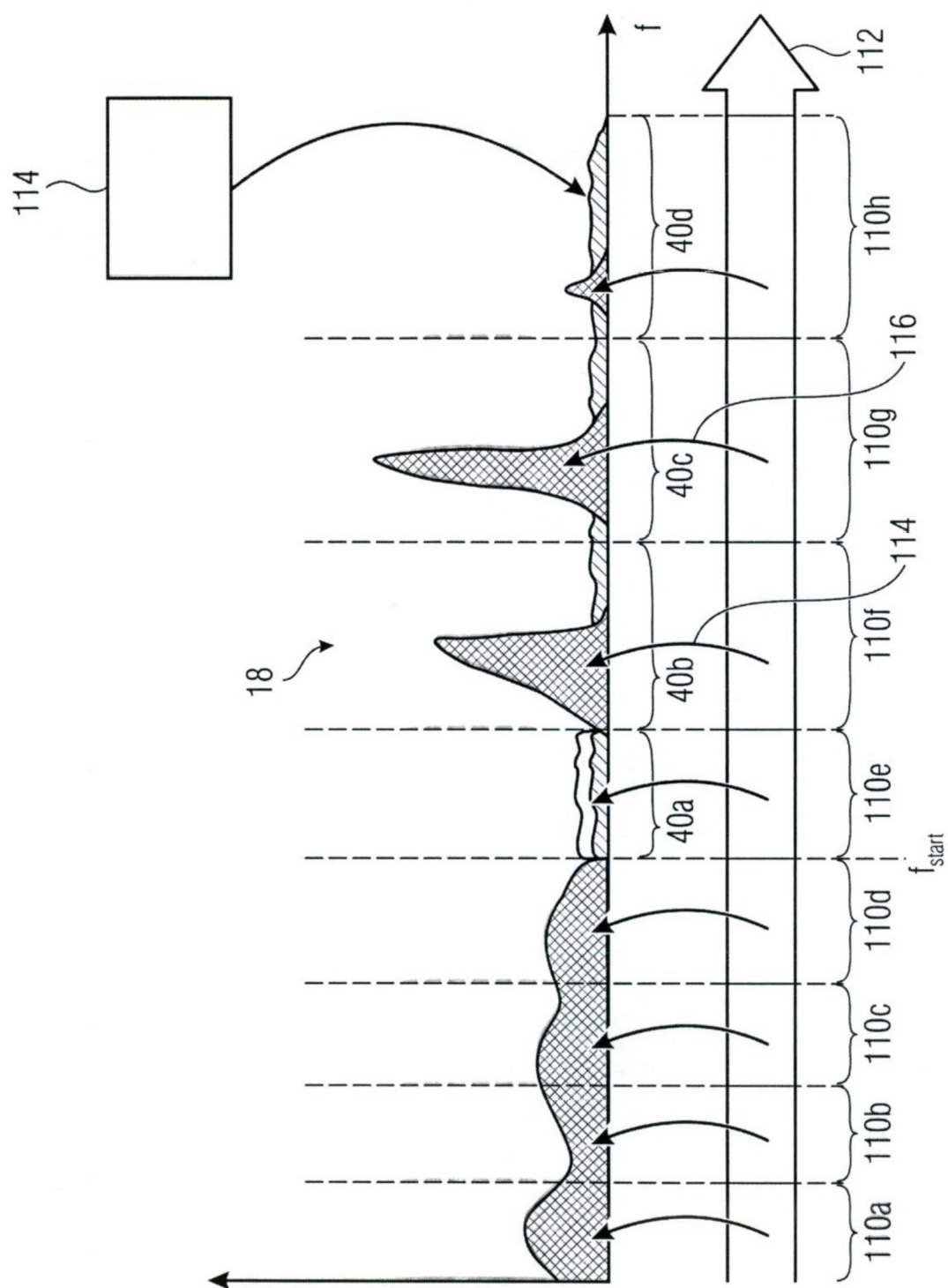


FIG 10

