



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

(11) PI 1008520-3 B1



(22) Data do Depósito: 02/02/2010

(45) Data de Concessão: 05/05/2020

(54) Título: MÉTODO E DISPOSITIVO DE EXTENSÃO DE LARGURA DE BANDA

(51) Int.Cl.: G10L 21/02; G10L 19/06; G10L 19/08.

(30) Prioridade Unionista: 04/02/2009 US 12/365,457.

(73) Titular(es): GOOGLE TECHNOLOGY HOLDINGS LLC.

(72) Inventor(es): TENKASI V. RAMABADRAN; MARK A. JASIUK.

(86) Pedido PCT: PCT US2010022879 de 02/02/2010

(87) Publicação PCT: WO 2010/091013 de 12/08/2010

(85) Data do Início da Fase Nacional: 04/08/2011

(57) Resumo: MÉTODO DE EXTENSÃO DE LARGURA DE BANDA E APARELHO PARA UM CODIFICADOR DE TRANSFORMADA DISCRETA DE COSSENO MODIFICADA Um método incluiu definir uma banda de transição para um sinal que tem um espectro dentro de uma primeira banda de frequência, onde a banda de transição é definida como uma porção da primeira banda de frequência e está localizada próxima a uma banda de frequência adjacente que é adjacente à primeira banda de frequência. O método analisa a banda de transição para obter um envelope espectral de banda de transição e um espectro de excitação de banda de transição; estima um envelope espectral de banda de frequência adjacente; gera um espectro de excitação de banda de frequência adjacente por repetição periódica de ao menos uma parte do espectro de excitação de banda de transição com um período de repetição determinado por uma frequência de timbre do sinal; e combina o envelope espectral de banda de frequência adjacente e o espectro de excitação de banda de frequência adjacente para obter um O espectro de sinal de banda de frequência adjacente . Uma lógica de processamento de sinal para realizar o método é também revelada.

MÉTODO E DISPOSITIVO DE EXTENSÃO DE LARGURA DE BANDA

REFERÊNCIA CRUZADA A PEDIDOS RELACIONADOS

A presente revelação está relacionada a: Pedido de
5 Patente N° U.S. 11/946.978, Número de Protocolo:
CML04909EV, depositada em 29 de Novembro de 2007,
denominada MÉTODO E APARELHO PARA FACILITAR A PROVISÃO E
USO DE UM VALOR DE ENERGIA PARA DETERMINAR UM FORMATO DE
ENVELOPE ESPECTRAL PARA CONTEÚDO DE LARGURA DE BANDA FORA
10 DE SINAL, Pedido de Patente N° 12/024.620, Número de
Protocolo: CML04911EV, depositada em 1 de Fevereiro de
2008, denominada MÉTODO E APARELHO PARA ESTIMAR ENERGIA DE
BANDA ALTA EM UM SISTEMA DE EXTENSÃO DE LARGURA DE BANDA;
Pedido de Patente N° U.S. 12/027.571, Número de Protocolo:
15 CML06672AUD, depositada em 7 de Fevereiro de 2008,
denominada MÉTODO E APARELHO PARA ESTIMAR ENERGIA DE BANDA
ALTA EM UM SISTEMA DE EXTENSÃO DE LARGURA DE BANDA; todos
os quais são incorporados a título de referência ao
presente.

20 CAMPO DA REVELAÇÃO

A presente revelação está relacionada a codificadores
de áudio e renderizar conteúdo audível e mais
particularmente a técnicas de extensão de largura de banda
para codificadores de áudio.

25 ANTECEDENTES

A fala telefônica por telefones móveis tem usualmente
utilizado somente uma porção do espectro de som audível,
por exemplo, a fala de banda estreita dentro do espectro de
áudio de 300 a 3400 Hz. Comparada à fala normal, tal fala
30 de banda estreita tem uma qualidade abafada e

inteligibilidade reduzida. Sendo assim, vários métodos para estender a largura de banda da saída de codificadores de fala, chamada de "extensão de largura de banda" ou "BWE," podem ser aplicados para artificialmente melhorar a qualidade de som percebida da saída do codificador.

Embora esquemas de BWE possam ser paramétricos ou não paramétricos, a maioria dos esquemas de BWE conhecidos é paramétrica. Os parâmetros que surgem do modelo de filtro por fonte de produção de fala onde o sinal de fala é considerado como um sinal de fonte de excitação que foi acusticamente filtrado pelo trato vocal. O trato vocal pode ser moldado por um filtro de todos os polos, por exemplo, com uso de técnicas de previsão linear (LP) para computar os coeficientes de filtro. Os coeficientes de LP efetivamente criam parâmetros para as informações de envelope espectral de fala. Outros métodos paramétricos utilizam frequências espectrais de linha (LSF), coeficientes cepstrais de mel-frequência (MFCC) e amostras de envelope de log-espectral (LES) para moldar o envelope espectral de fala.

Muitos codificadores de áudio/fala atuais utilizam a representação de Transformada Discreta de Cosseno Modificada (MDCT) do sinal de entrada e então são necessário métodos de BWE que poderiam ser aplicados à MDCT baseados nos codificadores de áudio/fala.

BREVE DESCRIÇÃO DOS DESENHOS

A FIGURA 1 é um diagrama de um sinal de áudio que tem uma banda de transição próxima à banda de frequência alta que é usada nas modalidades para estimar o espectro de sinal de banda de frequência alta.

A FIGURA 2 é um fluxograma de operação básica de um codificador em concordância com as modalidades.

A FIGURA 3 é um fluxograma que mostra detalhes a mais da operação de um codificador em concordância com as modalidades.

A FIGURA 4 é um diagrama em bloco de um dispositivo de comunicação que emprega um codificador em concordância com as modalidades.

A FIGURA 5 é um diagrama em bloco de um codificador em concordância com as modalidades.

A FIGURA 6 é um diagrama em bloco de um codificador em concordância com uma modalidade.

DESCRIÇÃO DETALHADA

A presente revelação fornece um método para extensão de largura de banda em um codificador e inclui definir uma banda de transição para um sinal que tem um espectro dentro de uma primeira banda de frequência, onde a banda de transição é definida como uma porção da primeira banda de frequência e está localizada próxima a uma banda de frequência adjacente que é adjacente à primeira banda de frequência. O método analisa a banda de transição para obter um envelope espectral de banda de transição e um espectro de excitação de banda de transição; estima um envelope espectral de banda de frequência adjacente; gera um espectro de excitação de banda de frequência adjacente por repetição periódica de ao menos uma parte do espectro de excitação de banda de transição com uma frequência de repetição determinada por uma frequência de timbre do sinal; e combina o envelope espectral de banda de frequência adjacente e o espectro de excitação de banda de

frequência adjacente para obter um espectro de sinal de banda de frequência adjacente. Uma lógica de processamento de sinal para realizar o método é também revelada.

