

(12) FASCÍCULO DE PATENTE DE INVENÇÃO

(22) Data de pedido: 2014.01.22	(73) Titular(es): TELEFONAKTIEBOLAGET LM ERICSSON (PUBL) 164 83 STOCKHOLM	SE
(30) Prioridade(s): 2013.02.05 US 201361760814 P 2013.02.05 US 201361760822 P 2013.02.05 US 201361761051 P	(72) Inventor(es): STEFAN BRUHN JONAS SVEDBERG	SE SE
(43) Data de publicação do pedido: 2015.12.16	(74) Mandatário: LUÍS MANUEL DE ALMADA DA SILVA CARVALHO RUA VÍCTOR CORDON, 14 1249-103 LISBOA	PT

(54) Epígrafe: **MÉTODO E DISPOSITIVO PARA CONTROLAR A OCULTAÇÃO DE PERDA DE TRAMA DE ÁUDIO**

(57) Resumo:

DE ACORDO COM UM EXEMPLO DE FORMA DE REALIZAÇÃO DA PRESENTE INVENÇÃO, É DESCRITO UM MÉTODO E UM DISPOSITIVO DO MESMO PARA O CONTROLO DE UM MÉTODO DE OCULTAÇÃO PARA UMA TRAMA PERDIDA DE ÁUDIO DE UM SINAL DE ÁUDIO RECEBIDO. UM MÉTODO PARA UM DESCODIFICADOR DE OCULTAÇÃO DE UMA TRAMA DE ÁUDIO PERDIDA COMPREENDE DETECTAR NUMA PROPRIEDADE DO SINAL DE ÁUDIO RECEBIDO ANTERIORMENTE E RECONSTRUÍDO, OU NUMA PROPRIEDADE ESTATÍSTICA DAS PERDAS DE TRAMA OBSERVADAS, UMA CONDIÇÃO PARA A QUAL A SUBSTITUIÇÃO DE UMA TRAMA PERDIDA PROPORCIONA QUALIDADE RELATIVAMENTE REDUZIDA. CASO UMA TAL CONDIÇÃO SEJA DETECTADA, O MÉTODO DE OCULTAÇÃO É MODIFICADO AJUSTANDO SELECTIVAMENTE UMA FASE OU MAGNITUDE DE ESPECTRO DE UM ESPECTRO DE TRAMA DE SUBSTITUIÇÃO.

RESUMO**"MÉTODO E DISPOSITIVO PARA CONTROLAR A OCULTAÇÃO DE PERDA
DE TRAMA DE ÁUDIO"**

De acordo com um exemplo de forma de realização da presente invenção, é descrito um método e um dispositivo do mesmo para o controlo de um método de ocultação para uma trama perdida de áudio de um sinal de áudio recebido. Um método para um descodificador de ocultação de uma trama de áudio perdida compreende detectar numa propriedade do sinal de áudio recebido anteriormente e reconstruído, ou numa propriedade estatística das perdas de trama observadas, uma condição para a qual a substituição de uma trama perdida proporciona qualidade relativamente reduzida. Caso uma tal condição seja detectada, o método de ocultação é modificado ajustando selectivamente uma fase ou magnitude de espectro de um espectro de trama de substituição.

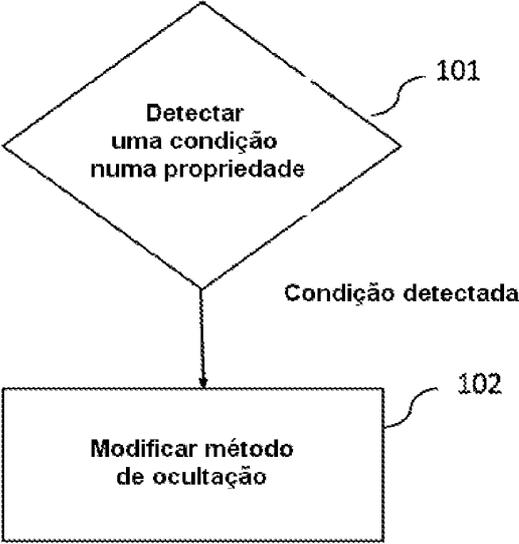


FIG. 10

DESCRIÇÃO

"MÉTODO E DISPOSITIVO PARA CONTROLAR A OCULTAÇÃO DE PERDA DE TRAMA DE ÁUDIO"

Campo técnico

O pedido refere-se a métodos e dispositivos para o controlo de um método de ocultação para uma trama perdida de áudio de um sinal de áudio recebido.

Antecedentes

Os sistemas de comunicações áudio convencionais transmitem sinais de voz e áudio em tramas, o que significa que o lado de envio organiza primeiro o sinal em segmentos curtos ou tramas de, por exemplo, 20-40 ms que, subsequentemente, são codificadas e transmitidas como uma unidade lógica em, por exemplo, um pacote de transmissão. O receptor descodifica cada uma destas unidades e reconstrói as tramas de sinal correspondentes, que por sua vez são finalmente emitidas como uma sequência contínua reconstruída de amostras de sinal. Antes da codificação há geralmente um passo de conversão de analógico para digital (A/D) que converte o sinal de voz ou áudio analógico de um microfone numa sequência de amostras de áudio. Por outro lado, na extremidade de recepção, existe tipicamente um

passo de conversão final D/A que converte a sequência de amostras reconstruídas de sinal digital num sinal analógico contínuo no tempo para a reprodução em altifalante.

No entanto, um tal sistema de transmissão para sinais de voz e de áudio pode sofrer de erros de transmissão, o que poderia levar a uma situação em que uma ou várias das tramas transmitidas não estão disponíveis no receptor para a reconstrução. Nesse caso, o decodificador tem que gerar um sinal de substituição para cada uma das tramas apagadas, ou seja não disponíveis. Isto é feito na denominada unidade de perda de trama ou de ocultação de erro do decodificador de sinais do lado do receptor. O objectivo da ocultação da perda de trama é fazer com que a perda de trama o mais inaudível quanto possível e, por conseguinte, minimizar tanto quanto possível o impacto da perda de trama sobre a qualidade do sinal reconstruído.

Os métodos de ocultação de perda de trama convencionais podem depender da estrutura ou arquitectura do codec, por exemplo por aplicação de uma forma de repetição de parâmetros de codec recebidos anteriormente. Tais técnicas de repetição de parâmetro são claramente dependentes dos parâmetros específicos do codec utilizado e, portanto, não são facilmente aplicável para outros codecs com uma estrutura diferente. Os métodos actuais de ocultação de perda de trama podem, por exemplo, aplicar o conceito de congelação e extrapolar os parâmetros de uma trama recebida anteriormente de modo a gerar uma trama de

substituição para a trama perdida.

Estes métodos de ocultação de perda de trama do estado da técnica incorporam alguns esquemas de manipulação de perda de rajada. Em geral, depois de várias perdas de tramas numa linha, o sinal sintetizado é atenuado até que esteja completamente silenciado após longas rajadas de erros. Além disso, os parâmetros de codificação que são essencialmente repetidos e extrapolados são modificados de tal modo que a atenuação é conseguida e os picos espectrais são nivelados.

As técnicas actuais de ocultação de perda de trama do estado da técnica aplicam tipicamente o conceito de congelação e extrapolação dos parâmetros de uma trama recebida anteriormente de modo a gerar uma trama de substituição para a trama perdida. Muitos codecs de voz paramétricos tais como codecs preditivos lineares como AMR ou AMR-WB congelam normalmente os parâmetros recebidos anteriormente ou usam alguma extrapolação dos mesmos e usam o descodificador com eles. Em essência, o princípio é o de ter um determinado modelo para a codificação/descodificação e aplicar o mesmo modelo com parâmetros congelados ou extrapolados. As técnicas de ocultação de perda de trama da AMR e AMR-WB podem ser consideradas como representativas. Elas estão especificadas em pormenor nas especificações das normas correspondentes.

Muitos codecs fora da classe de codecs de áudio

aplicam-se para técnicas de codificação do domínio da frequência. Isto significa que depois de alguma transformação de domínio de frequência é aplicado um modelo de codificação nos parâmetros espectrais. O descodificador reconstrói o espectro do sinal dos parâmetros recebidos e finalmente transforma o espectro de volta para um sinal de tempo. Normalmente, o sinal de tempo é reconstruído trama a trama. Tais tramas são combinadas por técnicas de sobreposição-adição para o sinal reconstruído final. Mesmo no caso de codecs de áudio, a ocultação de erros do estado da técnica aplica normalmente o mesmo ou pelo menos um modelo de descodificação semelhante para tramas perdidas. Os parâmetros no domínio da frequência de uma trama recebida anteriormente são congelados ou adequadamente extrapolados e depois usados na conversão no domínio da frequência-tempo. Exemplos de tais técnicas são proporcionados com os codecs de áudio 3GPP de acordo com as normas 3GPP.

A técnica anterior US 2004/122680 descreve um sistema para a ocultação de erro de trama que ensina a ajustar a magnitude da trama de substituição, dependendo do número de tramas perdidas consecutivas.

Sumário

As soluções actuais do estado da técnica para ocultação de perda de trama sofrem normalmente de deficiências de qualidade. O problema principal é que o

parâmetro da técnica de congelamento e extrapolação e a reaplicação do mesmo modelo descodificador, mesmo para tramas perdidas, nem sempre garante uma evolução de sinal suave e fiel a partir das tramas de sinal previamente descodificados para a trama perdida. Isto conduz normalmente a descontinuidades do sinal audível com o correspondente impacto na qualidade.

São descritos novos esquemas de ocultação de perda de tramas para sistemas de transmissão de voz e áudio. Os novos esquemas melhoram a qualidade em caso de perda de tramas sobre a qualidade alcançável com as técnicas de perda de tramas da técnica anterior

O objectivo das presentes formas de realização é o de controlar um esquema de ocultação perda de trama que é de preferência do tipo dos novos métodos relacionados descritos de tal modo que é alcançada a melhor qualidade de som do sinal reconstruído. As formas de realização têm como objectivo otimizar esta qualidade reconstrução tanto no que diz respeito às propriedades do sinal como na distribuição temporal das perdas de tramas. Particularmente problemáticos para a ocultação de perda de tramas para proporcionar boa qualidade são os casos em que o sinal de áudio tem propriedades que variam fortemente tais como onset ou deslocações de energia ou é espectralmente muito flutuante. Neste caso, os métodos de ocultação descritos podem repetir o onset, deslocação ou a flutuação espectral que conduz a grandes desvios do sinal original e perda de

qualidade correspondente.

Um outro caso problemático é se ocorrerem numa fila rajadas de perdas de tramas. Conceitualmente, o esquema para ocultação de perda de tramas de acordo com os métodos descritos pode lidar com tais casos, sendo que no entanto se verifica que podem ainda ocorrer perturbações tonais irritantes. É um outro objectivo da presente forma de realização atenuar tais perturbações ao mais alto grau possível.

De acordo com um primeiro aspecto, é descrito um método para a ocultação de uma trama de áudio perdida de acordo com a reivindicação 1.

De acordo com um segundo aspecto, um dispositivo encontra-se realizado para implementar uma ocultação de um nome de trama de áudio perdida, tal como descrito na reivindicação 12.

De acordo com um terceiro aspecto, é definido um programa de computador para ocultar uma trama de áudio perdida, sendo que o programa de computador compreende instruções que, quando executados por um processador, faz com que o processador oculte uma trama de áudio perdida, de acordo com o primeiro aspecto descrito acima.

De acordo com um quarto aspecto, um produto de programas informáticos compreende um suporte legível por

computador que armazena um programa de computador de acordo com o quarto aspecto acima descrito. Uma vantagem com uma forma de realização consigna o controlo de adaptações de métodos de ocultação de perda de trama que permite atenuar o impacto audível de perda de trama na transmissão de sinais de voz e áudio codificados e ainda mais sobre a qualidade conseguida com apenas os métodos de ocultação descritos. A vantagem geral das formas de realização é a de proporcionar uma evolução suave e fiel do sinal reconstruído mesmo para tramas perdidas. O impacto audível de perdas de tramas é muito reduzido em comparação com a utilização de técnicas do estado da técnica.

Breve descrição dos desenhos

Para uma compreensão mais completa dos exemplos de formas de realização da presente invenção, será feita agora referência à seguinte descrição realizada em ligação com os desenhos anexos. As figuras representam:

Figura 1 função de janela rectangular

Figura 2 combinação de janela de Hamming com a janela rectangular.

Figura 3 exemplo de um espectro de magnitude de uma função de janela.

Figura 4 espectro de linha de um exemplo de sinal sinusoidal com uma frequência f_k .

Figura 5 espectro de um sinal sinusoidal em janela com a frequência f_k .

Figura 6 barras que correspondem à magnitude de pontos de grelha de um DFT, com base numa trama de análise.

Figura 7 ajustamento de parábola por meio de pontos de rede DFT P1, P2 e P3.