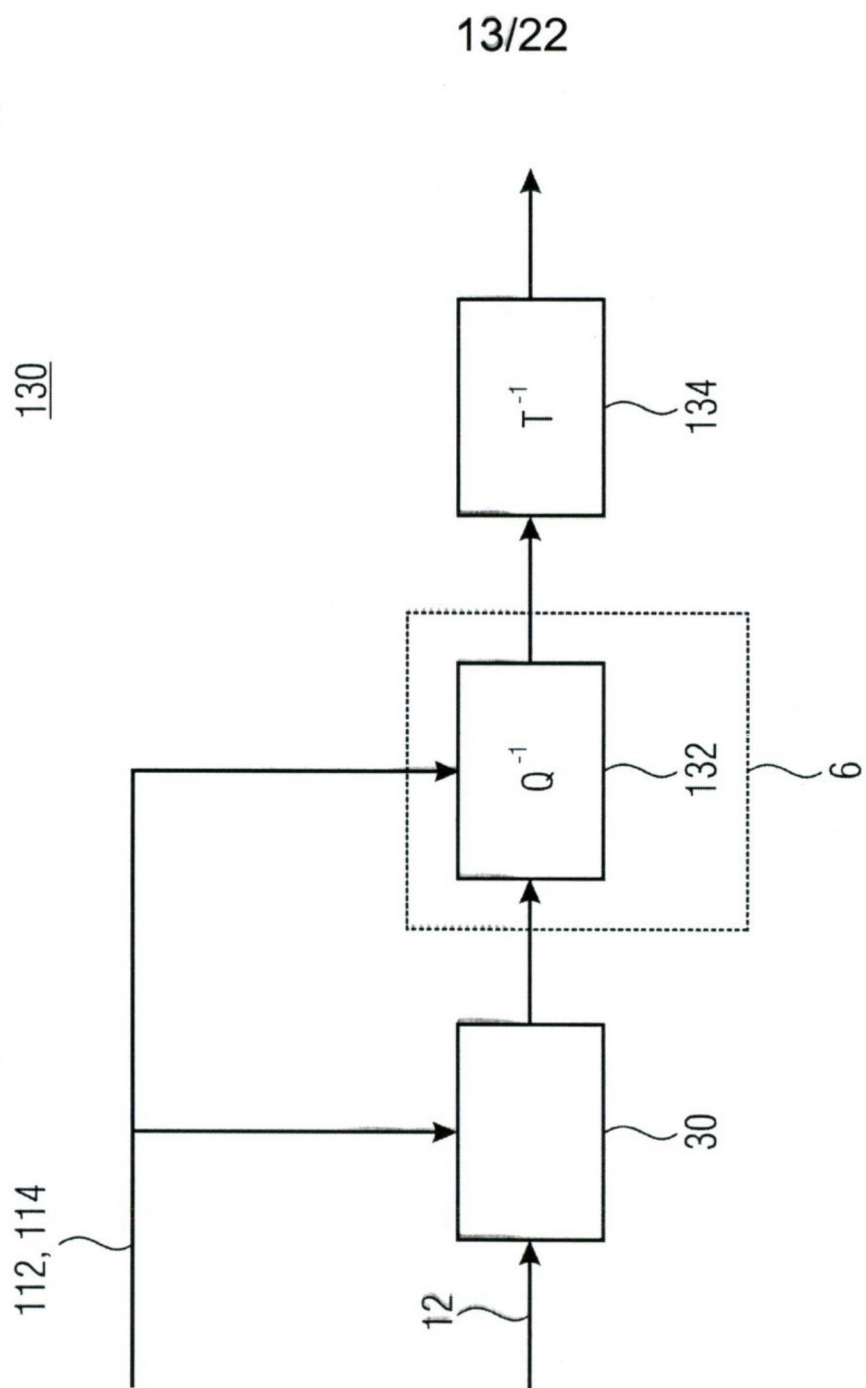


FIG 11

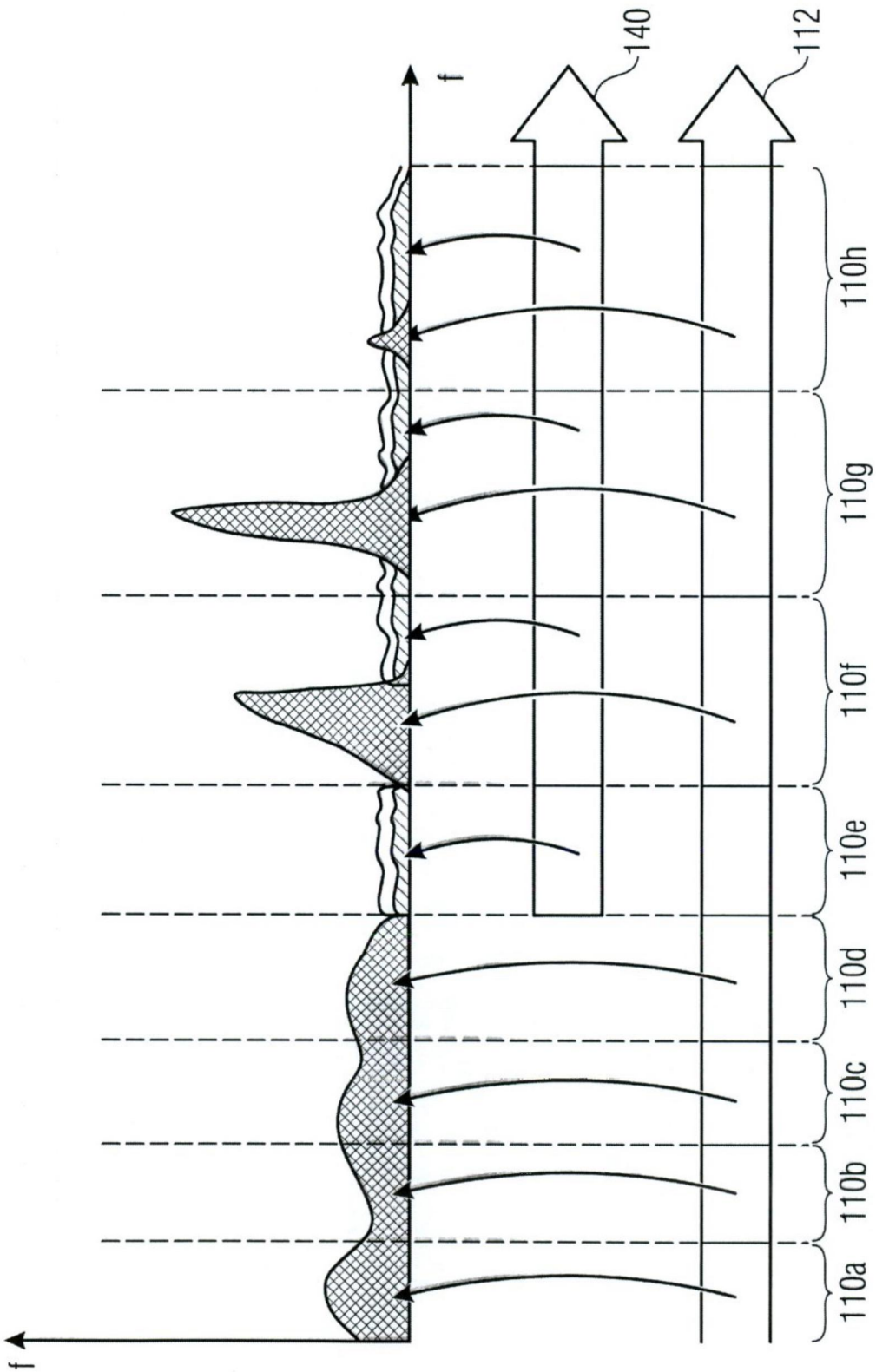