Em concordância com as modalidades, a extensão de largura de banda pode ser implantada, com uso de ao menos dos coeficientes de MDCT quantificados gerados por um codificador de fala ou áudio que molda uma banda de frequência, tal como 4 a 7 kHz, para prever coeficientes de MDCT que moldam outra banda de frequência, tal como 7 a 14 kHz.

Voltando-se agora aos desenhos em que numerais similares representam componentes similares, a FIGURA 1 é um gráfico 100, que não está em escala, que representa um sinal de áudio 101 sobre um espectro audível 102 que abrange de 0 a Y kHz. O sinal 101 tem uma porção de banda baixa 104 e uma porção de banda alta 105 que não é reproduzida como parte da fala de banda baixa. Em concordância com as modalidades, uma banda de transição 103 é selecionada e utilizada para estimar a porção de banda alta 105. O sinal de entrada pode ser obtido de várias maneiras. Por exemplo, o sinal 101 pode ser fala recebido por um canal sem fio digital de um sistema de comunicação, enviado a uma estação móvel. O sinal 101 pode ser também obtido a partir de memória, por exemplo, em um dispositivo de reprodução de áudio a partir de um arquivo de áudio armazenado.

A FIGURA 2 ilustra a operação básica de um codificador em concordância com as modalidades. Em 201, uma banda de transição 103 é definida dentro de uma primeira banda de frequência 104 do sinal 101. A banda de transição 103 é

definida como uma porção da primeira banda de frequência e está localizada próxima à banda de frequência adjacente (tal como a porção de banda alta 105). Em 203, a banda de transição 103 é analisada para obter dados espectrais de banda de transição, e, em 205, o espectro de sinal de banda de frequência adjacente é gerado com uso dos dados espectrais de banda de transição.

A FIGURA 3 ilustra detalhes a mais da operação para uma modalidade. Em 301, uma banda de transição é definida similarmente a 201. Em 303, a banda de transição é analisada para se obter dados espectrais de banda de transição que incluem o envelope espectral de banda de transição e um espectro de excitação de banda de transição. Em 305, o envelope espectral de banda de frequência adjacente é estimado. O espectro de excitação de banda de frequência adjacente é então gerado, conforme mostrado em 307, por repetição periódica de ao menos uma parte do espectro de excitação de banda de transição com uma frequência de repetição determinada por uma frequência de timbre do sinal de entrada. Conforme mostrado em 309, o envelope espectral de banda de frequência adjacente e o espectro de excitação de banda de frequência adjacente podem ser combinados para se obter um espectro de sinal para a banda de frequência adjacente.

A FIGURA 4 é um diagrama em bloco que ilustra os componentes de um dispositivo eletrônico 400 em concordância com as modalidades. O dispositivo eletrônico pode ser uma estação móvel, um computador portátil, um assistente digital pessoal (PDA), um rádio, um reproduutor de áudio (tal como um reproduutor de MP3) ou qualquer outro

dispositivo adequado que possa receber um sinal de áudio, seja por meio de transmissão com fio ou sem fio, e decodificar o sinal de áudio com uso dos métodos e aparelhos das modalidades reveladas no presente. O dispositivo eletrônico 400 incluirá uma porção de entrada 403 onde um sinal de áudio é fornecido a uma lógica de processamento de sinal 405 em concordância com as modalidades.

Deve-se entender que a FIGURA 4, assim como a FIGURA 5 e a FIGURA 6, são para fins ilustrativos somente, para o fim de ilustrar a uma pessoa versada na técnica a lógica necessária para fazer e usar as modalidades descritas no presente. Sendo assim, as Figuras no presente não pretendem ser diagramas esquemáticos completos de todos os componentes necessários para for, por exemplo, implantar um dispositivo eletrônico, mas ao invés disso mostrar somente aquele que é necessário para facilitar um entendimento, por uma pessoa versada na técnica, de como fazer e usar as modalidades reveladas no presente. Sendo assim, deve-se também entender que várias disposições de lógica e quaisquer componentes internos mostrados e qualquer conectividade correspondente entre estes, podem ser utilizadas e que tais disposições e conectividade correspondente permaneceriam em concordância com as modalidades reveladas no presente.

O termo "lógica" conforme usado no presente inclui software e/ou firmware que executem um ou mais processadores programáveis, ASICs, DSPs, lógica conectada diretamente ou combinações destes. Sendo assim, em concordância com as modalidades, qualquer lógica descrita,

incluindo, por exemplo, lógica de processamento de sinal 405, pode ser implantada em qualquer maneira apropriada e poderia permanecer em concordância com as modalidades reveladas no presente.

5 O dispositivo eletrônico 400 pode incluir um receptor, ou transceptor, porção de extremidade frontal 401 e qualquer antena ou antenas necessárias para receber um sinal. Sendo assim, o receptor 401 e/ou lógica de entrada 403, individualmente ou em combinação, incluirão toda a
10 lógica necessária para fornecer sinais de áudio apropriados para a lógica de processamento de sinal 405 adequada para processamento adicional pela lógica de processamento de sinal 405. A lógica de processamento de sinal 405 pode também incluir um livro de código ou livros de código 407 e
15 tabelas de pesquisa 409 em algumas modalidades. As tabelas de pesquisa 409 podem ser tabelas de pesquisa de envelope espectral.

A FIGURA 5 fornece detalhes a mais da lógica de processamento de sinal 405. A lógica de processamento de
20 sinal 405 inclui uma lógica de estimação e controle 500, que determina um conjunto de coeficientes de MDCT para representar a porção de banda alta de um sinal de áudio. Uma MDCT Inversa, IMDCT 501, é usada para converter o sinal para o domínio de tempo que é então combinado com a porção
25 de banda baixa do sinal de áudio 503 por meio de uma operação de soma 505 para obter um sinal de áudio de largura de banda estendida. O sinal de áudio de largura de banda estendida é então emitido para uma lógica de saída de áudio (não mostrada).