Figura 8 ajuste de um lóbulo principal do espectro de uma janela.

Figura 9 ajuste da função de aproximação do lóbulo principal P através dos pontos de rede DFT P1 e P2.

Figura 10 diagrama de fluxos que ilustra um exemplo de método de acordo com as formas de realização da invenção para o controlo de um método de ocultação para uma trama perdida de áudio de um sinal de áudio recebido.

Figura 10 diagrama de fluxos que ilustra um outro exemplo de método de acordo com as formas de realização da invenção para o controlo de um método de ocultação para uma trama perdida de áudio de um sinal de áudio recebido.

Figura 12 outro exemplo de forma de realização da invenção.

Figura 13 exemplo de um dispositivo de acordo com uma forma de realização da invenção.

Figura 14 outro exemplo de um dispositivo de acordo com uma forma de realização da invenção.

Figura 15 outro exemplo de um dispositivo de acordo com uma forma de realização da invenção.

Descrição pormenorizada

O novo esquema de controlo para as novas técnicas de ocultação de perda de trama descritas envolve os seguintes passos tal como mostrado na figura 10. Deve notar-se que o método pode ser implementado num controlador num descodificador.

1. Detectar condições nas propriedades do sinal de áudio recebido anteriormente e reconstruído ou nas propriedades estatísticas das perdas de trama observadas para a substituição de uma trama perdida de acordo com os métodos descritos proporciona qualidade relativamente reduzida, 101.

2. Caso uma tal condição seja detectada no passo 1, codificar o elemento dos métodos de acordo com os quais o espectro de trama de substituição é calculado por $Z(m) = Y(m) \cdot e^{j\theta_k}$ ajustando selectivamente as fases ou as magnitudes de espectro, 102.

Análise sinusoidal

Um primeiro passo da técnica de ocultação de perda de trama à qual a nova técnica de controlo pode ser aplicada, envolve uma análise sinusoidal de uma parte do sinal recebido anteriormente. O objectivo desta análise sinusoidal é o de encontrar as frequências das sinusóides principais do referido sinal, e a suposição subjacente é que o sinal é constituído por um número limitado de sinusóides individuais, isto é que é um sinal de multi-seno

do seguinte tipo:

$$s(n) = \sum_{k=1}^K a_k \cdot \cos\left(2\pi \frac{f_k}{f_s} \cdot n + \varphi_k\right)$$

Nesta equação K é o número de sinusóides que se assume que o sinal é constituído. Para cada um das sinusóides com índice $k = 1 \dots K$, a_k é a amplitude, f_k é a frequência, e φ_k é a fase. A frequência de amostragem é denominada de f_s e o índice das amostras $s(n)$ do sinal discreto de tempo do tempo de n .

É de importância principal encontrar frequências tão exactas das sinusóides quanto possível. Enquanto que um sinal sinusoidal ideal teria um espectro de linha com frequências de linha f_k , encontrar os seus valores verdadeiros iriam em princípio requer tempo de medição infinito. Por isso, é na prática difícil de encontrar estas frequências, uma vez que elas só pode ser estimadas com base num período de medição curto, o que corresponde ao segmento de sinal utilizado para a análise sinusoidal aqui descrita; este segmento de sinal é daqui em diante designado como uma trama de análise. Uma outra dificuldade é que o sinal pode, na prática, ser variante no tempo, o que significa que os parâmetros da equação acima variam ao longo do tempo. Assim, por um lado é desejável a utilização de uma trama de análise longa tornando a medição mais precisa; por outro lado seria necessário um período de medição curto de modo de lidar melhor com possíveis variações de sinal. Uma boa solução de compromisso é a de

usar um comprimento de trama de análise na ordem de, por exemplo, 20-40 ms.

Uma possibilidade preferida para identificar as frequências das sinusóides f_k é fazer uma análise no domínio da frequência da trama de análise. Para este fim a trama de análise é transformada no domínio da frequência, por exemplo por meio de DFT ou DCT ou transformações no domínio da frequência. No caso de ser utilizado um DFT da trama de análise, o espectro é dado por:

$$X(m) = DFT(w(n) \cdot x(n)) = \sum_{n=0}^{L-1} e^{-j\frac{2\pi}{L}mn} \cdot w(n) \cdot x(n)$$

Nesta equação $w(n)$ indica a função de janela com a qual a trama de análise de comprimento L é extraída e pesada. As funções de janela típicas são, por exemplo, janelas rectangulares que são iguais a 1 para $n \in [0 \dots L-1]$ e de outro modo 0, tal como mostrado na figura 1. Assume-se aqui que os índices de tempo do sinal de áudio recebido anteriormente são definidos de tal modo que a trama de é referenciada pelos índices de tempo $n=0 \dots L-1$. Outras funções de janela, que podem ser mais adequadas para a análise espectral são, por exemplo, janela de Hamming, janela de Hamming, janela de Kaiser ou janela de Blackman. Uma função de janela que se descobriu ser particularmente útil é uma combinação de janela de Hamming com a janela rectangular. Esta janela tem uma forma de flanco ascendente como a metade esquerda de uma janela de Hamming de

comprimento L_1 e uma forma de flanco descendente como a metade direita de uma janela de Hamming de comprimento L_1 e entre os flancos ascendentes e descendentes a janela é igual a 1 para o comprimento de $L-L_1$, tal como se apresentado na figura 2.

Os picos do espectro de magnitude da trama de análise de janela $|X(m)|$ constituem uma aproximação das frequências f_k sinusoidais necessárias. A precisão desta aproximação está contudo limitada pelo espaçamento de frequência da DFT. Com a DFT com um comprimento de bloco L a precisão está limitada a $\frac{f_s}{2L}$

Os testes mostram que este nível de precisão pode ser muito baixo no âmbito dos métodos aqui descritos. A precisão melhorada pode ser obtida com base nos resultados da seguinte consideração:

O espectro da trama de análise de janela é dado pela convolução do espectro da função de janela com o espectro de linha do sinal de modelo sinusoidal $S(\Omega)$, subsequentemente amostrado nos pontos da rede de DFT:

$$X(m) = \int_{-\pi}^{\pi} \delta(\Omega - m \cdot \frac{2\pi}{L}) \cdot (W(\Omega) * S(\Omega)) \cdot d\Omega$$

Ao utilizar a expressão do espectro do sinal de modelo sinusoidal, isto pode ser escrito como

$$X(m) = \frac{1}{2} \int_{2\pi} \delta(\Omega - m \cdot \frac{2\pi}{L}) \cdot \sum_{k=1}^K a_k \cdot \left(W(\Omega + 2\pi \frac{f_k}{f_s}) \cdot e^{-j\varphi_k} + W(\Omega - 2\pi \frac{f_k}{f_s}) \cdot e^{j\varphi_k} \right) \cdot d\Omega$$

Assim, a amostra do espectro é dada por

$$X(m) = \frac{1}{2} \sum_{k=1}^K a_k \cdot \left(W(2\pi(\frac{m}{L} + \frac{f_k}{f_s})) \cdot e^{-j\varphi_k} + W(2\pi(\frac{m}{L} - \frac{f_k}{f_s})) \cdot e^{j\varphi_k} \right),$$

com

$$m=0 \dots L-1.$$

Com base nesta análise, assume-se que os picos observados no espectro de magnitude da haste da trama de análise de um sinal sinusoidal de janela com K sinusóides onde as verdadeiras frequências sinusóides se encontram na vizinhança dos picos.

Deixemos m_k ser o índice DFT (ponto de rede) do $k^{\text{gésimo}}$ pico observado, então a frequência correspondente é $\hat{f}_k = \frac{m_k}{L} \cdot f_s$ a qual pode ser considerada como uma aproximação da verdadeira frequência sinusoidal f_k . A verdadeira frequência sinusóide f_k pode ser assumida com estando dentro do intervalo

$$\left[\left(m_k - \frac{1}{2} \right) \cdot \frac{f_s}{L}, \left(m_k + \frac{1}{2} \right) \cdot \frac{f_s}{L} \right]$$

Para maior visibilidade note-se que a convolução do espectro da função de janela com o espectro do espectro

de linha do sinal de modelo sinusoidal pode ser entendida como uma sobreposição de versões deslocadas de frequência do espectro de função de janela, em que as frequências de deslocação são as frequências das sinusóides. Esta sobreposição é então amostrada nos pontos de rede DFT. Estes passos são ilustrados pelas seguintes figuras. A figura 3 mostra um exemplo do espectro de magnitude de uma função de janela. A figura 4 mostra o espectro da magnitude (espectro de linha) de um exemplo de sinal sinusoidal com uma única sinusóide de frequência. A figura 5 mostra o espectro de magnitude do sinal sinusoidal de janela que replica e sobrepõe o espectro de janela deslocado na frequência nas frequências da sinusóide. As barras na figura 6 correspondem à magnitude dos pontos da rede do DFT da sinusóide de janela que são obtidos através do cálculo da DFT da trama de análise. Deve notar-se que todos os espectros são periódicos com o parâmetro de frequência normalizada Ω onde $\Omega = 2\pi$ que corresponde à frequência de amostragem f_s .

A descrição anterior e a ilustração da figura 6 sugerem que uma melhor aproximação das verdadeiras frequências sinusoidais só pode ser encontrada através do aumento da resolução da pesquisa sobre a resolução da frequência da transformação usada de domínio da frequência.

Uma forma preferida de encontrar melhores aproximações das frequências f_k das sinusóides é aplicar a interpolação parabólica. Uma dessas soluções consiste em

ajustar parábolas através dos pontos da rede do espectro de magnitude DFT que circundam os picos e calcular as respectivas frequências pertencentes aos valores máximos da parábola. A escolha adequada para a ordem das parábolas é 2. Pormenorizadamente, pode ser aplicado o seguinte procedimento:

1. Identificar os picos de DFT da trama de análise de janela. A pesquisa de pico irá entregar o número de picos K e os índices DFT correspondentes dos picos. A pesquisa de picos pode tipicamente ser feita no espectro de magnitude DFT ou no espectro de magnitude logarítmica DFT.

2. Para cada pico k (com $k = 1 \dots K$) com o índice DFT correspondente m_k ajusta uma parábola através dos três pontos $\{P1; P2; P3\} = \{(m_k - 1, \log(|X(m_k - 1)|)|); (m_k, \log(|X(m_k)|)|); (m_k + 1, \log(|X(m_k + 1)|)|)\}$. Isto resulta em coeficientes de parábola $b_k(0)$, $b_k(1)$, $b_k(2)$ da parábola definida por

$$p_k(q) = \sum_{i=0}^2 b_k(i) \cdot q^i$$

Este encaixe de parábola encontra-se ilustrado na figura 7.

3. Para cada uma das parábolas K calcular o índice de frequência interpolada m_k que corresponde ao valor de

q para o qual a parábola tem o seu máximo. Use $\hat{f}_k = \hat{m}_k \cdot f_s/L$ como aproximação para a frequência sinusoidal f_k

A abordagem descrita proporciona bons resultados mas pode ter algumas limitações uma vez que as parábolas não aproximam a forma do lóbulo principal do espectro de magnitude $|W(\Omega)|$ da função de janela. Um esquema alternativo de fazer isso é uma estimativa melhorada de frequência usando uma aproximação de lóbulo principal, descrita como se segue. A ideia principal desta alternativa é uma função de encaixe $P(q)$, que aproxima o lóbulo principal de $|W(\frac{2\pi}{L} \cdot q)|$ através dos pontos de rede do espectro de magnitude DFT que circundam os picos e calcula as respectivas frequências que pertencem à máxima da função. A função $P(q)$ pode ser idêntica ao espectro de magnitude deslocada na frequência $|W(\frac{2\pi}{L} \cdot (q - \hat{q}))|$ da função de janela. Por simplicidade numérica deve no entanto em vez disso ser por exemplo um polinómio que permite o cálculo directo do máximo da função. Pode ser aplicado o seguinte procedimento pormenorizado:

1. Identificar os picos de DFT da trama de análise de janela. A pesquisa de pico irá entregar o número de picos K e os índices DFT correspondentes dos picos. A pesquisa de picos pode tipicamente ser feita no espectro de magnitude DFT ou no espectro de magnitude logarítmica DFT.

2. Derivar a função $P(q)$ que se aproxima do espectro de magnitude $|W(\frac{2\pi}{L} \cdot q)|$ da função de janela ou do espectro de magnitude logarítmico $|W(\frac{2\pi}{L} \cdot q)|$ para um determinado intervalo (q_1, q_2) . A escolha da função de aproximação que se aproxima do lóbulo principal do espectro de janela encontra-se ilustrada pela figura 8.