FIG 12

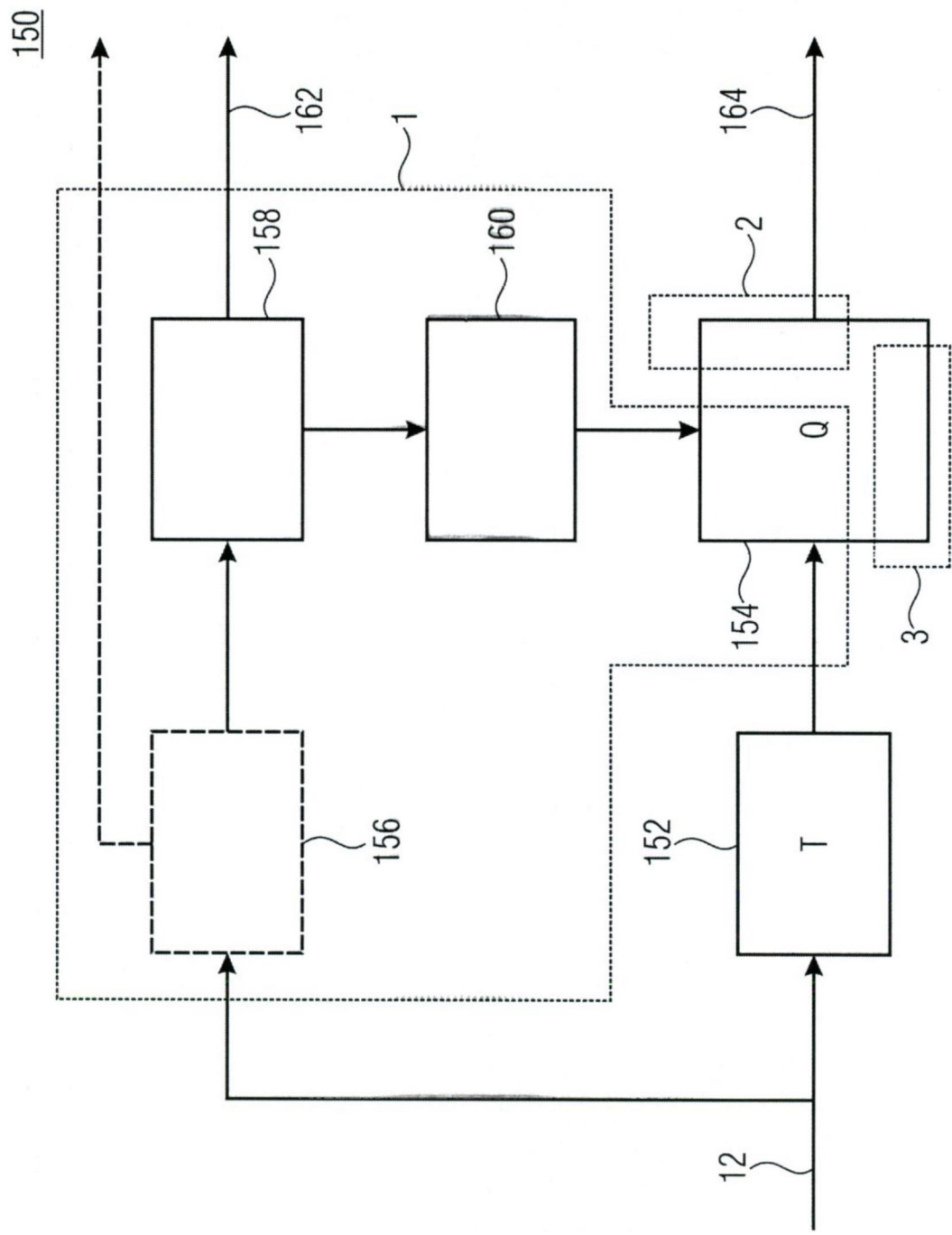


FIG 13