30 Os detalhes adicionais de algumas modalidades são

ilustrados pela FIGURA 6, embora algumas lógicas ilustradas possam não estar, e não precisem estar, presentes em todas as modalidades. Para fins de ilustração, a seguir, é considera-se que a banda baixa cobre a faixa de 50 Hz a 7 kHz (nominalmente chamada como o espectro de áudio/fala de banda larga) e considera-se que a banda alta cobre a faixa de 7 kHz a 14 kHz. A combinação de bandas baixas e altas, isto é, a faixa de 50 Hz a 14 kHz, é nominalmente chamada como o espectro de áudio/fala de superbanda larga. É claro que outras escolhas para as bandas baixas e altas são possíveis e permaneceriam em concordância com as modalidades. Além disso, para fins de ilustração, o bloco de entrada 403, que é parte do codificador de linha de base, é mostrado para fornecer os seguintes sinais: i) o sinal de áudio/fala de banda larga decodificado S_{wb} , ii) os coeficientes de MDCT que correspondem a ao menos a banda de transição, e iii) a frequência de timbre 606 ou o atraso/período de timbre correspondente. O bloco de entrada 403, em algumas modalidades, pode fornecer somente o sinal de áudio/fala de banda larga decodificado e os outros sinais podem ser, neste caso, derivados deste no decodificador. Conforme ilustrado na FIGURA 6, a partir do bloco de entrada 403, um conjunto de coeficientes de MDCT quantificado é selecionado em 601 para representar uma banda de transição. Por exemplo, a banda de frequência de 4 a 7 kHz pode ser utilizada como uma banda de transição; no entanto, outras porções espectrais podem ser usadas e permaneceriam em concordância com as modalidades.

A seguir os coeficientes de MDCT da banda de transição selecionada são usados, juntamente com os parâmetros

-- computados a partir da fala/áudio de banda larga
-- decodificado (por exemplo, até 7 kHz), para gerar um
conjunto estimado de coeficientes de MDCT de forma a
especificar o conteúdo de sinal na banda adjacente, por
5 exemplo, de 7 a 14 kHz. Os coeficientes de MDCT da banda de
transição selecionada são assim fornecidos para a lógica de
análise de banda de transição 603 e a lógica estimadora de
energia de banda de transição 615. A energia nos
coeficientes de MDCT quantificados, que representa a banda
10 de transição, é computada pela lógica estimadora de energia
de banda de transição 615. A saída da lógica estimadora de
energia de banda de transição 615 é um valor de energia e é
proximamente relacionada a, embora não idêntica, à energia
na banda de transição do sinal de fala/áudio de banda larga
15 decodificado.

O valor de energia determinado em 615 é inserido no
previsor de energia de banda alta 611, que é um previsor de
energia não linear que computa a energia coeficientes de
MDCT que moldam a banda adjacente, por exemplo, a banda de
20 frequência de 7 a 14 kHz. Em algumas modalidades, para
melhorar o desempenho do previsor de energia de banda alta
611, o previsor de energia de banda alta 611 pode usar
cruzamentos por zero a partir da fala decodificada,
calculados pelo calculador de cruzamentos por zero 619, em
25 conjunto com o formato de envelope espectral da porção
espectral da banda de transição determinado pelo estimador
de formato de banda de transição 609. Dependendo do valor
do cruzamento por zero e do formato da banda de transição,
diferentes previsores não lineares são usados assim levando
30 a um desempenho de previsor acentuado. No desenho dos

previsores, uma grande base de dados de treinamento é primeiramente dividida em um número de partições com base no valor de cruzamento por zero e no formato da banda de transição e para cada uma das partições assim gerada, coeficientes de previsor separados são computados.

Especificamente, a saída do calculador de cruzamentos por zero 619 pode ser quantificada com uso de um quantificador escalar de 8 níveis que quantifica os cruzamentos por zero de quadro e, desta forma, o estimador de formato da banda de transição 609 pode ser um quantificador de vetor de envelope espectral de 8 formatos 8 (VQ) que classifica o formato de envelope espectral. Assim a cada quadro no máximo 64 (isto é, 8x8) previsores não lineares são fornecidos e um previsor que corresponde à partição selecionada é empregado naquele quadro. Na maioria das modalidades, menos de 64 previsores são usados, porque algumas das some 64 partições não estão atribuídas a um número suficiente de quadros a partir da base de dados de treinamento para garantir sua inclusão, e essas partições podem ser consequentemente unidas às partições próximas. Um previsor de energia separado (não mostrado), treinado por quadros de energia baixa, pode ser usado para tais quadros de baixa energia em concordância com as modalidades.

Para computar o envelope espectral que corresponde à banda de transição (4 a 7 kHz), os coeficientes de MDCT, que representam o sinal naquela banda, são primeiramente processados no bloco 603 por um operador de valor absoluto. Em seguida, os coeficientes de MDCT processados que são de valor zero são identificados e as magnitudes zeradas são substituídas por valores obtidos através de uma

interpolação linear entre as magnitudes de MDCT de valor não zero que se ligam, que foram escalonadas para baixo (por exemplo, por um fator de 5) antes de se aplicar o operador de interpolação linear. A eliminação de coeficientes de MDCT de valor zero, conforme descrito anteriormente, reduz a faixa dinâmica do espectro de magnitude de MDCT e melhora a eficiência de moldagem do envelope espectral computado a partir dos coeficientes de MDCT modificados.

Os coeficientes de MDCT modificados são então convertidos ao domínio dB, por meio do operador $20 \cdot \log_{10}(x)$ (não mostrado). Na banda de 7 a 8 kHz, o espectro dB é obtido por dobra espectral sobre um índice de frequência que corresponde a 7 kHz, para reduzir ainda mais a faixa dinâmica do envelope espectral a ser computado para a banda de frequência de 4 a 7 kHz. Uma Transformada de Fourier Discreta Inversa (IDFT) é a próxima aplicada ao espectro dB assim construído para a banda de frequência de 4 a 8 kHz, para computar os primeiros 8 coeficientes (pseudo-)cepstrais. O envelope espectral dB é então calculado realizando-se uma operação de Transformada de Fourier Discreta (DFT) mediante os coeficientes cepstrais.