3. Para cada pico k (com $k = 1 \dots K$) com o índice DFT m_k correspondente ajusta a função de deslocação de frequência $P(q - \hat{q}_k)$ através dos dois pontos de rede DFT que cercam o verdadeiro pico esperado do espectro contínuo do sinal sinusoidal de janela. Por isso, se $|X(m_k - 1)|$ for maior do que $|X(m_k + 1)|$ ajusta $P(q - \hat{q}_k)$ através dos pontos $\{P_1; P_2\} = \{(m_k - 1, \log(|X(m_k - 1)|)); (m_k, \log(|X(m_k)|))\}$ e de outra forma através dos pontos $\{P_1; P_2\} = \{(m_k, \log(|X(m_k)|)); (m_k + 1, \log(|X(m_k + 1)|))\}$. $P(q)$ pode, por simplicidade, ser escolhido para ser um polinómio de ordem 2 ou 4. Isto torna a aproximação no passo 2 um cálculo de regressão linear simples e o cálculo de \hat{q}_k directo. O intervalo (q_1, q_2) pode ser escolhido para ser fixo e igual para todos os picos, por exemplo, $(q_1, q_2) = (-1, 1)$, ou adaptativo.

Na abordagem adaptativa o intervalo pode ser escolhido de tal modo que a função $P(q - q_k)$ se ajusta ao lóbulo principal do espectro função de janela na gama dos

pontos de rede DFT relevantes $\{P1; P2\}$. O processo de ajuste encontra-se apresentado na figura 9.

4. Para cada um dos K parâmetros \hat{q}_k de desvio de frequência para os quais se espera o espectro contínuo do sinal sinusoidal de janela para ter o seu pico calcular $\hat{f}_k = q_k \cdot f_s/L$ como aproximação para a frequência sinusóide f_k .

Há muitos casos em que o sinal transmitido é harmónico o que significa que o sinal consiste em ondas sinusoidais cujas frequências são múltiplos inteiros de alguma frequência fundamental f_0 . Este é o caso quando o sinal é muito periódico, como por exemplo para voz sonora ou os tons sustentados de algum instrumento musical. Isto significa que as frequências do modelo sinusoidal das formas de realização não são independentes, mas sim têm um relacionamento harmónico e derivam da mesma frequência fundamental. Tomar esta propriedade harmónica em consideração pode, conseqüentemente, melhorar substancialmente a análise das frequências de componentes sinusoidais.

Uma possibilidade de melhoramento é descrita como se segue:

1. Verificar se o sinal é harmónico. Isto pode, por exemplo, ser feito através da avaliação da periodicidade do sinal antes da perda de trama. Um

método simples é a realização de uma análise de autocorrelação do sinal. O valor máximo de uma tal função de autocorrelação para algum desfasamento temporal $\tau > 0$ pode ser utilizado como um indicador. Se o valor deste máximo exceder um dado limiar, o sinal pode ser considerado harmónica. O desfasamento temporal correspondente τ corresponde então ao período do sinal que está relacionado com a frequência

fundamental através de
$$f_0 = \frac{f_s}{\tau}$$

Muitos métodos de codificação de voz preditivos lineares aplicam a denominada previsão de altura de som aberta ou de ciclo fechado ou codificação CELP usando livros de códigos adaptativos. O ganho de altura de som e os parâmetros de atraso de altura de som associados derivados por tais métodos de codificação são também indicadores úteis se o sinal for harmónica e, respectivamente, para o desfasamento temporal.

Um outro método para a obtenção de f_0 encontra-se descrito a seguir.

2. Para cada índice de componente harmónica j dentro do intervalo de número inteiro $1 \dots J_{\max}$ verificar se existe um pico no espectro de magnitude (logarítmica) DFT da trama de análise na vizinhança da componente harmónica $f_j = j \cdot f_0$. A proximidade de f_j pode ser

definida como a gama delta em de torno f_j onde delta $\frac{f_s}{L}$ corresponde à resolução de frequência do DFT $\frac{f_s}{L}$ ou seja, o intervalo

$$\left[j \cdot f_0 - \frac{f_s}{2 \cdot L}, j \cdot f_0 + \frac{f_s}{2 \cdot L} \right]$$

No caso de estar presente de tal pico com frequência sinusoidal estimada \hat{f}_k correspondente, substituir \hat{f}_k por $\hat{f}_k = j \cdot f_0$.

Para o procedimento de dois passos indicado acima existe também a possibilidade de verificar se o sinal é um componente harmónico e a derivação da frequência fundamental implicitamente, e, possivelmente de uma forma iterativa sem necessariamente usar indicadores de algum método separado. Um exemplo de uma tal técnica é dado como se segue:

Para cada $f_{0,p}$ de um conjunto de valores candidatos $\{f_{0,1} \dots f_{0,p}\}$ aplicar o passo 2 do procedimento, embora sem substituir \hat{f}_k mas com contagem de quantos picos DFT se encontram presentes dentro da vizinhança em torno das frequências de componente harmónico, ou seja, os múltiplos inteiros de $f_{0,p}$. Identificar a frequência fundamental $f_{0,pmax}$ para o qual é obtido o maior número de picos em ou em torno das frequências harmónicas. Se este maior número de picos exceder um

dado limiar, então o sinal pode ser considerado como sendo um componente harmónico. Neste caso, $f_{0,pmax}$ pode ser assumido como sendo a frequência fundamental que com o passo 2 é então executado levando a frequências sinusoidais melhoradas \hat{f}_k . Uma alternativa mais preferível é, contudo primeiro otimizar a frequência fundamental f_0 com base nas frequências de pico f_k que se descobriu coincidirem com as frequências harmónicas. Assumir um conjunto de M componentes harmónicos, ou seja, múltiplos inteiros $\{n_1 \dots n_M\}$ de alguma frequência fundamental que foram encontrados para coincidir com algum conjunto de M picos espectrais em frequências $\hat{f}_{k(m)}$, $m = 1 \dots M$, em seguida, a frequência fundamental subjacente (otimizada) $f_{0,opt}$ pode ser calculada para minimizar o erro entre as frequências harmónicas e as frequências de pico espectral. Se o erro a ser minimizado for o

erro quadrático médio
$$E_2 = \sum_{m=1}^M (n_m \cdot f_0 - \hat{f}_{k(m)})^2$$
 então a frequência fundamental óptima é calculada como

$$f_{0,opt} = \frac{\sum_{m=1}^M n_m \cdot \hat{f}_{k(m)}}{\sum_{m=1}^M n_m^2}$$

O conjunto inicial de valores candidatos $\{f_{0,1} \dots f_{0,p}\}$ pode ser obtido a partir das frequências dos picos DFT ou das frequências sinusoidais estimadas \hat{f}_k .

Uma outra possibilidade para melhorar a precisão das frequências sinusoidais estimadas f_k é a de considerar a sua evolução temporal. Para este fim, as estimativas das frequências sinusoidais de várias tramas de análise podem ser combinadas, por exemplo, por meio de cálculo da média ou previsão. Antes do cálculo da média ou previsão pode ser aplicada um rastreamento de pico que liga os picos espectrais estimados às respectivas mesmas sinusoidais subjacentes.

Aplicar o modelo sinusoidal

A aplicação de um modelo sinusoidal a fim de executar uma operação de ocultação perda de trama aqui descrita pode ser descrita como se segue.

Assume-se que um determinado segmento do sinal codificado não pode ser reconstruído por o decodificador uma vez a informação codificada correspondente não está disponível. Assume-se ainda que uma parte do sinal antes deste segmento se encontra disponível. Deixe $y(n)$, com $n = 0 \dots N-1$ ser o segmento indisponível para o qual uma trama $z(n)$ de substituição tem de ser gerada e $y(n)$ com $n < 0$ ser o sinal disponível previamente decodificado. Então, num primeiro passo é extraída uma trama protótipo do sinal disponível de comprimento L e índice inicial n_{-1} com uma função de janela $w(n)$ e transformada em domínio de frequência, por exemplo por meio de DFT:

$$Y_{-1}(m) = \sum_{n=0}^{L-1} y(n - n_{-1}) \cdot w(n) \cdot e^{-j\frac{2\pi}{L}nm}$$

A função de janela pode ser uma das funções de janela descritas acima na análise sinusoidal. De preferência, a fim de economizar complexidade numérica, a trama transformada de domínio de frequência deve ser idêntica à utilizada durante a análise sinusoidal.

Num passo seguinte é aplicado o modelo pressuposto sinusoidal. De acordo com isto o DFT da trama do protótipo pode ser escrito como se segue:

$$Y_{-1}(m) = \frac{1}{2} \sum_{k=1}^K a_k \cdot \left(W\left(2\pi\left(\frac{m}{L} + \frac{f_k}{f_s}\right)\right) \cdot e^{-j\varphi_k} + W\left(2\pi\left(\frac{m}{L} - \frac{f_k}{f_s}\right)\right) \cdot e^{j\varphi_k} \right)$$

O próximo passo é o de perceber que o espectro da função de janela utilizada tem apenas uma contribuição significativa numa gama de frequências próxima de zero. Tal como ilustrado na figura 3, o espectro de magnitude da função de janela é grande para frequências próximas de zero e pequeno de outra forma (dentro da gama de frequências normalizada de $-\pi$ a π , correspondente a metade da frequência de amostragem). Assim, como uma aproximação, assume-se que o espectro de janela $W(m)$ é não-zero apenas durante um intervalo $M = [-m_{min}, m_{max}]$, com m_{min} e m_{max} sendo pequenos números positivos. Em particular, é utilizada uma aproximação do espectro de função de janela de tal forma que para cada k as contribuições do espectro de janela

deslocado na expressão acima são estritamente de não-sobreposição. Assim, na equação acima para cada índice de frequência há sempre apenas no máximo a contribuição de uma soma, ou seja de um espectro de janela deslocado. Isto significa que a expressão acima reduz-se à seguinte expressão aproximada:

$$\hat{Y}_1(m) = \frac{a_k}{2} \cdot W\left(2\pi\left(\frac{m}{L} - \frac{f_k}{f_s}\right)\right) \cdot e^{j\omega_k m}$$

para $m \in M_k$ não negativo e para cada k .

Aqui, M_k indica o intervalo do número inteiro $N_k = \left\{ \text{round}\left(\frac{f_k}{f_s} \cdot L\right) - m_{\min,k}, \text{round}\left(\frac{f_k}{f_s} \cdot L\right) + m_{\max,k} \right\}$ em que $m_{\min,k}$ e $m_{\max,k}$ cumpre a restrição explicada acima de tal modo que os intervalos não se sobrepõem. Uma escolha adequada para $m_{\min,k}$ e $m_{\max,k}$ é a configurá-los para um pequeno valor inteiro δ , por exemplo $\delta = 3$. Se, contudo, os índices de DFT relacionados com duas frequências sinusoidais vizinhas f_k e f_{k+1} forem inferiores a 2δ , então δ está definido para $\text{floor}\left(\frac{\text{round}\left(\frac{f_{k+1}}{f_s} \cdot L\right) - \text{round}\left(\frac{f_k}{f_s} \cdot L\right)}{2}\right)$ de tal forma que é assegurado que os intervalos não se sobrepõem. A função piso ($\text{floor}(\cdot)$) é o número inteiro mais próximo ao argumento da função que é menor ou igual ao mesmo.

O passo seguinte de acordo com a forma de realização é o de aplicar o modelo sinusoidal de acordo com a expressão acima e evoluir as suas sinusoidais K no tempo. A suposição de que os índices de tempo do segmento apagado

em comparação com os índices de tempo da trama de protótipo difere em n_{-1} amostras significa que as fases das sinusoidais avançam em

$$\theta_k = 2\pi \cdot \frac{f_k}{f_s} n_{-1}$$

Assim, o espectro de DFT do modelo sinusoidal evoluído é dado por:

$$Y_0(m) = \frac{1}{2} \sum_{k=1}^K a_k \left(W\left(2\pi\left(\frac{m}{L} + \frac{f_k}{f_s}\right)\right) \cdot e^{-j(\theta_k + \phi_k)} + W\left(2\pi\left(\frac{m}{L} - \frac{f_k}{f_s}\right)\right) \cdot e^{j(\theta_k + \phi_k)} \right)$$

Aplicar novamente a aproximação de acordo com a qual o espectro da função de janela deslocada não se sobrepõe dá:

$$\tilde{Y}_0(m) = \frac{a_k}{2} \cdot W\left(2\pi\left(\frac{m}{L} - \frac{f_k}{f_s}\right)\right) \cdot e^{j(\theta_k + \phi_k)} \quad \text{para } m \in M_k \text{ não negativo e para cada } k.$$

Comparar o DFT da trama protótipo $Y_{-1}(m)$ com o DFT do modelo sinusoidal evoluído $Y_0(m)$ utilizando a aproximação, verificou-se que o espectro de magnitude permanece inalterado enquanto que a fase é deslocada por $\theta_k = 2\pi \cdot \frac{f_k}{f_s} n_{-1}$, para cada $m \in M_k$. Deste modo, os coeficientes de espectro de frequência da trama de protótipo na vizinhança de cada sinusoidal estão deslocados proporcionais à frequência sinusoidal f_k e a diferença de tempo entre a trama de áudio perdida e a trama de protótipo

$n-1$.