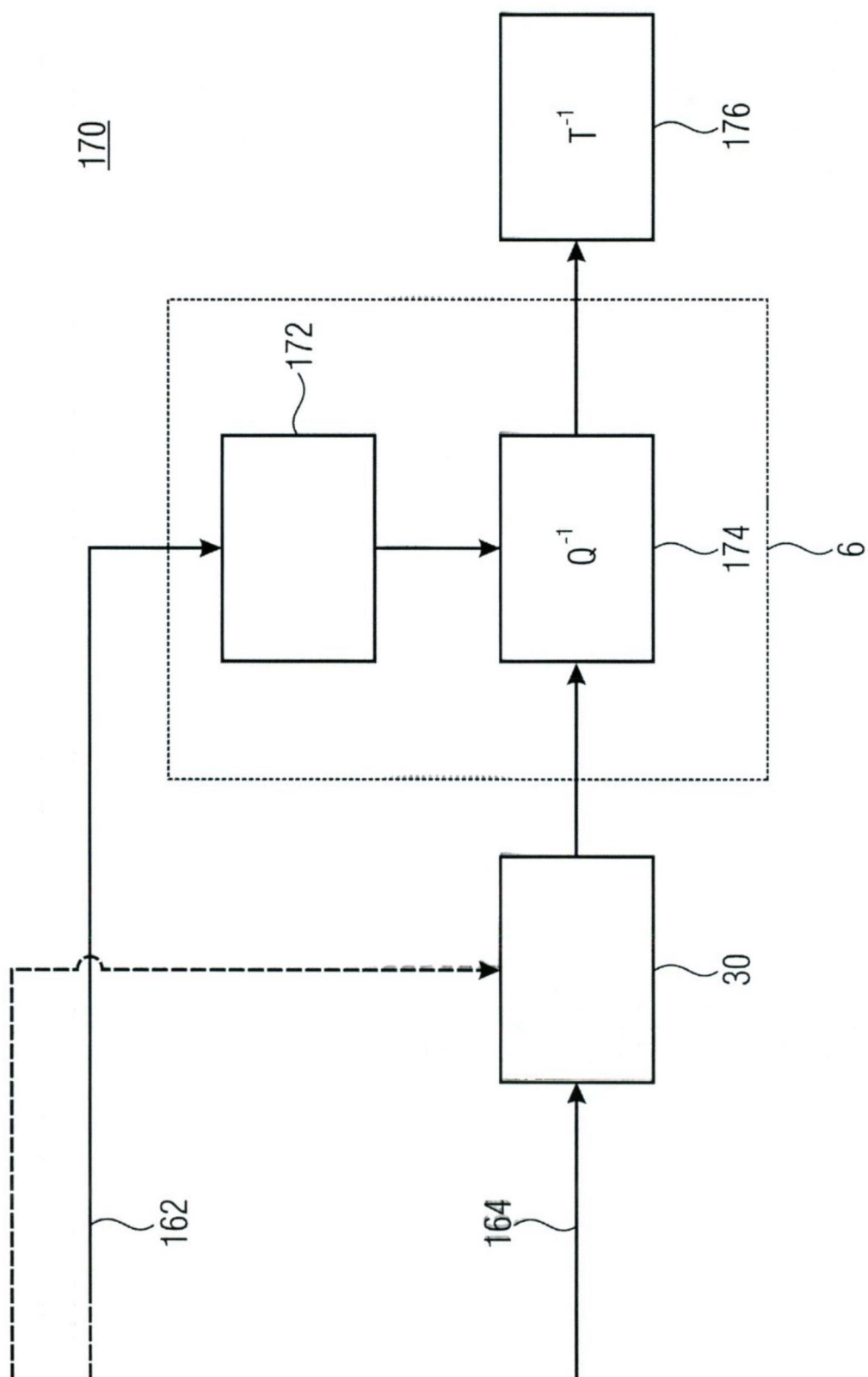


FIG 14