O envelope espectral de MDCT de banda de transição resultante é usado de duas formas. Primeiro, forma uma entrada para o quantificador de vetor de envelope espectral de banda de transição, isto é, para o estimador de formato da banda de transição 609, que retorna um índice do envelope espectral pré-armazenado (um dos 8) que está mais perto para o envelope espectral de entrada. Esse índice juntamente com um índice (um dos 8) retornados por um

quantificador escalar dos cruzamentos por zero computados a partir da fala decodificada, é usado para selecionar um dos no máximo 64 previsores de energia não linear, conforme previamente detalhado. Em segundo lugar, o envelope espectral computado é usado para planificar o envelope espectral dos coeficientes de MDCT da banda de transição. Uma forma na qual isto pode ser feito é dividir cada coeficiente de MDCT da banda de transição por seu valor de envelope espectral correspondente. A planificação pode ser também implantada no domínio de log, em cujo caso a operação de divisão é substituída por uma operação de subtração. Na última implantação, os signos de coeficiente de MDCT (ou polaridades) são salvos para posterior reestabelecimento, porque a conversão para o domínio de log requer saídas de valor positivo. Nas modalidades, a planificação é implantada no domínio de log.

Os coeficientes de MDCT de banda por transição planificados (que representam o espectro de excitação de MDCT de banda de transição) emitidos pelo bloco 603 são então usados para gerar os coeficientes de MDCT que moldam o sinal de excitação na banda de 7 a 14 kHz. Em uma modalidade a faixa de MDCT indica correspondente à banda de transição pode ser 160 a 279, assumindo-se que o índice de MDCT inicial seja de tamanho de quadro de 0 e 20 ms a amostragem de 32 kHz. Dados os coeficientes de MDCT de banda por transição planificados, os coeficientes de MDCT que representam a excitação para índices de 280 a 559 que correspondem à banda de 7 a 14 kHz são gerados, com uso do seguinte mapeamento:

$$\text{MDCT}_{\text{exc}}(i) = \text{MDCT}_{\text{exc}}(i-D), \quad i = 280, \dots, 559, \quad D \leq 120.$$

O valor de atraso de frequência D , para um dado quadro, é computado a partir do valor de atraso de preditor a longo prazo (LTP) para o último sub-quadro do quadro de 20 ms que é parte das informações transmitidas de codec de núcleo. A partir desse atraso de LTP decodificado, um valor de frequência de timbre estimado para o quadro é computado, e o maior múltiplo inteiro deste valor de frequência de timbre é identificado, para render um valor de atraso de frequência inteiro correspondente D (definido no domínio de índice de MDCT) que é menos ou igual a 120. Esta abordagem garante o reuso das informações de MDCT de banda de transição planejadas assim preservando a relação harmônica entre os coeficientes de MDCT na banda de 4 a 7 kHz e os coeficientes de MDCT que são estimados para a banda de 7 a 14 kHz. Alternativamente, os coeficientes de MDCT computados a partir de uma entrada de sequência de ruído branco podem ser usados para formar uma estimativa de coeficientes de MDCT planejados na banda de 7 a 14 kHz. De qualquer forma, uma estimativa dos coeficientes de MDCT representativos das informações de excitação na banda de 7 a 14 kHz é formada pelo gerador de excitação de banda alta 605.

O valor de energia previsto dos coeficientes de MDCT na banda de 7 a 14 kHz emitido pelo preditor de energia não linear pode ser adaptado pela lógica adaptadora de energia 617 com base nas características do sinal de banda larga decodificado para minimizar artefatos e acentuar a qualidade da fala de saída de largura de banda estendida. Para este fim, o adaptador de energia 617 recebe as seguintes entradas em adição ao valor de energia de banda

alta previsto: i) o desvio padrão σ do erro de previsão a partir do previsor de energia de banda alta 611, ii) o nível de voz v a partir do estimador de nível de voz 621, iii) a saída d do detector de ativação/explosão 623, e iv) a saída ss do detector de transmissão/estado estável 625.

Dado o valor de energia previsto e adaptado dos coeficientes de MDCT na banda de 7 a 14 kHz, o envelope espectral consistente com esse valor de energia é selecionado a partir de um livro de códigos 407. Tal livro de códigos de envelopes espectrais que molda os envelopes espectrais que caracterizam os coeficientes de MDCT na banda de 7 a 14 kHz e classificados de acordo com os valores de energia nessa banda é treinado fora de linha. O envelope que corresponde à classe de energia mais próxima do valor de energia previsto e adaptado é selecionado pelo seletor de envelope de banda alta 613.

O envelope espectral selecionado é fornecido pelo seletor de envelope de banda alta 613 para o gerador de MDCT de banda alta 607, e é então aplicado para formatar os coeficientes de MDCT que moldam a excitação planejados na banda de 7 a 14 kHz. Os coeficientes de MDCT formatados que correspondem à banda de 7 a 14 kHz que representa o espectro de MDCT de banda alta são em seguida aplicados a uma transformada inversa de cosseno modificada (IMDCT) 501, para formar um sinal de domínio de tempo que tem conteúdo na banda de 7 a 14 kHz. Este sinal é então combinado por, por exemplo, operação de soma 505, com o sinal de banda larga decodificado que tem conteúdo até 7 kHz, isto é, a porção de banda baixa 503, para formar o sinal de largura de banda estendida que contém informações até 14 kHz.

Por uma abordagem, o valor de energia previsto e adaptado anteriormente mencionado pode servir para facilitar acessar uma tabela de pesquisa 409 que contém uma pluralidade de formatos de envelope espectral candidatos correspondentes. Para suportar tal abordagem, este aparelho por também compreender, se desejado, uma ou mais tabelas de pesquisa 409 que são acopladas de forma operável à lógica de processamento de sinal 405. Assim configurada, a lógica de processamento de sinal 405 pode facilmente acessar as tabelas de pesquisa 409 conforme apropriado.

Deve-se entender que o processamento de sinal discutido anteriormente pode ser realizado por uma estação móvel em comunicação sem fio com uma estação base. Por exemplo, a estação base pode transmitir o sinal de áudio digital de banda larga ou banda estreita por meio convencional para a estação móvel. Uma vez recebido, a lógica de processamento de sinal dentro da estação móvel realiza as operações de requisito para gerar a versão estendida da largura de banda do sinal de áudio digital que é mais clara e mais audível agradando um usuário da estação móvel.