Assim, de acordo com a forma de realização a trama de substituição pode ser calculada pela seguinte expressão:

$z(n) = \text{IDTF}\{Z(m)\}$ com $Z(m) = Y(m) \cdot e^{j\theta_k}$ para $m \in M_k$ não negativo e para cada k .

Uma forma de realização específica aborda a randomização da fase para os índices de DFT que não pertencem a qualquer intervalo M_k . Tal como descrito acima, os intervalos M_k , $k = 1 \dots K$ têm que ser ajustados de tal modo que eles não se encontram totalmente não-sobrepostos o que é feito usando algum parâmetro δ que controla o tamanho dos intervalos. Pode acontecer que δ seja menor em relação à distância da frequência das duas sinusoidais vizinhas. Assim, neste caso acontece que existe uma lacuna entre dois intervalos. Por conseguinte, para os índices m DFT correspondentes não se encontra definido qualquer desvio de fase de acordo com a expressão acima $Z(m) = Y(m) \cdot e^{j\theta_k}$. Uma escolha apropriada de acordo com esta forma de realização é a de fazer a escolha aleatória da fase para estes índices, obtendo-se $Z(m) = Y(m) \cdot e^{j2\pi \text{rand}(\cdot)}$, em que a função $\text{rand}(\cdot)$ devolve algum número aleatório.

Verificou-se benéfico para a qualidade dos sinais reconstruídos otimizar o tamanho dos intervalos M_k . Em particular, os intervalos devem ser maiores se o sinal for

muito tonal, ou seja, quando ele tem picos espectrais claros e distintos. Este é o caso, por exemplo, quando o sinal é harmónico com uma periodicidade clara. Em outros casos, onde o sinal tem estrutura espectral menos pronunciada com máximos espectrais mais amplos, verificou-se que o uso de pequenos intervalos conduz a uma melhor qualidade. Esta constatação conduz a uma melhoria adicional de acordo com a qual o tamanho do intervalo se encontra adaptado de acordo com as propriedades do sinal. Uma forma de realização é a utilização de um detector de tonalidade ou periodicidade. Se este detector identificar o sinal como tonal, o parâmetro δ que controla o tamanho do intervalo é definido como um valor relativamente grande. Caso contrário, o parâmetro δ é definido com valores relativamente pequenos.

Com base no acima mencionado, os métodos de ocultação de perda de trama de áudio compreendem os seguintes passos:

1. Analisar um segmento do sinal disponível, sintetizado anteriormente para se obter as frequências sinusoidais f_k constituintes de um modelo sinusoidal, utilizando opcionalmente uma estimativa melhorada de frequência.
2. Extrair uma trama de protótipo y_{-1} do sinal previamente sintetizado disponível e calcular o DFT dessa trama.

3. Calcular o deslocamento de fase θ_k para cada sinusóide k em resposta à frequência sinusoidal f_k e o avanço de tempo n_{-1} entre a trama do protótipo e a trama de substituição. Opcionalmente, neste passo, o tamanho do intervalo M pode ter sido adaptado em resposta à tonalidade do sinal de áudio.

4. Para cada sinusóide k que avança a fase da trama DFT do protótipo com θ_k selectivamente para os índices DFT relacionados com uma proximidade em torno da frequência sinusóide f_k .

5. Calcular a DFT inverso do espectro obtido no passo 4.

Análise e detecção da propriedade de perda de sinal

Os métodos descritos acima são baseados no pressuposto de que as propriedades do sinal de áudio não mudam significativamente durante a curta duração da trama de sinal recebida anteriormente e reconstruída e uma trama perdida. Nesse caso é uma escolha muito boa reter o espectro de magnitude da trama reconstruída previamente e evoluir as fases dos componentes principais sinusoidais detectados no sinal reconstruído anteriormente. Existem no entanto casos em que essa suposição é errada que são, por

exemplo transientes com mudanças repentinas de energia ou mudanças repentinas espectrais.

Uma primeira forma de realização de um detector de transientes de acordo com a invenção pode, conseqüentemente, ser baseada em variações de energia dentro do sinal reconstruído anteriormente. Este método, ilustrado na figura 11, calcula a energia numa parte esquerda e uma parte direita de alguma trama de análise 113. A trama de análise pode ser idêntica à trama utilizada para a análise sinusoidal acima descrita. Uma parte (esquerda ou direita) da trama de análise pode ser a primeira ou, respectivamente, a última metade da trama de análise ou, por exemplo, o primeiro ou, respectivamente, o último quarto da trama de análise, 110. O respectivo cálculo de energia é feito pela soma dos quadrados das amostras nestas tramas parciais:

$$E_{left} = \sum_{n=0}^{N_{part}-1} y^2(n - n_{left}), \text{ e } E_{right} = \sum_{n=0}^{N_{part}-1} y^2(n - n_{right}).$$

Nisto $y(n)$ indica a trama de análise, n_{left} e n_{right} indica os respectivos índices iniciais das tramas parciais que são ambos de tamanho N_{part} .

Agora as energias esquerda e direita das tramas parciais são usadas para a detecção de uma descontinuidade de sinal. Isto é feito por meio do cálculo do rácio

$$R_{L/r} = \frac{E_{offr}}{E_{right}}$$

Pode ser detectada uma descontinuidade com queda repentina de energia (offset) se o rácio $R_{L/r}$, exceder algum limiar (por exemplo 10), 115. Da mesma forma pode ser detectada uma descontinuidade com aumento súbito de energia (onset), se o rácio $R_{L/r}$, for inferior a algum outro limiar (por exemplo 0,1), 117.

No contexto dos métodos de ocultação acima descritos, verificou-se que o rácio de energia acima definido pode, em muitos casos, ser um indicador muito insensível. Em particular nos sinais reais e especialmente música há casos em que um tom em alguma frequência surge de repente enquanto algum outro tom em alguma outra frequência pára de repente. Analisando uma tal trama de sinal com o rácio de energia acima definido iria em qualquer caso conduzir a um resultado de detecção errado para pelo menos um dos tons uma vez que este indicador não é sensível a diferentes frequências.

Uma solução para este problema encontra-se descrita na forma de realização seguinte. A detecção transitória é agora feita no plano da frequência do tempo. A trama de análise é novamente dividida numa trama parcial direita e esquerda, 110. Todavia agora, estas duas tramas parciais são (após janelamento adequado com por exemplo uma

janela de Hamming, 111) transformadas no domínio da frequência, por exemplo por meio de um N_{part} -ponto DFT, 112.

$$Y_{left}(m) = DFT\{y(n - n_{left})\}_{N_{part}}$$

e

$$Y_{right}(m) = DFT\{y(n - n_{right})\}_{N_{part}}, \text{ e } m = 0 \dots N_{part} - 1.$$

Agora a detecção transitória pode ser feita selectivamente na frequência para cada posição DFT com o índice m . Usando os poderes do espectro de magnitude da trama parcial esquerda e direita, para cada índice DFT m pode ser calculado 113 um respectivo rácio de energia como

$$R_{left}(m) = \frac{|Y_{left}(m)|^2}{|Y_{right}(m)|^2}$$

As experiências mostram que a detecção transitória selectiva da frequência com DFT é relativamente imprecisa devido a flutuações estatísticas (erros de estimativa). Verificou-se que a qualidade da operação é bastante melhorada ao fazer a detecção de transitório selectivo da frequência com base em bandas de frequência. Deixemos $I_k = [m_{k-1} + 1, \dots, m_k]$ especificar o k ésimo intervalo $k = 1 \dots K$, cobrindo as posições DFT de $m_{k-1} + 1$ para m_k , pois esses intervalos definem bandas de frequências K . A detecção transitória selectiva do grupo de frequências pode agora ser baseada no rácio relativo à banda entre as

respectivas energias de banda das tramas parciais esquerda e direita:

$$R_{i/c,band}(k) = \frac{\sum_{m \in I_k} |Y_{left}(m)|^2}{\sum_{m \in I_k} |Y_{right}(m)|^2}$$

É de notar que o intervalo $I_k = [m_{k-1} + 1, \dots, m_k]$ corresponde à banda de frequência $B_k = [\frac{m_{k-1}+1}{N_{part}} \cdot f_s, \dots, \frac{m_k}{N_{part}} \cdot f_s]$ em que f_s indica a frequência de amostragem de áudio.

O menor limite de banda de frequência inferior m_0 pode ser definido como 0 mas pode também ser definido como um índice de DFT que corresponde a uma frequência maior de modo a atenuar os erros de estimativa que crescem com frequências mais baixas. O maior limite de banda de frequência superior m_K pode ser definido como $\frac{N_{part}}{2}$ mas é de preferência escolhido para corresponder a alguma frequência inferior em que um transiente ainda tem um efeito audível significativo.

Uma escolha apropriada para estes tamanhos ou larguras de banda de frequência é para torná-los de tamanho igual, com por exemplo uma largura de vários 100 Hz. Uma outra forma preferida é fazer com que as larguras de banda de frequência sigam o tamanho das bandas críticas de audição humana, ou seja relacioná-las com a resolução da frequência do sistema auditivo. Isso significa aproximadamente fazer as larguras de banda de frequência

iguais para frequências até 1 kHz e aumentá-las exponencialmente acima de 1 kHz. O aumento exponencial significa, por exemplo, duplicar a largura de banda de frequência quando se incrementa o índice de banda k .

Tal como descrito na primeira forma de realização do detector de transientes que se baseou num rácio de energia de duas tramas parciais, qualquer dos rácios relacionados com as energias da banda ou energias da posição DFT de duas tramas parciais são comparados a certos limiares. É utilizado um respectivo limiar superior para (seleccionado pela frequência) detecção de deslocamento 115 e um respectivo limiar mais baixo para (seleccionado pela frequência) detecção de onset 117.

Um indicador adicional dependente do sinal de áudio que é adequado para uma adaptação do método de ocultação de perda de trama pode ser baseado nos parâmetros de codec transmitidos para o decodificador. Por exemplo, o codec pode ser um codec multimodo como ITU-T G.718. Tal codec pode utilizar modos particulares de codec para diferentes tipos de sinais e uma mudança do modo de codec numa trama pouco antes da perda de trama poder ser considerada como um indicador para um transiente.

Um outro indicador útil para a adaptação da ocultação da perda de trama é um parâmetro de codec relacionado com uma propriedade de vocalização e o sinal transmitido. A vocalização refere-se a voz altamente

periódica que é gerada por uma excitação glotal periódica do tracto vocal humano.

Um indicador adicional preferido é se o conteúdo do sinal for estimado como sendo música ou voz. Um tal indicador pode ser obtido a partir de um classificador de sinal que pode, tipicamente, fazer parte do codec. No caso de o codec efectuar uma tal classificação e tornar uma decisão de classificação correspondente disponível como um parâmetro de codificação para o descodificador, este parâmetro é de preferência utilizado como indicador do teor de sinal a ser usado para a adaptação do método de ocultação de perda de trama.

Um outro indicador que é de preferência utilizado para adaptação dos métodos de ocultação de perda de trama é o burstiness (ritmo de pico/ritmo médio) das perdas de trama. Burstiness de perdas de tramas significa que ocorrem diversas perdas de tramas numa fila, o que torna difícil para o método de ocultação de perda de trama utilizar partes válidas de sinal recentemente descodificado para o seu funcionamento. Um indicador do estado da técnica é o número n_{burst} das perdas de trama observadas numa fila. Este contador é incrementado com um em cada perda de trama e reposto a zero após a recepção de uma trama válida. Este indicador também é usado no contexto do presente exemplo de formas de realização da invenção.

Adaptação do método de ocultação de perda de

tramas

No caso dos passos executados acima indicarem uma condição que sugere uma adaptação da operação de ocultação de perda de trama, o cálculo do espectro da trama de substituição é modificado.

Embora o cálculo original do espectro da trama de substituição seja feito de acordo com a expressão $Z(m) = Y(m) \cdot e^{j\theta_k}$, agora é introduzida uma adaptação que modifica tanto magnitude como a fase. A magnitude é modificada por meio de dimensionamento com dois factores $\alpha(m)$ e $\beta(m)$ e a fase é modificada com um componente aditivo de fase $\mathcal{G}(m)$. Isto conduz ao seguinte cálculo modificado da trama de substituição:

$$Z(m) = \alpha(m) \cdot \beta(m) \cdot Y(m) \cdot e^{j(\theta_k + \mathcal{G}(m))}.$$

É para ser notado que os métodos de ocultação originais perda de trama (não-adaptados) são utilizados se $\alpha(m) = 1$, $\beta(m) = 1$, e $\mathcal{G}(m) = 0$. Estes valores respectivos são, portanto, o padrão.