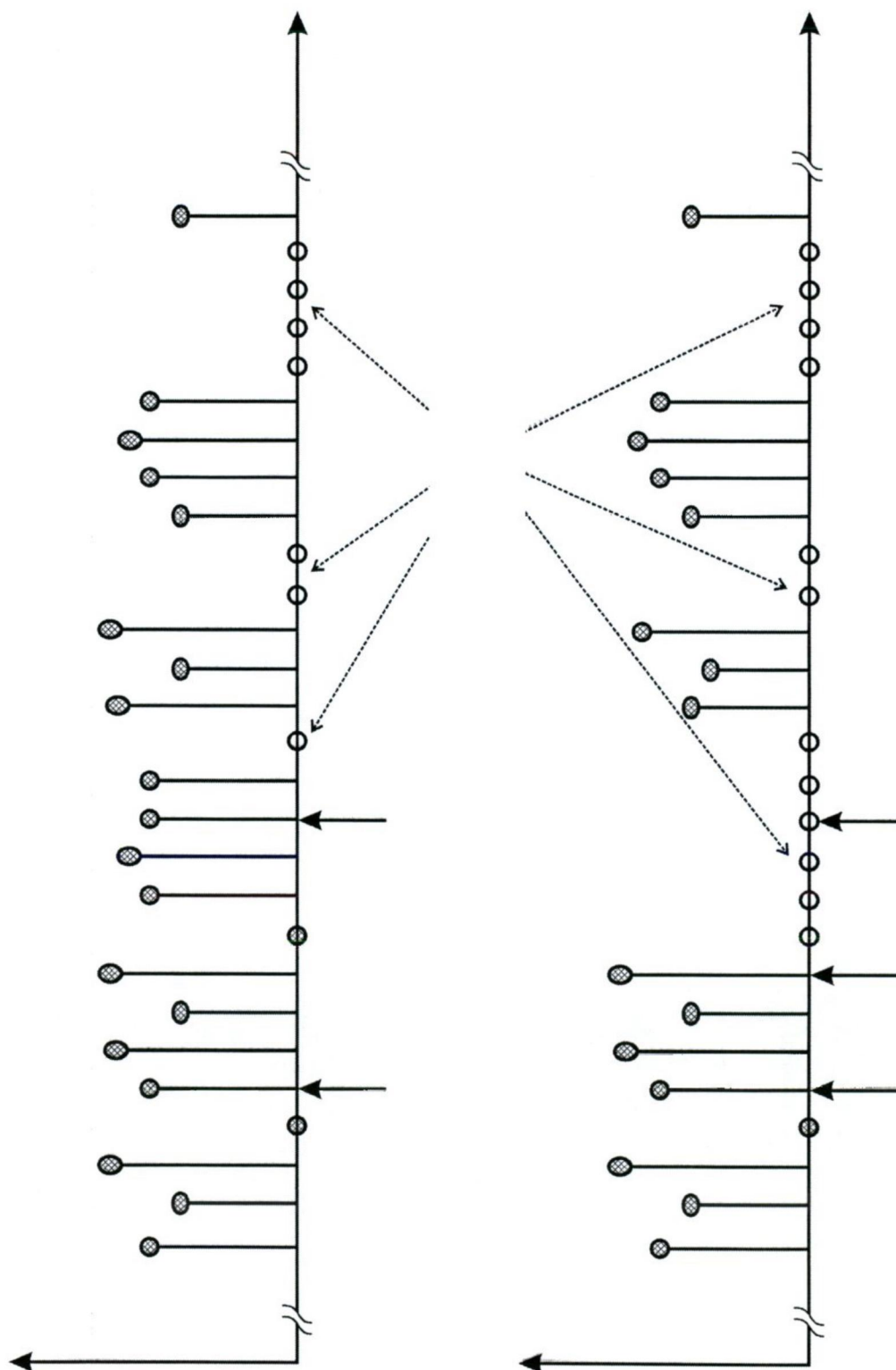
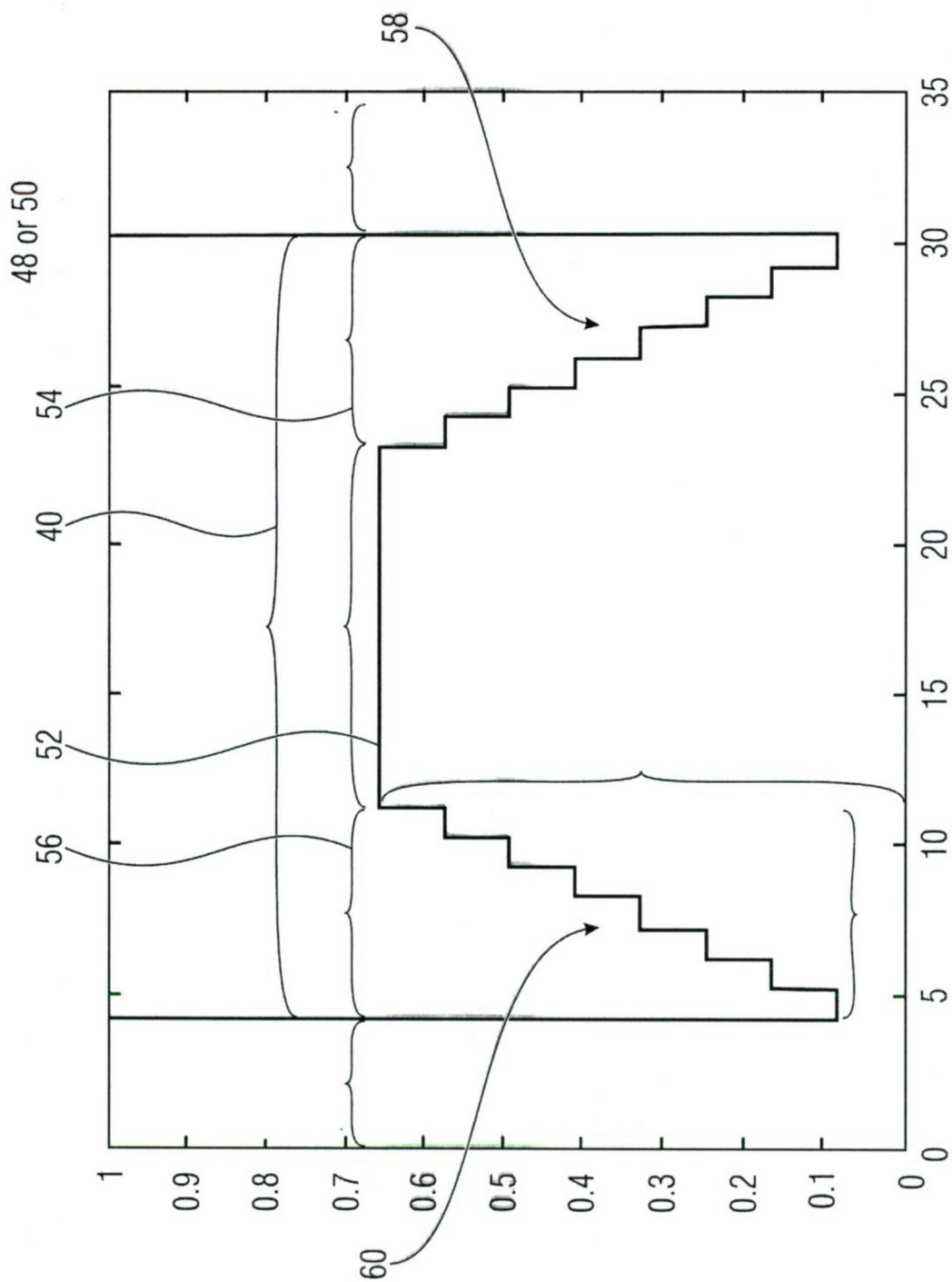