Adicionalmente em algumas modalidades, um estimador de nível de voz 621 pode ser usado em conjunto com o gerador de excitação de banda alta 605. Por exemplo, um nível de voz de 0, que indica fala não-vozeada, pode ser usado para determinar o uso de excitação de ruído. Similarmente, um nível de voz de 1 que indica fala vozeada, pode ser usado para determinar o uso de excitação de banda alta derivada da excitação de banda de transição descrita anteriormente. Quando o nível de voz está entre 0 e 1 indicando fala

vozeada misturada, várias excitações podem estar misturadas em proporção adequada conforme determinado pelo nível de voz e usado. A excitação de ruído pode ser função de ruído pseudo aleatória e, conforme descrito anteriormente, pode ser considerada como orifícios de preenchimento ou conexão no espectro baseado no nível de voz. Uma excitação de banda alta misturada é assim adequada para sons vozeados, não-vozeados e vozeados misturados.

A FIGURA 6 mostra a Lógica de Estimação e Controle 500 como compreendendo a lógica seletora de coeficiente de MDCT de banda de transição 601, a lógica de análise de banda de transição 603, o gerador de excitação de banda alta 605, o gerador de coeficiente de MDCT de banda alta 607, o estimador de formato da banda de transição 609, o previsor de energia de banda alta 611, o seletor de envelope de banda alta 613, o estimador de energia de banda de transição 615, o adaptador de energia 617, o calculador de cruzamentos por zero 619, o estimador de nível de voz 621, o detector de ativação/explosão 623, e detector de Transição/SS 625.

A entrada 403 fornece o sinal de fala/áudio de banda larga decodificado S_{wb} , os coeficientes de MDCT correspondentes a ao menos a banda de transição e a frequência de timbre (ou atraso) para cada quadro. A lógica seletora de MDCT de banda de transição 601 é parte do codificador de linha de base e fornece um conjunto de coeficientes de MDCT para a banda de transição para a lógica de análise de banda de transição 603 e para o estimador de energia de banda de transição 615.

Estimação de nível de voz: Para estimar o nível de

voz, um calculador de cruzamento por zero 619 pode calcular o número de cruzamentos por zero zc em cada quadro da fala de banda larga swb conforme segue:

$$zc = \frac{1}{2(N-1)} \sum_{n=0}^{N-2} |Sgn(s_{wb}(n)) - Sgn(s_{wb}(n+1))|$$

onde

$$Sgn(s_{wb}(n)) = \begin{cases} 1 & \text{if } s_{wb}(n) \geq 0 \\ -1 & \text{if } s_{wb}(n) < 0 \end{cases}$$

onde n é o índice de amostra e N é o tamanho de quadros em amostras. O tamanho de quadro e sobreposição por cento usados na Lógica de Estimação e Controle 500 são determinados pelo codificador de linha de base, por exemplo, $Nu = 640$ em frequência de amostragem de 32 kHz e 50% de sobreposição. O valor do parâmetro zc calculado conforme as faixas anteriores de 0 a 1. A partir do parâmetro zc , um estimador de nível de voz 621 pode estimar o nível de voz v conforme segue.

$$v = \begin{cases} 1 & \text{if } zc < ZC_{low} \\ 0 & \text{if } zc > ZC_{high} \\ 1 - \left[\frac{zc - ZC_{low}}{ZC_{high} - ZC_{low}} \right] & \text{de outra forma} \end{cases}$$

onde, ZC_{low} (baixo) e ZC_{high} (alto) representam os limites baixo e alto apropriadamente escolhidos respectivamente, por exemplo, $ZC_{baixo} = 0,125$ e $ZC_{alto} = 0,30$.

A fim de estimar a energia de banda alta, um estimador de energia de transição de banda 615 estima a energia de transição de banda a partir dos coeficientes de MDCT de banda de transição. A transição de banda é definida aqui como uma banda de frequência que está contida dentro da

banda larga e perto da banda alta, isto é, serve como uma transição para a banda alta, (que, neste exemplo ilustrativo é de cerca de 7000 a 14.000 Hz). Uma forma de calcular a energia de transição de banda E_{tb} é somar as energias dos componentes espectrais, isto é, coeficientes de MDCT, dentro da transição de banda.

A partir da energia de transição de banda E_{tb} em dB (decibéis), a energia de banda alta E_{hw} em dB é estimada conforme

$$E_{kbo} = \alpha E_{tb} + \beta$$

onde, os coeficientes α e β são selecionado para minimizar o erro quadrado de meio entre os valores verdadeiros e estimado da energia de banda alta por uma quantidade de quadros a partir de uma base de dados de áudio/fala de treinamento.

A precisão de estimação pode ser ainda mais acentuada exportando informações contextuais a partir de parâmetros de fala adicionais tal como o parâmetro de cruzamento por zero zc e o formato espectral de banda de transição como pode ser fornecido por um estimador de formato de banda de transição 609. O parâmetro de cruzamento por zero, conforme discutido anteriormente, é indicativo do nível de voz da fala. O estimador de formato da banda de transição 609 fornece uma representação de alta resolução do formado de envelope de banda de transição. Por exemplo, uma representação quantificada de vetor dos formatos de envelope espectral de banda de transição (em dB) pode ser usada. O livro de códigos de quantificador de vetor (VQ) consiste de 8 formatos chamados de parâmetros de formato de envelope espectral de banda de transição tbs que são

computados a partir de uma grande base de dados de treinamento. Um plano de parâmetro zc - tbs correspondente pode ser formado com uso dos parâmetros zc e tbs para alcançar desempenho melhorado. Conforme descrito anteriormente, o plano zc - tbs é dividido em 64 partições que correspondem a 8 níveis quantificados escalares de zc e a 8 formatos de tbs . Algumas das partições podem ser unidas com as partições próximas para falta de pontos de dados suficientes a partir da base de dados de treinamento. Para cada uma das partições remanescentes no plano zc - tbs , coeficientes previsores separados são computados.

O previsor de energia de banda alta 611 pode fornecer melhoria adicional na precisão de estimação usando forças maiores de E_{tb} para estimar E_{hb0} , por exemplo,

$$E_{hb0} = \alpha_4 E_{tb}^4 + \alpha_3 E_{tb}^3 + \alpha_2 E_{tb}^2 + \alpha_1 E_{tb} + \beta.$$

Neste caso, cinco coeficientes diferentes, a saber, α_4 , α_3 , α_2 , α_1 , e β , são selecionados para cada partição do plano de parâmetro zc - tbs . Como as equações anteriores para estimar E_{hb0} são não lineares, deve ser tomado cuidado especial para ajustar a energia de banda alta estimada conforme o nível de sinal de entrada, isto é, a energia, se altera. Uma forma de alcançar isto é estimar o nível de sinal de entrada em dB, ajustar o E_{tb} para cima ou para baixo para corresponder ao nível de sinal nominal, estimar o E_{hb1} , e ajustar o E_{hb0} para baixo ou para cima para corresponder ao nível de sinal atual.