O objectivo geral com a introdução de adaptações de magnitude é evitar perturbações audíveis do método de ocultação de perda de trama. Tais perturbações podem ser sons musicais ou tonais ou sons estranhos decorrentes de repetições de sons transitórios. Tais perturbações iriam por sua vez conduzir a degradações de qualidade, sendo que

é objectivo evitar as mesmas através das adaptações descritas. Um modo adequado para tais adaptações é o de modificar o espectro de magnitude da trama de substituição a um grau adequado.

A figura 12 ilustra uma forma de realização da modificação do método de ocultação. A adaptação da magnitude, 123 , é de preferência feita se o contador de perda de rajada n_{burst} exceder algum limiar thr_{burst} , por exemplo $thr_{burst} = 3$, 121 . Nesse caso é utilizado um valor menor do que 1 para o factor de atenuação, por exemplo, $\alpha(m) = 0,1$.

Verificou-se no entanto que é benéfico realizar a atenuação com um grau que aumenta gradualmente. Uma forma de realização preferida que realiza isto é definir um parâmetro logarítmico que especifica um aumento logarítmico na atenuação por trama, att_per_frame . Depois, no caso do contador de rajadas exceder o limiar, é calculado o factor de atenuação que aumenta gradualmente em

$$\alpha(m) = 10^{c \cdot att_per_frame \cdot (n_{burst} - thr_{burst})}$$

Aqui, a constante c é simplesmente uma constante de escalonamento que permite especificar o parâmetro att_per_frame , por exemplo em decibéis (dB).

Uma adaptação preferida adicional é feita em resposta ao indicador se o sinal é estimado como sendo de

música ou voz. Para conteúdo de música, em comparação com o conteúdo de voz é preferível aumentar o limiar thr_{burst} e diminuir a atenuação por trama. Isto é equivalente a executar a adaptação do método de ocultação de perda de trama com um menor grau. O antecedente deste tipo de adaptação é que a música é geralmente menos sensível a rajadas de perda mais longas do que a voz. Assim, o original, ou seja, o método de ocultação de perda de trama não modificada é ainda preferível para este caso, pelo menos para um número maior de perdas de tramas numa fila.

Uma outra adaptação do método de ocultação em relação ao factor de atenuação de magnitude é de preferência feita no caso de ter sido detectado um transiente com base em que o indicador de $R//r, band(k)$ ou, alternativamente, $R//r(m)$ ou $R//r$ tiverem passado um limiar, 122. Nesse caso, uma acção de adaptação adequada, 125, consiste em modificar o segundo factor de atenuação de magnitude, $\beta(m)$ de tal modo que a atenuação total é controlada pelo produto de dois factores $\alpha(m)$.

$\beta(m)$ é definido em resposta a um transiente indicado. Caso seja detectado um deslocamento será de preferência escolhido o factor $\beta(m)$ para reflectir a diminuição de energia do deslocamento. Uma escolha adequada é definir $\beta(m)$ para a mudança detectada de ganho:

$$\beta(m) = \sqrt{R_{thr,band}(k)}, \text{ para } m \in I_k, k = 1 \dots K.$$

Caso seja detectado um onset verificou-se ser vantajoso limitar o aumento de energia na trama de substituição. Nesse caso o factor pode ser definido para um valor fixo de, por exemplo, 1, o que significa que não existe qualquer atenuação mas também não existe qualquer amplificação.

No exemplo acima, é de notar que o factor de atenuação de amplitude é de preferência aplicado selectivamente na frequência, isto é com factores calculados individualmente para cada banda de frequência. Caso a abordagem de banda não for utilizada, podem ainda ser obtidos de um modo semelhante os factores de atenuação de amplitude correspondentes. $\beta(m)$ pode então ser definido individualmente para cada posição DFT caso seja utilizada a detecção transitória selectivamente na frequência ao nível da posição DFT. Ou, caso não seja utilizada qualquer detecção transitória selectivamente na frequência $\beta(m)$ pode ser globalmente idêntico para todos os m .

Uma outra adaptação preferida do factor de atenuação de amplitude é feita em conjunto com uma modificação da fase por meio do componente de fase adicional $\mathcal{G}(m)$ 127. No caso de para um determinado m for utilizada uma tal modificação de fase, o factor de atenuação $\beta(m)$ é ainda mais reduzido. De preferência, é tomado em consideração mesmo o grau de modificação de fase. Se a modificação de fase é apenas moderada, $\beta(m)$ só é ligeiramente reduzido, enquanto que a modificação de fase é

forte, $\beta(m)$ é reduzida a um grau maior.

O objectivo geral com a introdução de adaptações de fase é evitar tonalidade ou periodicidade do sinal demasiado fortes nas tramas de substituição geradas, que por sua vez levariam a degradação de qualidade. Um modo adequado para tais adaptações é o de randomizar ou aplicar ruído aleatório à fase a um grau adequado.

Uma tal aplicação de ruído aleatório à fase é realizada se o componente de fase adicional $\mathcal{G}(m)$ for definido como um valor aleatório dimensionado com algum factor de controlo: $\mathcal{G}(m) = a(m) \cdot \text{rand}(\bullet)$.

O valor aleatório obtido pela função $\text{rand}(\bullet)$ é, por exemplo, gerado por algum gerador de número pseudo-aleatório. É aqui assumido que proporciona um número aleatório no intervalo $[0, 2\pi]$.

O factor de dimensionamento $a(m)$ na equação acima controla o grau pelo qual à fase inicial θ_k é aplicado um ruído aleatório. As seguintes formas de realização solucionam a adaptação de fase por meio de controlo deste factor de dimensionamento. O controlo do factor de dimensionamento é feito de um modo análogo ao controlo dos factores de modificação de magnitude acima descritos.

De acordo com um primeiro factor de dimensionamento de forma de realização, $a(m)$ é adaptado em

resposta ao contador de perda de rajada. Se o contador de perda de rajada n_{burst} exceder algum limiar thr_{burst} , por exemplo $thr_{burst} = 3$, é utilizado um valor maior do que 0, por exemplo $a(m) = 0,2$.

Verificou-se no entanto que é benéfico realizar a aplicação de ruído aleatório com um grau que aumenta gradualmente. Uma forma de realização preferida que realiza isto é definir um parâmetro que especifica um aumento na aplicação de ruído aleatório por trama, $dith_increase_per_frame$. Depois, no caso do contador de rajadas exceder o limiar é calculado o factor de aplicação de ruído aleatório que aumenta gradualmente em

$$a(m) = dith_increase_per_frame \cdot (n_{burst} - thr_{burst}).$$

É de notar, na fórmula acima, que $a(m)$ tem de ser limitado a um valor máximo de 1 para o qual a aplicação de ruído aleatório de fase completa é alcançada.

É de notar que o valor limiar de perda de rajada thr_{burst} utilizado para iniciar a aplicação de ruído aleatório de fase pode ser o mesmo limiar que aquele utilizado para atenuação de magnitude. No entanto, pode ser obtida melhor qualidade ajustando estes limiares para valores óptimos individualmente, o que geralmente significa que estes limiares podem ser diferentes.

Uma adaptação preferida adicional é feita em

resposta ao indicador se o sinal é estimado como sendo de música ou voz. Para conteúdo de música, em comparação com o conteúdo de voz é preferível aumentar o limiar thr_{burst} o que significa que a aplicação de ruído aleatório na fase para música, quando comparado com voz é feita somente no caso de mais tramas perdidas numa fila. Isto é equivalente a executar a adaptação do método de ocultação de perda de trama para música com um menor grau. O antecedente deste tipo de adaptação é que a música é geralmente menos sensível a rajadas de perda mais longas do que a voz. Assim, o original, ou seja, método de ocultação de perda de trama não modificada é ainda preferível para este caso, pelo menos para um número maior de perdas de tramas numa fila.

Uma outra forma de realização preferida é para adaptar a aplicação de ruído aleatório da fase em resposta a um transiente detectado. Nesse caso, pode ser usado um grau mais forte de aplicação de ruído aleatório de fase para as posições m DFT para as quais um transiente é indicado quer para cada posição, as posições DFT da banda de frequência correspondente ou de toda a trama.

Parte dos esquemas descritos consignam otimização do método de ocultação de perda de tramas para sinais harmônicos e particularmente para discurso expresso.

No caso de não serem realizados métodos que utilizam uma estimativa melhorada de frequência tal como

descrito acima, uma outra possibilidade para o método de ocultação de perda de trama para otimizar a qualidade de sinais de discurso expresso é a de mudar para algum outro método de ocultação de perda de trama que se encontra especificamente projectado e optimizado para a voz em vez de para sinais de áudio gerais que contêm música e voz. Nesse caso, o indicador de que o sinal compreende um sinal de discurso expresso é usado para seleccionar um outro esquema de ocultação de perda de trama optimizado para a voz em vez dos esquemas descritos acima.

As formas de realização aplicam-se a um controlador num descodificador, tal como ilustrado na figura 13. A figura 13 diagrama de blocos esquemático de um descodificador de acordo com as formas de realização. O descodificador 130 compreende uma unidade de entrada 132 configurada para receber um sinal de áudio codificado. A figura ilustra a ocultação perda de trama por uma unidade lógica de ocultação de perda de trama 134, que indica que o descodificador está configurado para implementar uma ocultação de uma trama de áudio perdida, de acordo com as formas de realização acima descritas. Além disso, o descodificador compreende um controlador 136 para implementar as formas de realização descritas acima. O controlador 136 encontra-se configurado para detectar condições nas propriedades do sinal de áudio recebido anteriormente e reconstruído ou nas propriedades estatísticas das perdas de trama observadas para as quais a substituição de uma trama perdida de acordo com os métodos

descritos proporciona qualidade relativamente reduzida. Caso uma tal condição seja detectada, o controlador 136 encontra-se configurado para modificar o elemento dos métodos de ocultação de acordo com os quais o espectro de trama de substituição é calculado por $Z(m) = Y(m) \cdot e^{j\theta_k}$, ajustando selectivamente as fases ou as magnitudes de espectro. A detecção pode ser realizada por uma unidade de detecção 146 e a modificação pode ser realizada por uma unidade modificadora 148 tal como ilustrado na figura 14.

O descodificador com as suas unidades de inclusão pode ser implementado em hardware. Existem várias variantes de elementos de circuito que podem ser utilizados e combinados para atingir as funções das unidades do descodificador. Tais variantes estão abrangidas pelas formas de realização. Os exemplos particulares de implementação de hardware do descodificador é a implementação em hardware de processador de sinais digitais (DSP) e tecnologia de circuito integrado, incluindo ambos circuito electrónico de uso geral e circuitos de aplicação específica.

O descodificador 150 aqui descrito pode alternativamente ser implementado, por exemplo, tal como ilustrado na figura 15, ou seja por um ou mais de um processador 154 e software adequado 155 com armazenamento ou memória apropriados 156, por conseguinte a fim de reconstituir o sinal de áudio que inclui a realização de ocultação de perda de trama de áudio de acordo com as

formas de realização aqui descritas, tal como mostrado na figura 13. O sinal de áudio codificado que entra é recebido por uma entrada (IN) 152, à qual o processador 154 e a memória 156 se encontram ligados. O sinal de áudio decodificado e reconstruído obtido a partir do software é emitido da saída (OUT) 158.

A tecnologia descrita acima destina-se a ser utilizada, por exemplo, num receptor, o qual pode ser utilizado num dispositivo móvel (por exemplo, telefone móvel, portátil) ou um dispositivo estacionário, tal como um computador pessoal.

Deverá ser entendido que a escolha de unidades ou módulos que interagem, assim como a nomenclatura das unidades são apenas com a finalidade de exemplificar, e podem ser configuradas em várias formas alternativas de modo a serem passíveis de executar as acções de procedimento sugeridas.

Deverá também notar-se que as unidades ou módulos descritos na presente memória descritiva devem ser considerados como entidades lógicas e não com necessidade como entidades físicas separadas. Faz-se observar que o âmbito da tecnologia aqui descrita abrange totalmente outras formas de realização que podem tornar-se óbvias para os técnicos, e que o âmbito da presente memória descritiva não é portanto para ser limitado.

A referência a um elemento no singular não se destina a significar "um e apenas um" a menos que explicitamente assim indicado, mas sim "um ou mais". Todos os equivalentes estruturais e funcionais aos elementos das formas de realização acima descritas que são conhecidos dos técnicos são aqui expressamente incorporados como referência e destinam-se a ser englobados por este meio. Além disso, não é necessário a um dispositivo ou método tratar cada e todos os problemas que se pretende resolver com a tecnologia aqui descrita, para que possa ser englobado pela presente.