FIG 15



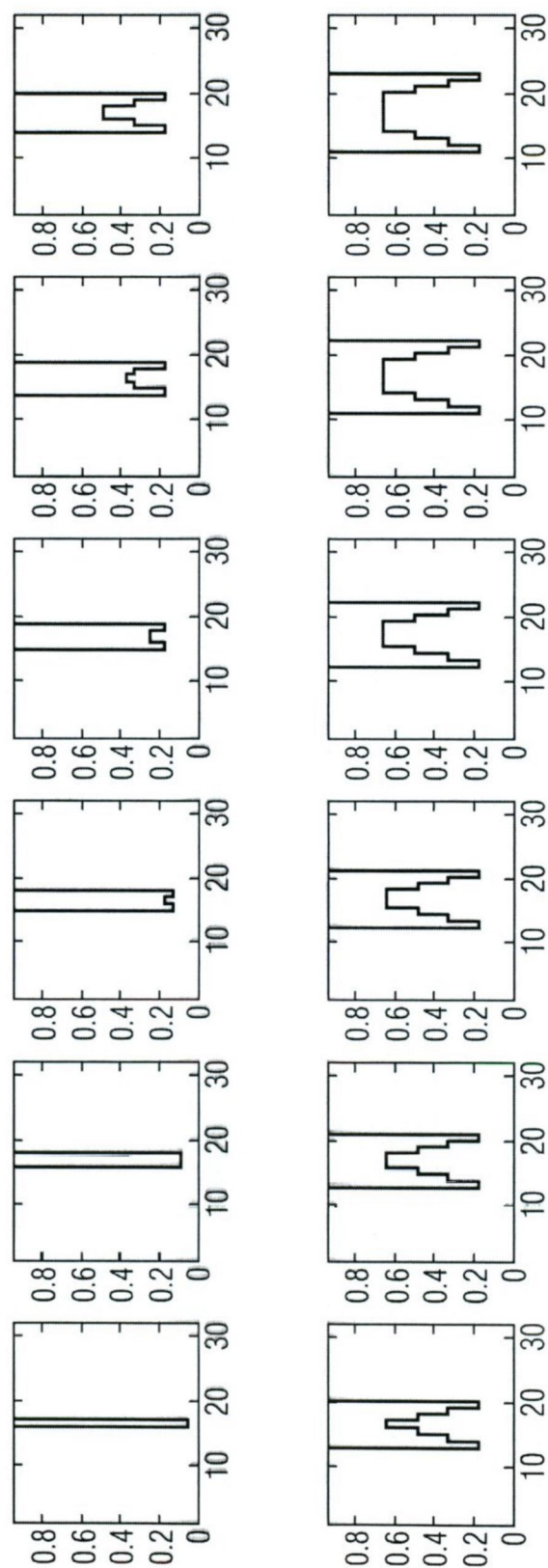


FIG 17A