A estimação da energia de banda alta é tendente a erros. Como a superestimação leva a artefatos, a energia de banda alta estimada é induzida a ser inferior por uma quantidade proporcional ao desvio padrão do erro de

estimação de UW. Isto é, a energia de banda alta é adaptada no adaptador de energia 617 como:

$$E_{hb1} = E_{hb0} - \lambda \cdot \sigma$$

onde, E_{hb1} é a energia de banda alta adaptada em dB,

5 E_{hb0} é a energia de banda alta estimada em dB, $\lambda > 0$ é um fator de proporcionalidade e σ é o desvio padrão do erro de estimação em dB. Assim, após determinar o nível de energia de banda alta estimada, o nível de energia de banda alta estimada é modificado com base em uma precisão de estimação da energia de banda alta estimada. Com referência à FIGURA 6, o previsor de energia de banda alta 611 adicionalmente determina uma medida de falta de confiabilidade na estimação do nível de energia de banda alta e o adaptador de energia 617 induz o nível de energia de banda alta 15 estimada a ser inferior por uma quantidade proporcional à medida de falta de confiabilidade. Em uma modalidade, a medida de falta de confiabilidade compreende um desvio padrão σ do erro no nível de energia de banda alta estimada. Outras medidas de falta de confiabilidade podem 20 ser da mesma forma empregadas sem se afastar do escopo das modalidades.

Por "induzir para baixo" a energia de banda alta estimada, a probabilidade (ou número de ocorrências) de superestimação de energia é reduzida, assim reduzindo o 25 número de artefatos. Além disso, a quantidade pela qual a energia de banda alta estimada é reduzida é proporcional a quão boa a estimativa é - uma estimativa mais confiável (isto é, baixo valor de σ) é reduzida por um menor quantidade que uma estimativa menos confiável. Enquanto 30 desenha o previsor de energia de banda alta 611, o valor de

σ correspondente a cada partição do plano de parâmetro zc -tbs é computado a partir da base de dados de fala de treinamento e armazenado para uso posterior em "indução para baixo" da energia de banda alta estimada. O valor de σ das (≤ 64) partições do plano de parâmetro zc -tbs, por exemplo, abrange de cerca de 4 dB a cerca de 8 dB com uma valor médio de cerca de 5,9 dB. Um valor adequado de λ para este previsor de energia de banda alta, por exemplo, é de 1,2.

Em uma abordagem de técnica anterior, a superestimação da energia de banda alta é tratada por uso de uma função de custo assimétrica que penaliza os erros superestimados mais que erros subestimados no desenho do previsor de energia de banda alta 611. Comparada a esta abordagem de técnica anterior, a abordagem de "indução para baixo" descrita no presente tem as seguintes vantagens: (A) O desenho do previsor de energia de banda alta 611 é mais simples porque é baseado na função de custo de "erro quadrado" simétrica padrão; (B) A "indução para baixo" é feita durante a fase operacional (e não implicitamente durante a fase de) e sendo assim a quantidade de "indução para baixo" pode ser facilmente controlada conforme desejado; e (C) A dependência da quantidade de "indução para baixo" para a confiabilidade da estimativa é explícita e direta (ao invés de implicitamente dependente da função de custo específica usada durante a fase de desenho).

Além de reduzir os artefatos devido a superestimação de energia, a abordagem de "indução para baixo" descrita anteriormente tem um benefício adicional para quadros vozeados - ou seja mascarando quaisquer erros na estimação

de formato de envelope espectral de banda alta e assim
reduzindo os artefatos "ruidosos" resultantes. No entanto,
para quadros não-vozeados, se a redução na energia de banda
alta estimada for muito alta, a fala de saída de largura de
5 banda estendida não soa mais como fala de banda superlarga.
Para contar isto, a energia de banda alta estimada é
además adaptada no adaptador de energia 617 dependendo de
seu nível de voz conforme

$$E_{hb2} = E_{hb1} + (1-v) \cdot \delta_1 + v \cdot \delta_2$$

10 onde, E_{hb2} é a energia de banda alta adaptada por nível
de voz em dB, v é o nível de voz na faixa de 0 para fala
não-vozeada a 1 para fala vozeada, e δ_1 e δ_2 ($\delta_1 > \delta_2$) são
constantes em dB. A escolha de δ_1 e δ_2 depende do valor de
 λ usado para a "indução para baixo" e é determinado
15 empiricamente para render a fala de saída de melhor som.
Por exemplo, quando λ é escolhido como 1,2, δ_1 e δ_2 podem
ser escolhidos como 3,0 e -3,0 respectivamente. Note-se que
outras escolhas para o valor de λ podem resultar em
diferentes escolhas de δ_1 e δ_2 - os valores de δ_1 e δ_2
20 podem ser ambos positivos ou negativos ou de sinais
opostos. O nível de energia aumentado enfatiza tal fala na
saída de largura de banda estendida comparada a entrada de
banda larga e também ajuda a selecionar um formato de
envelope espectral mais apropriado para tais segmentos
25 surdos.

Com referência à FIGURA 6, o estimador de nível de voz
621 emite um nível de voz para o adaptador de energia 617
que modifica ainda mais o nível de energia de banda alta
estimada com base nas características de sinal de banda
30 larga modificando ainda mais o nível de energia de banda

alta estimada com base em um nível de voz. A modificação adicional pode compreender reduzir o nível de energia de banda alta para fala substancialmente vozeada e/ou aumentar o nível de energia de banda alta para fala substancialmente surda.

Enquanto o previsor de energia de banda alta seguido pelo adaptador de energia trabalha bastante bem para a maioria dos quadros, ocasionalmente há quadros para os quais a energia de banda alta é grosseiramente sub ou superestimada. Algumas modalidades podem então fornecer para tais erros de estimação e, ao menos parcialmente, corrigi-los com uso de uma lógica atenuadora de faixa de energia (não mostrada) que compreende um filtro de atenuação. Assim a etapa de modificar o nível de energia de banda alta estimada com base nas características de sinal de banda larga pode compreender atenuar o nível de energia de banda alta estimada (que foi previamente modificado conforme descrito anteriormente com base no desvio padrão da estimação σ e o nível de voz v), essencialmente reduzir uma diferença de energia entre quadros consecutivos.