Na descrição anterior, para efeitos de explicação e não limitação, são estabelecidos pormenores específicos tais como arquitecturas particulares, interfaces, técnicas, etc., de modo a proporcionar uma compreensão profunda da tecnologia descrita. No entanto, será evidente para os técnicos que a tecnologia descrita pode ser praticada em outras formas de realização e/ou combinações de formas de realização que variam destes pormenores específicos. Quer dizer, os técnicos serão capazes de conceber várias formas de realização que, embora não explicitamente descritas ou apresentadas aqui, realizam os princípios da tecnologia descrita. Em alguns casos encontram-se omitidas descrições pormenorizadas de dispositivos bem conhecidos, circuitos, e métodos, de modo a não obscurecer a descrição da tecnologia descrita com pormenores desnecessários. Todas as declarações da presente, princípios recitados, aspectos e formas de realização da tecnologia descrita, bem como

exemplos específicos da mesma, servem para abranger ambos os equivalentes estruturais e funcionais da mesma. Adicionalmente, pretende-se que tais equivalentes incluam ambos os equivalentes actualmente conhecidos, bem como equivalentes desenvolvidos no futuro, por exemplo, quaisquer elementos desenvolvidos que desempenhem a mesma função, independentemente da estrutura.

Deste modo, por exemplo, será apreciado pelos técnicos que as figuras na presente podem representar vistas conceituais de circuito ilustrativo ou outras unidades funcionais que realizam os princípios da tecnologia, e/ou vários processos que podem ser substancialmente representados em suportes legíveis por computador e executados por um computador ou processador, mesmo que tal computador ou processador possa não estar explicitamente apresentado nas figuras.

As funções dos vários elementos, incluindo os blocos funcionais podem ser proporcionadas através do uso de hardware, tal como hardware de circuito e/ou hardware passível de executar software na forma de instruções codificadas armazenadas num meio legível por computador. Assim, tais funções e blocos funcionais ilustrados são para ser entendidos como sendo ou implementados por hardware e/ou implementados por computador, e deste modo implementados através de máquina.

As formas de realização descritas acima deverão

ser entendidas como alguns exemplos ilustrativos da presente invenção. Será entendido pelos técnicos que várias modificações, combinações e alterações podem ser feitas às formas de realização sem fugir do escopo da presente invenção. Em particular, diferentes soluções parciais nas diferentes formas de realização podem ser combinadas noutras configurações, onde tecnicamente possível.

Lisboa, 25 de Novembro de 2016

REIVINDICAÇÕES

1. Método para o controlo de um método de ocultação para uma trama perdida de áudio de um sinal de áudio recebido, caracterizado por compreender:

- detectar (101, 122) numa propriedade de um sinal de áudio previamente recebido e reconstruído uma condição transitória que poderia levar a uma qualidade de reconstrução subóptima, quando é utilizado um método de ocultação original para criar uma trama de substituição; e
- modificar (102, 125) o método de ocultação original ajustando selectivamente uma magnitude de espectro de um espectro de trama de substituição, quando a condição transitória é detectada;
- detectar (101, 122) adicionalmente numa propriedade estatística de perdas de trama observadas uma segunda condição que poderia levar a uma qualidade de reconstrução subóptima, quando o método de ocultação original é utilizado para criar a trama de substituição; e
- modificar (102, 123, 127) adicionalmente o método de ocultação original ajustando selectivamente a magnitude de espectro do espectro de trama de substituição, quando é detectada a segunda condição.

2. Método de acordo com a reivindicação 1, caracterizado por o método de ocultação original

compreender:

- extrair um segmento de um sinal de áudio recebido previamente ou reconstituído, em que o referido segmento é usado como uma trama protótipo;
- aplicar um modelo sinusoidal à trama protótipo para obter frequências sinusoidais do modelo sinusoidal; e
- sinusóides obtidas evoluídas no tempo para criar a trama de substituição.

3. Método de acordo com a reivindicação 2, caracterizado por a evolução no tempo compreender avançar a fase de coeficientes espectrais relacionados com as sinusóides obtidas (k) de θ_k , e em que o cálculo do espectro de trama de substituição é executado de acordo com a expressão $Z(m) = Y(m) \cdot e^{i\theta_k}$, em que $Y(m)$ é uma representação do domínio de frequência da trama protótipo.

4. Método de acordo com qualquer das reivindicações 1 a 3, caracterizado por a condição transiente compreender uma deslocação detectada.

5. Método de acordo com qualquer das reivindicações 1 a 4, caracterizado por a detecção transiente ser realizada selectivamente na frequência com base numa banda de frequência.

6. Método de acordo com as reivindicações 4 ou 5, caracterizado por o ajuste selectivo da magnitude do espectro da trama de substituição ser realizado

selectivamente na banda de frequência em resposta a um transiente detectado na banda de frequência.

7. Método de acordo com qualquer das reivindicações anteriores, caracterizado por a segunda condição ser uma ocorrência de várias perdas seguidas de trama.

8. Método de acordo com a reivindicação 7, caracterizado por a magnitude do espectro ser ajustada em resposta a várias perdas seguidas de tramas ao executar a atenuação com um grau de aumento gradual.

9. Método de acordo com qualquer das reivindicações anteriores, caracterizado por o método de ocultação inicial ser adicionalmente modificado por selectivamente ajustar uma fase do espectro de trama de substituição, quando a segunda condição é detectada.

10. Método de acordo com a reivindicação 9, caracterizado por o ajuste da fase do espectro de trama de substituição compreender randomizar ou aplicar ruído aleatório num espectro de fase.

11. Método de acordo com a reivindicação 10, caracterizado por o espectro de fase ser ajustado executando a aplicação de ruído aleatório com um grau que aumenta gradualmente.

12. Dispositivo caracterizado por compreender:
um processador (154), e
uma memória (156) que armazena instruções (155) que,
quando executadas pelo processador, fazem o dispositivo

- detectar numa propriedade de um sinal de áudio previamente recebido e reconstruído uma condição transitória que poderia levar a uma qualidade de reconstrução subóptima, quando é utilizado um método de ocultação original para criar uma trama de substituição;
- modificar o método de ocultação original, quando a condição transitória é detectada, ajustando selectivamente uma magnitude de espectro de um espectro de trama de substituição;
- detectar adicionalmente numa propriedade estatística de perdas de trama observadas uma segunda condição que poderia levar a uma qualidade de reconstrução subóptima, quando o método de ocultação original é utilizado para criar a trama de substituição; e
- modificar adicionalmente o método de ocultação original, quando é detectada a segunda condição, ajustando selectivamente a magnitude do espectro de trama de substituição.

13. Dispositivo de acordo com a reivindicação 12, caracterizado por durante a criação da trama de substituição utilizando o método de ocultação inicial o dispositivo é levado a:

- extrair um segmento de um sinal de áudio recebido

previamente ou reconstituído,
em que o referido segmento é usado como uma trama
protótipo;

- aplicar um modelo sinusoidal à trama protótipo para obter frequências sinusoidais do modelo sinusoidal; e
- sinusóides obtidas evoluídas no tempo para criar a trama de substituição.

14. Dispositivo de acordo com a reivindicação 13, caracterizado por a evolução no tempo ser executada avançando a fase de coeficientes espectrais relacionados com as sinusóides obtidas (k) de θ_k , e em que o cálculo do espectro de trama de substituição é executado de acordo com a expressão $Z(m) = Y(m) \cdot e^{j\theta_k}$, em que $Y(m)$ é uma representação do domínio da frequência da trama protótipo.

15. Dispositivo de acordo com as reivindicações 12 a 14 compreendendo adicionalmente um detector transitório, caracterizado por o detector transitório se encontrar configurado para executar uma detecção transitória selectiva na frequência com base em bandas de frequência.

16. Dispositivo de acordo com a reivindicação 15, caracterizado por o ajuste selectivo da magnitude do espectro da trama de substituição ser realizado na banda de frequência selectivamente em resposta a um transiente detectado na banda de frequência.

17. Dispositivo de acordo com qualquer das reivindicações 12 a 16, caracterizado por a segunda condição ser uma ocorrência de várias perdas seguidas de trama.

18. Dispositivo de acordo com a reivindicação 17, caracterizado por a magnitude do espectro ser ajustada em resposta a várias perdas seguidas de tramas ao executar a atenuação com grau de aumento gradual.

19. Dispositivo de acordo com qualquer das reivindicações 12 a 18, caracterizado por o dispositivo estar configurado para modificar adicionalmente o método de ocultação original, quando é detectada a segunda condição, ajustando selectivamente uma fase do espectro de trama de substituição.

20. Dispositivo de acordo com a reivindicação 12, caracterizado por o dispositivo ser um descodificador num dispositivo móvel.

21. Programa de computador (155) caracterizado por compreender unidades de código legíveis em computador que quando executados num dispositivo fazem o dispositivo:

- detectar (101) numa propriedade de um sinal de áudio previamente recebido e reconstruído uma condição transitória que poderia levar a uma qualidade de reconstrução subóptima, quando é utilizado um método de ocultação original para criar uma trama de

substituição; e

- modificar (102) o método de ocultação original, quando a condição transitória é detectada, ajustando selectivamente uma magnitude de espectro de um espectro de trama de substituição;

- detectar adicionalmente numa propriedade estatística de perdas de trama observadas uma segunda condição que poderia levar a uma qualidade de reconstrução subóptima, quando o método de ocultação original é utilizado para criar a trama de substituição; e

- modificar adicionalmente o método de ocultação original, quando é detectada a segunda condição, ajustando selectivamente a magnitude do espectro de trama de substituição.

22. Produto de programa informático (156), caracterizado por compreender um suporte legível por computador e um programa de computador (155) de acordo com a reivindicação 22 armazenado num suporte legível por computador.