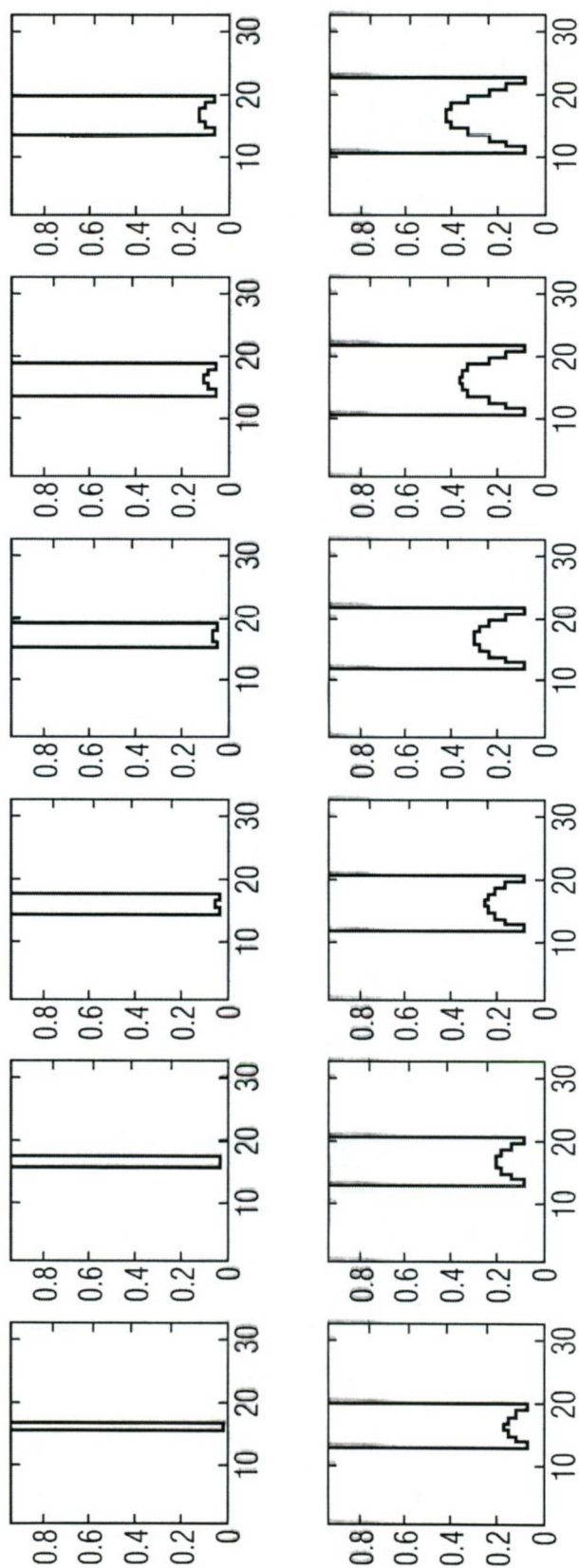


FIG 17B

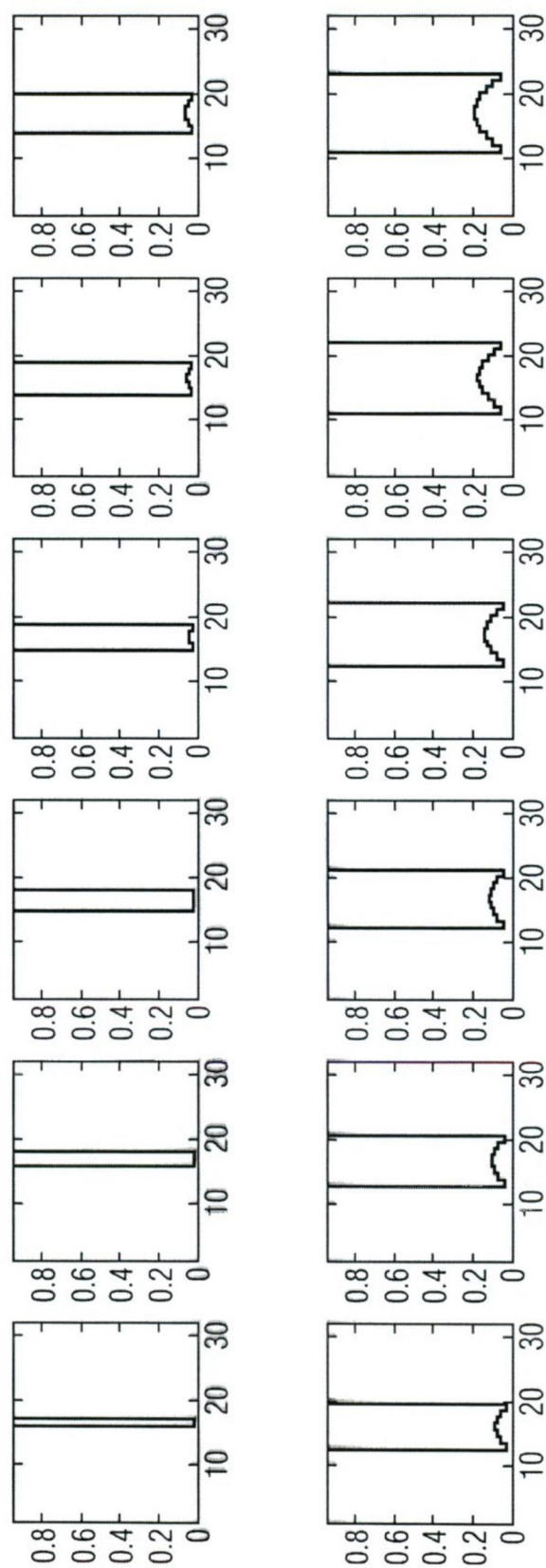