Por exemplo, a energia de banda alta adaptada por nível de voz E_{hb2} pode ser atenuada com uso de filtro de média de 3 pontos conforme

$$E_{hb3} = [E_{hb2}(k-1) + E_{hb2}(k) + E_{hb2}(k+1)] / 3$$

onde, E_{hb3} é a estimativa atenuada e k é o índice de quadro. A atenuação reduz a diferença de energia entre quadros consecutivos, especialmente quando uma estimativa é uma "exceção", isto é, a estimativa de energia de banda alta de um quadro é muito alta ou muito baixa comparada às estimativas dos quadros vizinhos. Assim, a atenuação ajuda

a reduzir a quantidade de artefatos na fala de largura de banda estendida de saída. O filtro de média de 3 pontos introduz um atraso de um quadro. Outros tipos de filtros com ou sem atraso podem ser também desenhados para atenuar a faixa de energia.

O valor de energia atenuada E_{hb3} pode ser ainda mais adaptado pelo adaptador de energia 617 para obter a estimativa de energia de banda alta final adaptada E_{hb} . Esta adaptação pode envolver tanto diminuir ou aumentar o valor de energia atenuado com base na saída de parâmetro ss pelo detector de estado estável/transmissão 625 e/ou a saída de parâmetro d pelo detector de ativação/explosão 623. Assim, a etapa de modificar o nível de energia de banda alta estimada com base nas características de sinal de banda larga pode incluir a etapa de modificar o nível de energia de banda alta estimada (ou o nível previamente modificado de energia de banda alta estimada) com base em se ou não um quadro está em estado estável ou transiente. Isto pode incluir reduzir o nível de energia de banda alta para quadros transientes e/ou aumentar o nível de energia de banda alta para quadros em estado estável, e pode ademais incluir modificar o nível de energia de banda alta estimada com base em uma ocorrência de uma ativação/explosão. Por uma abordagem, adaptar o valor de energia de banda alta altera não somente o nível de energia, mas também o formato de envelope espectral já que a seleção do espectro de banda alta pode estar presa à energia estimada.

Um quadro é definido como um quadro em estado estável se tem energia suficiente (isto é, é um quadro de fala e

não um quadro de silêncio) e está perto de cada um de seus quadros vizinhos tanto em um sentido espectral como em termos de energia. Dois quadros podem ser considerados espectralmente próximos se a distância de Itakura entre os dois quadros estiver abaixo de um limite especificado. Outros tipos de medidas de distância espectral podem ser também usados. Dois quadros são considerados próximos em termos de energia se a diferença nas energias de banda larga dos dois quadros estiver abaixo de um limite.

Qualquer quadro que não seja um quadro em estado estável é considerado um quadro de transição. Um quadro em estado estável é capaz de mascarar erros em estimação de energia de banda alta muito melhor que quadros transientes. Desta forma, a energia de banda alta estimada de um quadro é adaptada com base no parâmetro ss , isto é, dependendo de se é um quadro em estado estável ($ss = 1$) ou quadro de transição ($ss = 0$) conforme

$$E_{hb4} = \begin{cases} E_{hb3} + \mu_1 & \text{para quadros em estado estável } ss \\ \min(E_{hb3} - \mu_2, E_{hb2}) & \text{para quadros de transição } ss \end{cases}$$

20

onde, $\mu_2 > \mu_1 \geq 0$, são empiricamente escolhidos constantes em dB para atingir boa qualidade de fala de saída. Os valores de μ_1 e μ_2 dependem da escolha da constante de proporcionalidade λ usada para a "indução para baixo". Por exemplo, quando λ é escolhido como 1,2, δ_1 como 3,0 e δ_2 como -3,0, μ_1 e μ_2 podem ser escolhidos como 1,5 e 6,0 respectivamente. Note-se que neste exemplo aumentou-se levemente a energia de banda alta estimada para quadros em estado estável e diminui-se significativamente mais para quadro de transição. Note-se que outras escolhas para os

30

valores de λ , δ_1 e δ_2 podem resultar em diferentes escolhas para μ_1 e μ_2 - os valores de μ_1 e μ_2 podem ser ambos positivos ou negativos ou de sinais opostos. Ademais, note-se que outros critérios para identificar quadros em estado estável/de transição podem ser também usados.

Com base na saída do detector de ativação/explosão d , o nível de energia de banda alta estimada pode ser ajustado conforme segue: Quando $d = 1$, isto indica que o quadro correspondente contém uma ativação, por exemplo, transição de silêncio para som não-vozeado ou vozeado, ou um som explosivo. Uma ativação/explosão é detectada no quadro atual se a energia de banda larga do quadro precedente estiver abaixo de um certo limite e a diferença de energia entre os quadros atual e precedente exceder outro limite. Em outra implantação, as energias de banda de transição dos quadros atual e precedente são usadas para detectar uma ativação/explosão. Outros métodos para detectar uma ativação/explosão podem ser também empregados. Uma ativação/explosão representa um problema especial por causa das seguintes razões: A) A estimação de energia de banda alta próxima a ativação/explosão é difícil; B) Os artefatos do tipo pré-eco podem ocorrer na fala de saída por causa do processamento de bloco típico empregado; e C) Os sons explosivos (por exemplo, [p], [t], e [k]), após seu estouro de energia inicial, têm características similares a certas sibilantes (por exemplo, [s], [ʃ] e [ʒ]) na banda larga, mas bem diferentes na banda alta levando a superestimação de energia e consequentes artefatos. A adaptação de energia de banda alta para uma ativação/explosão ($d = 1$) é feita conforme segue:

$$E_{hb}(k) = \begin{cases} E_{\min} & \text{for } k = 1, \dots, K_{\min} \\ E_{hb4}(k) - \Delta & \text{para } k = K_{\min} + 1, \dots, K_T \text{ se } v(k) > V_1 \\ E_{hb4}(k) - \Delta + \Delta_T(k - K_T) & \text{para } k = K_T + 1, \dots, K_{\max} \text{ se } v(k) > V_1 \end{cases}$$