Lisboa, 25 de Novembro de 2016

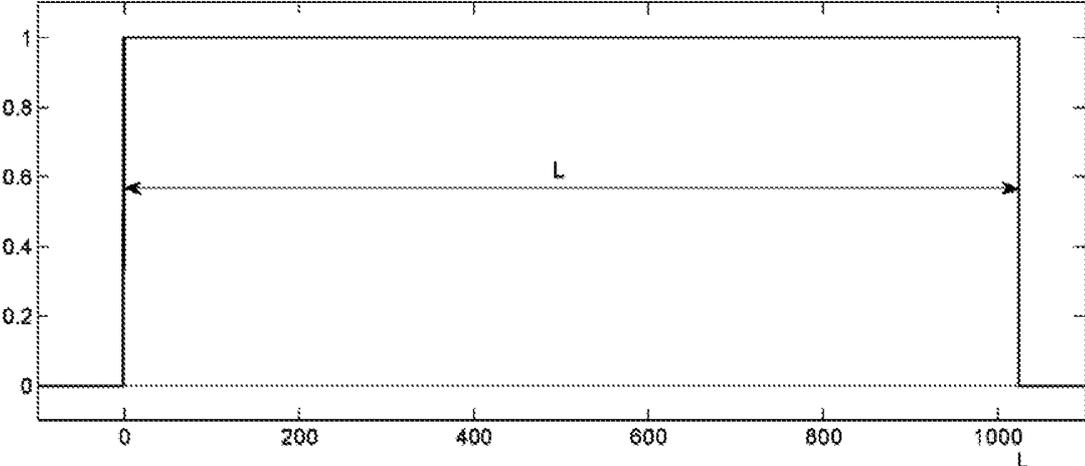


FIG. 1

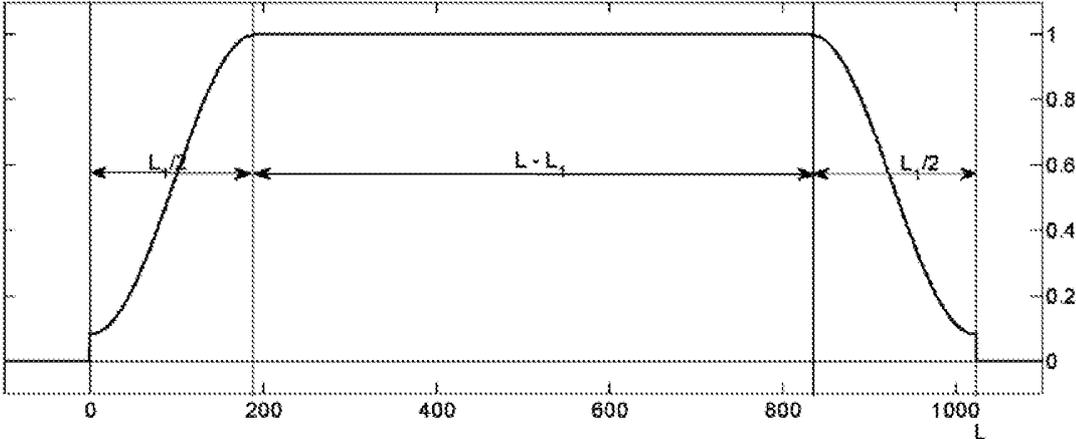


FIG. 2

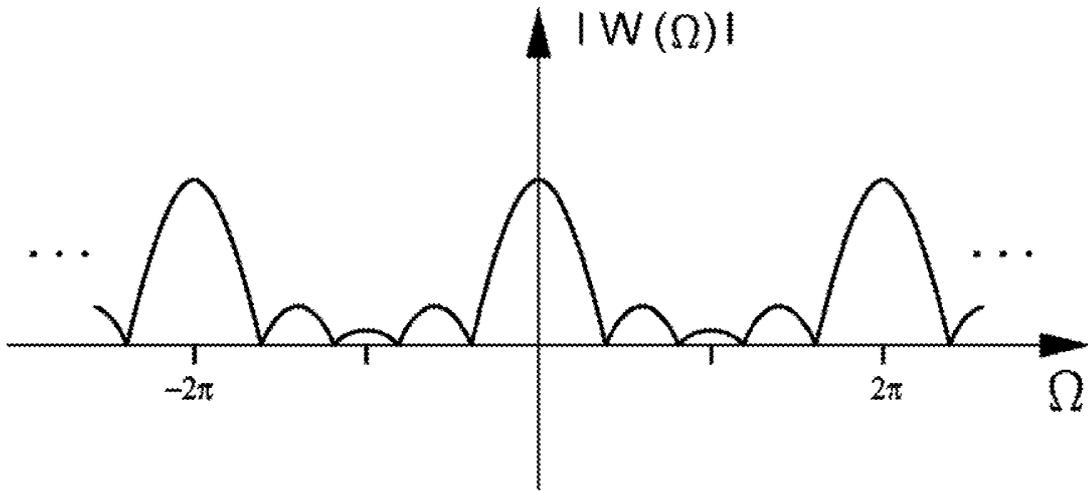


FIG. 3

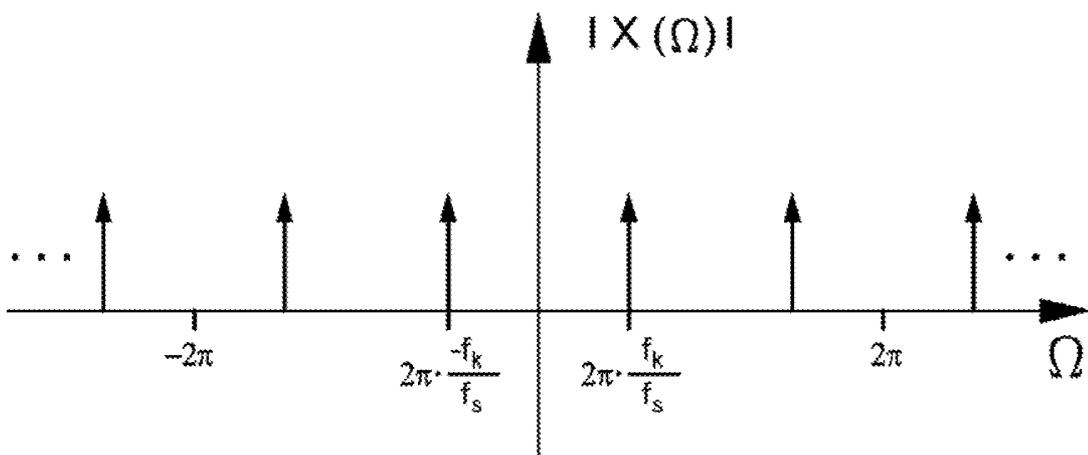


FIG. 4

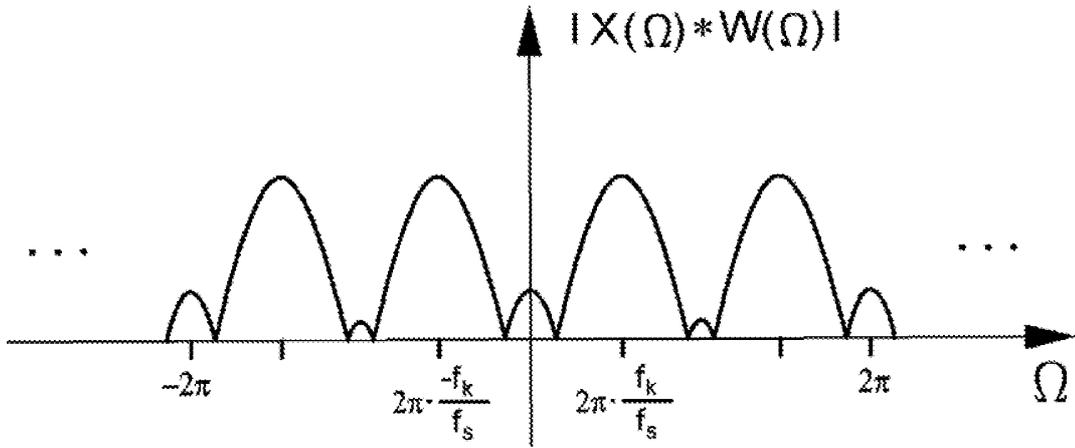


FIG. 5

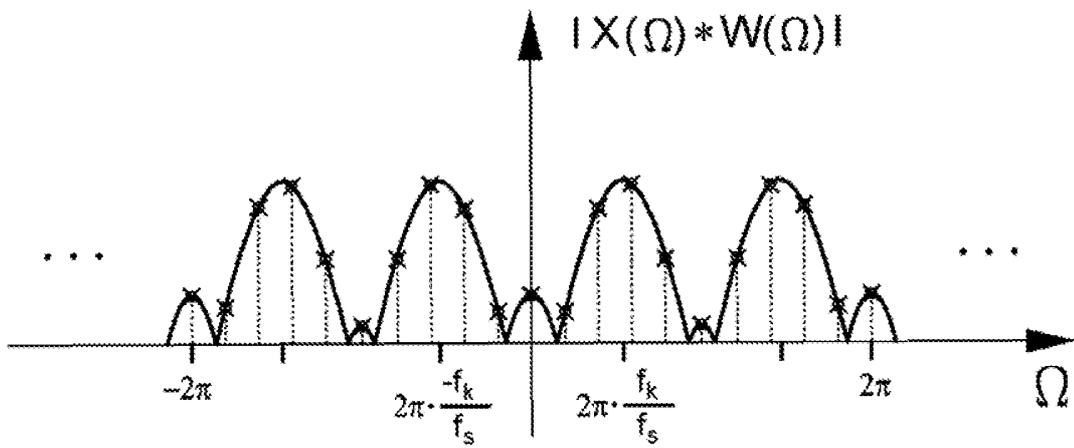
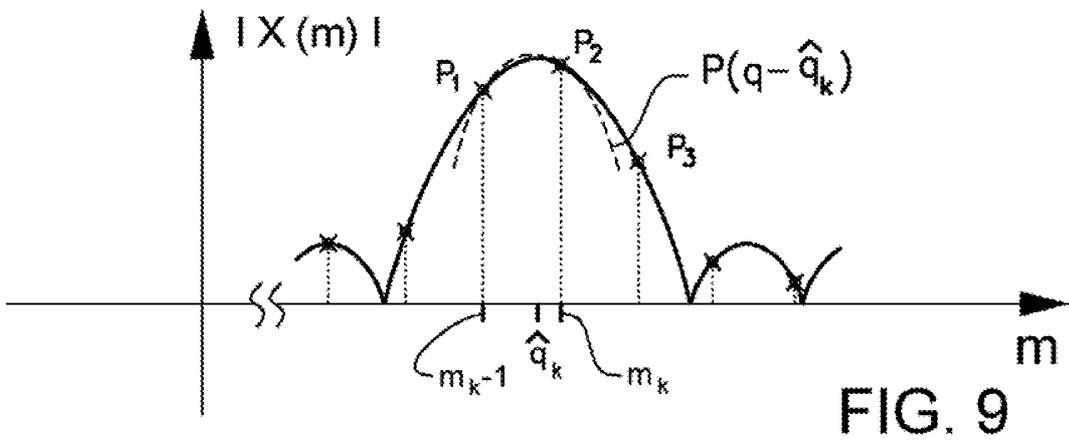
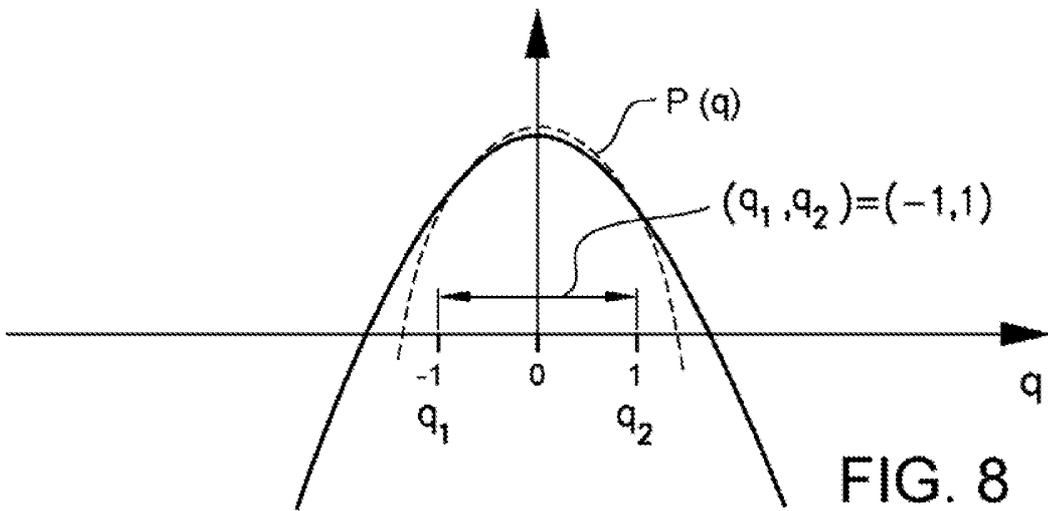
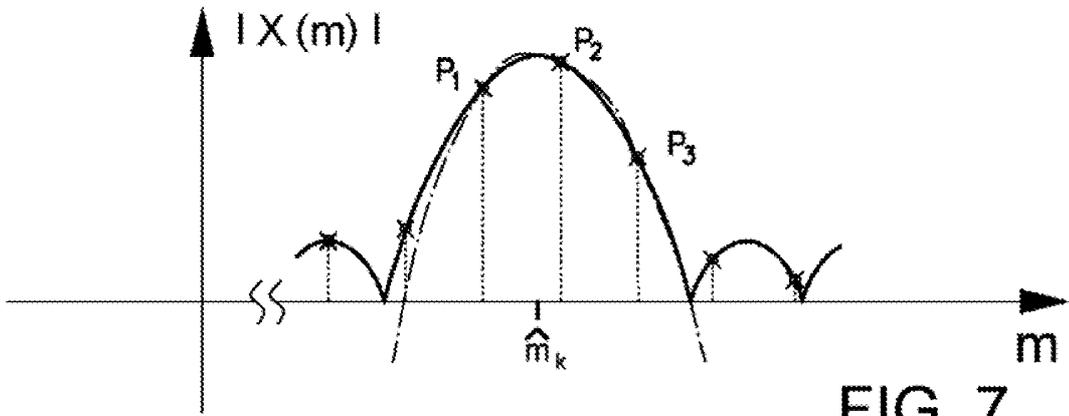