FIG 17C

22/22

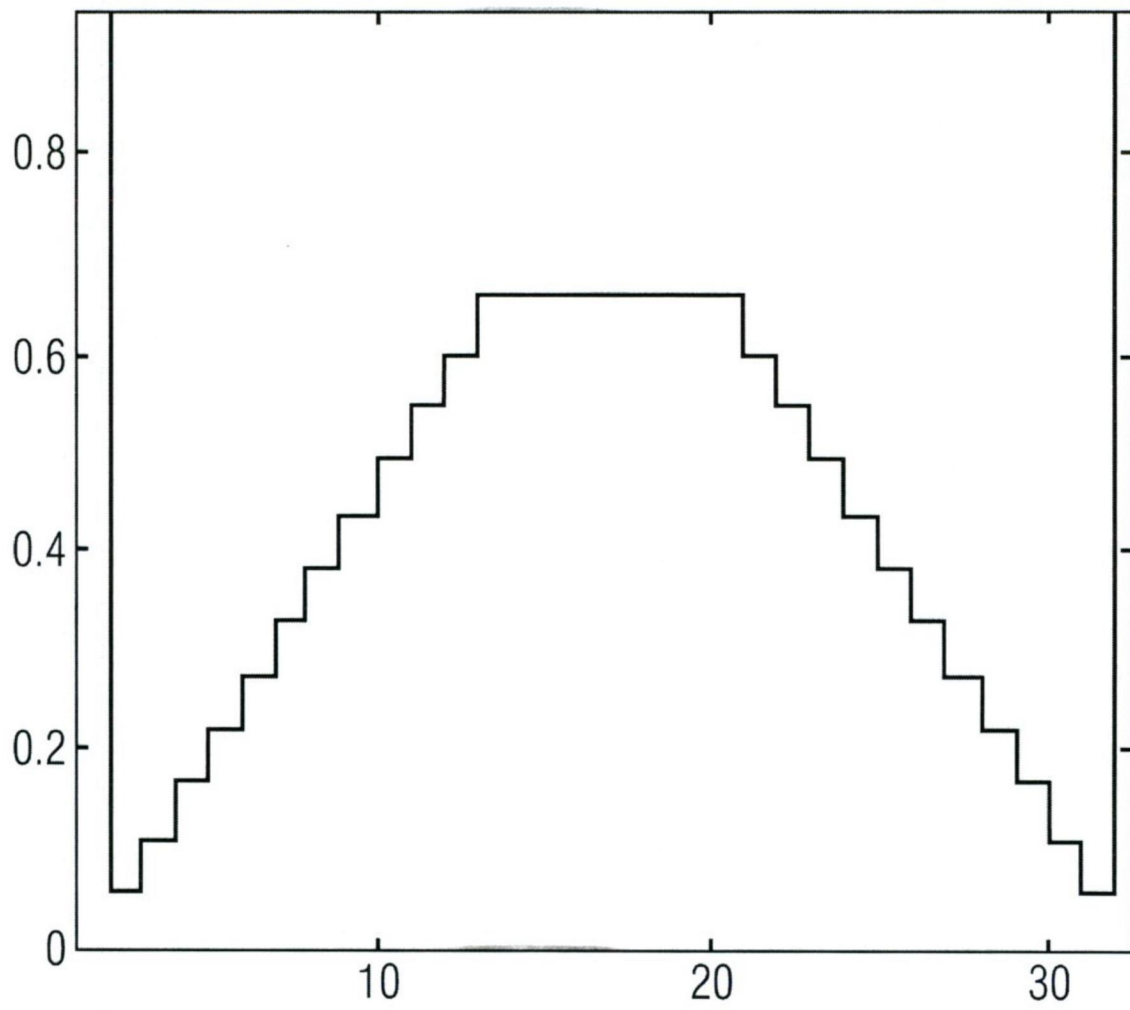


FIG 17D