FIG. 6



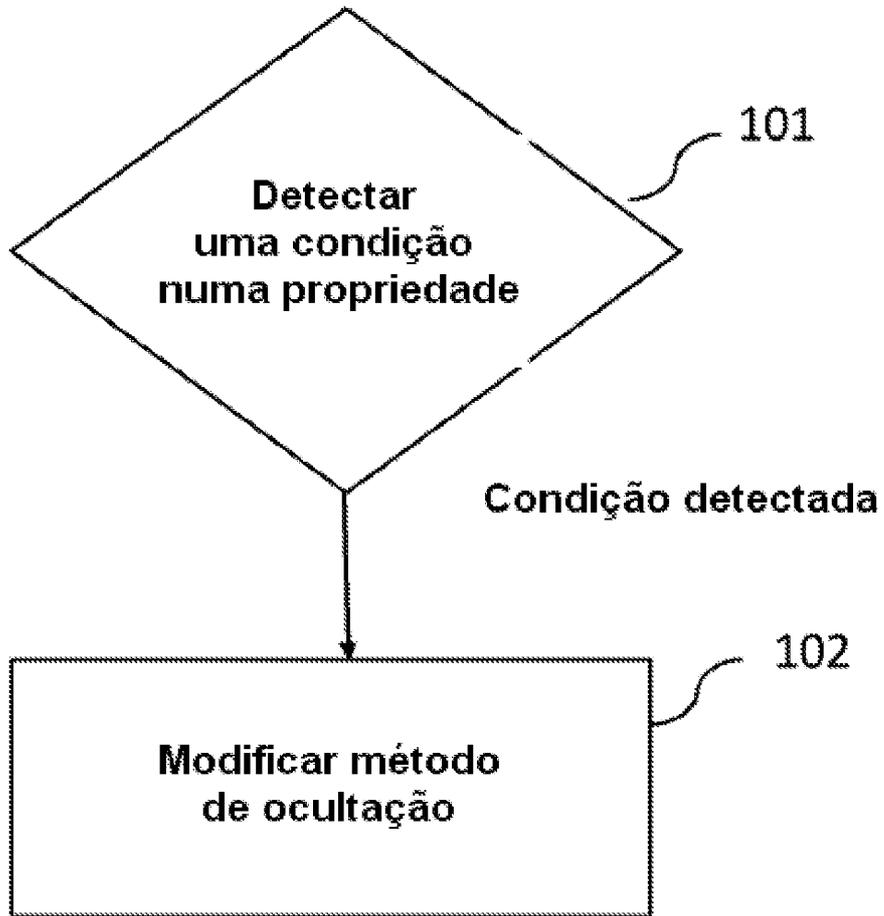


FIG. 10

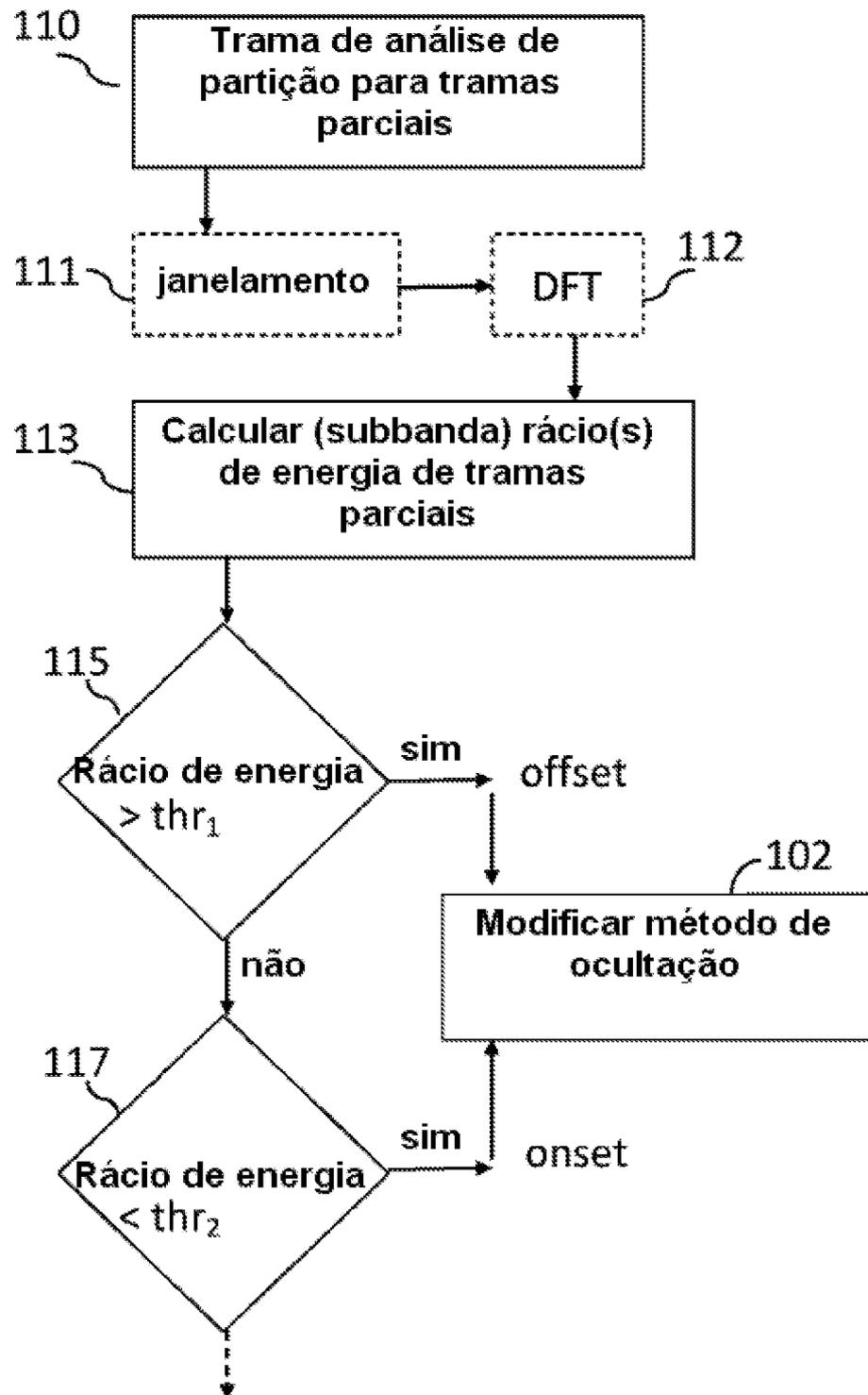


FIG. 11

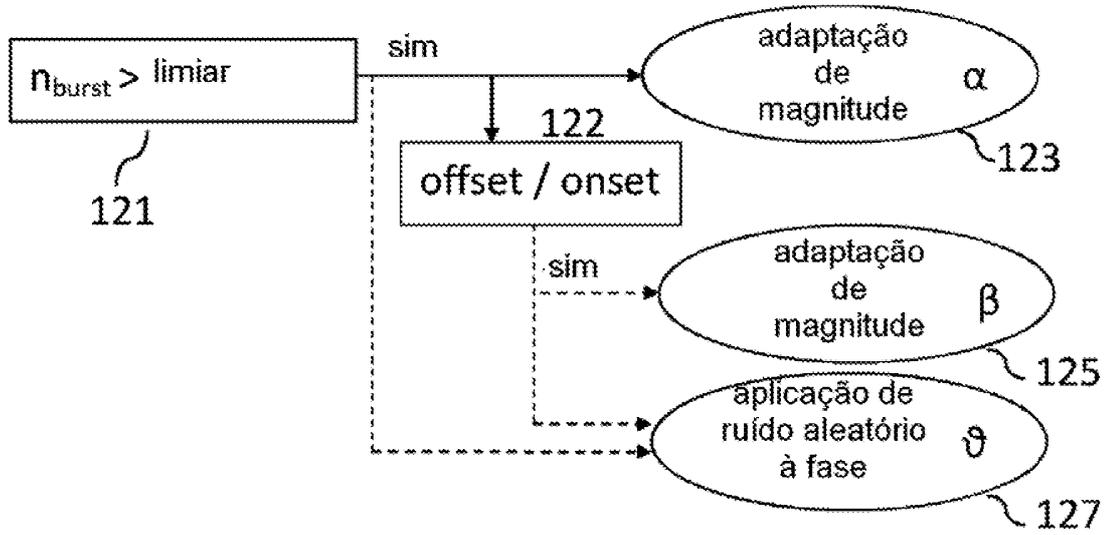


FIG. 12

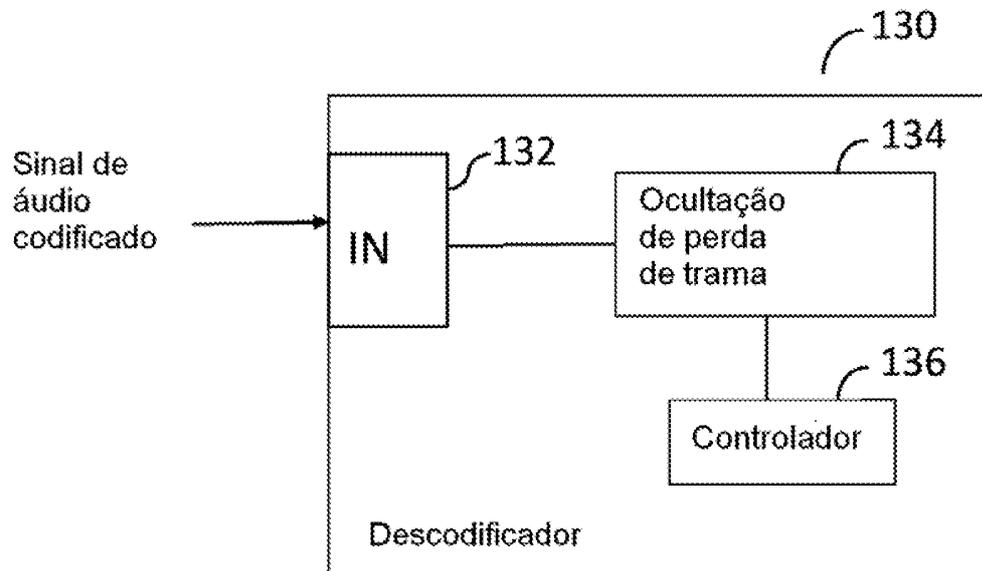


FIG. 13

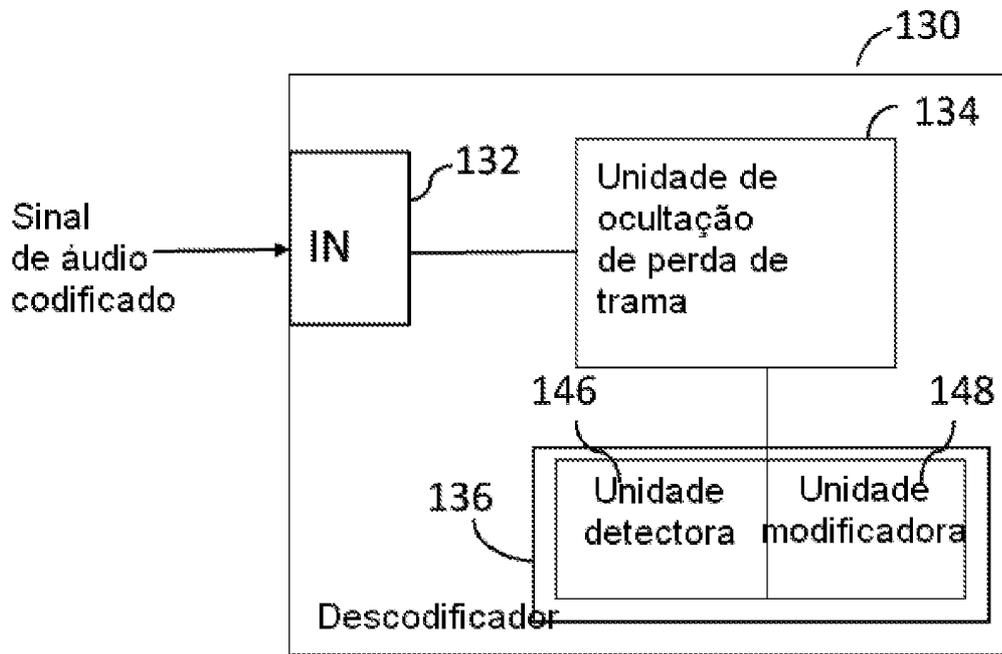


FIG. 14

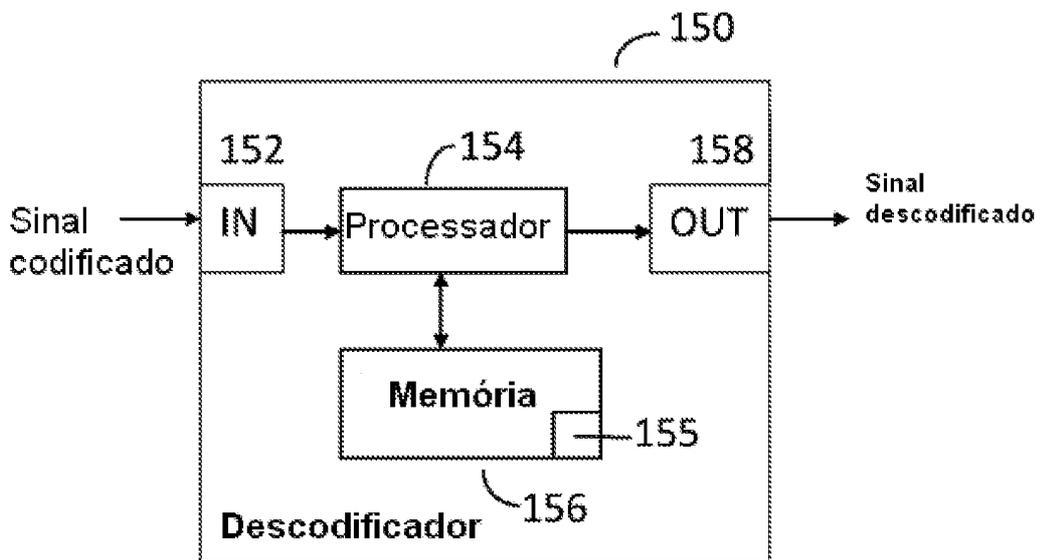


FIG. 15

REFERÊNCIAS CITADAS NA DESCRIÇÃO

Esta lista de referências citadas pelo requerente é apenas para conveniência do leitor. A mesma não faz parte do documento da patente europeia. Ainda que tenha sido tomado o devido cuidado ao compilar as referências, podem não estar excluídos erros ou omissões e o IEP declina quaisquer responsabilidades a esse respeito.

Documentos de patentes citados na Descrição

- **US 2004122680 A**