onde k é o índice de quadro. Para os primeiros quadros de $K_{\min}$ começando com o quadro ($k = 1$) em que a ativação/explosão é detectada, a energia de banda alta é definida para o valor mais baixo possível $E_{\min}$. Por exemplo, $E_{\min}$ pode ser definido para $-\infty$ dB ou para a energia do formato de envelope espectral de banda alta com a energia mais baixa. Para os quadros subsequentes (isto é, para a faixa dada por $k = K_{\min} + 1$ para $k = K_{\max}$), a adaptação de energia é feita somente enquanto o nível de voz $v(k)$ do quadro excede o limite V_1 . Ao invés do parâmetro de nível de voz, o parâmetro de cruzamento por zero zc com um limite apropriado pode ser também usado para este fim. Sempre que o nível de voz de um quadro dentro desta faixa se tornar menos ou igual a V_1 , a adaptação de energia de ativação é imediatamente parada, isto é, $E_{hb}(k)$ é definido igual a $E_{hb4}(k)$ até que a próxima ativação seja detectada. Se o nível de voz $v(k)$ for maior que V_1 , então para $k = K_{\min} + 1$ a $k = K_T$, a energia de banda alta é diminuída por uma quantidade fixada Δ . Para $k = K_T + 1$ a $k = K_{\max}$, a energia de banda alta é gradualmente aumentada a partir de $E_{hb4}(k) - \Delta$ em direção a $E_{hb4}(k)$ por meio da sequência pré-especificada $\Delta T(k - K_T)$ e em $k = K_{\max} + 1$, $E_{hb}(k)$ é definido igual a $E_{hb4}(k)$, e isto continua até a próxima ativação ser detectada. Os valores típicos dos parâmetros usados para ativação/explosão com base da adaptação de energia, por exemplo, são $K_{\min} = 2$, $K_T = 3$, $K_{\max} = 5$, $V_1 = 0,9$, $\Delta = -12$ dB, $\Delta T(1) = 6$ dB e $\Delta T(2) = 9,5$ dB. Para $d = 0$, nenhuma

adaptação adicional da energia é feita, isto é, E_{hb} é definido igual a E_{hb4} . Assim, a etapa de modificar o nível de energia de banda alta estimada com base nas características de sinal de banda larga pode compreender a etapa de modificar o nível de energia de banda alta estimada (ou o nível previamente modificado de energia de banda alta estimada) com base em uma ocorrência de uma ativação/explosão.

A adaptação da energia de banda alta estimada conforme descrito anteriormente ajuda a minimizar a quantidade de artefatos na fala de saída de largura de banda estendida e assim acentuar a qualidade. Embora a sequência de operações usadas para adaptar a energia de banda alta estimada tenha sido apresentada de uma forma particular, aqueles versados na técnica reconhecerão que tal especificidade com respeito a sequência não é um requisito, e sendo assim, outras sequências podem ser usadas e permaneceriam em concordância com as modalidades reveladas no presente. Além disso, as operações descritas para modificar o nível de energia de banda alta podem ser seletivamente aplicadas nas modalidades.

Sendo assim a lógica de processamento de sinal e métodos de operação foram revelados no presente para estimar uma porção espectral de banda alta, na faixa de cerca de 7 a 14 kHz, e determinar coeficientes de MDCT de forma que uma saída de áudio que tem uma porção espectral na banda alta possa ser fornecida. Outras variações que poderiam ser equivalente às modalidades reveladas no presente podem ocorrer àqueles versados na técnica e permaneceriam em concordância com o espírito e escopo das

modalidades as definidas no presente pelas seguintes
reivindicações.

REIVINDICAÇÕES

1. Método **caracterizado** por:

definir uma banda de transição para um sinal de áudio que tem um espectro dentro de uma primeira banda de frequência, dita banda de transição definida como uma porção de dita primeira banda de frequência, dita banda de transição sendo localizada próxima a uma banda de frequência adjacente que é adjacente a dita primeira banda de frequência;

10 analisar dita banda de transição para obter um envelope espectral de banda de transição e um espectro de excitação de banda de transição;

estimar um envelope espectral de banda de frequência adjacente;

15 gerar um espectro de excitação de banda de frequência adjacente por repetição periódica de ao menos uma parte do dito espectro de excitação de banda de transição com um período de repetição determinado por uma frequência de timbre de dito sinal de áudio; e

20 combinar o dito envelope espectral de banda de frequência adjacente e o dito espectro de excitação de banda de frequência adjacente para obter um espectro de sinal de banda de frequência adjacente.

2. Método, de acordo com a reivindicação 1,
25 **caracterizado** pelo fato de que a estimativa de um envelope espectral de banda de frequência adjacente compreende ainda estimar dita energia do sinal em dita banda de frequência adjacente.

3. Método, de acordo com a reivindicação 1,
30 **caracterizado** por compreender ainda combinar o dito

espectro com dita primeira banda de frequência e dito espectro de sinal de banda de frequência adjacente para obter um espectro de sinal de largura de banda estendida e um sinal de largura de banda estendida correspondente.

5 4. Método, de acordo com a reivindicação 3, **caracterizado** pelo fato de que a geração do dito espectro de excitação de banda de frequência adjacente, compreende ainda mesclar dito espectro de excitação de banda de frequência adjacente gerado por repetição periódica de ao
10 menos uma parte de dito espectro de excitação de banda de transição com um espectro de excitação de pseudo-ruído dentro de dita banda de frequência adjacente.

 5. Método, de acordo com a reivindicação 4, **caracterizado** por compreender ainda determinar uma razão de
15 mescla, para mesclar dito espectro de excitação de banda de frequência adjacente e dito espectro de excitação de pseudo-ruído, com uso de um nível de voz estimado a partir de dito sinal.

 6. Método, de acordo com a reivindicação 5,
20 **caracterizado** por compreender ainda preencher quaisquer orifícios em dito espectro de excitação de banda de frequência adjacente devido a orifícios correspondentes em dito espectro de excitação de banda de transição com uso de dito espectro de excitação de pseudo-ruído.

25 7. Dispositivo **caracterizado** por compreender:
uma entrada onde um sinal de áudio é fornecido; e
um processador acoplado à entrada, em que o processador é configurado para:

 definir uma banda de transição para o sinal de
30 áudio que tem um espectro dentro de uma primeira banda de

frequência, a dita banda de transição definida como uma porção da dita primeira banda de frequência, a dita banda de transição sendo localizada próxima a uma banda de frequência adjacente que é adjacente à dita primeira banda de frequência;

analisar a dita banda de transição para obter um envelope espectral de banda de transição e um espectro de excitação de banda de transição;

estimar um envelope espectral de banda de frequência adjacente;

gerar um espectro de excitação de banda de frequência adjacente por repetição periódica de ao menos uma parte de dito espectro de excitação de banda de transição com um período de repetição determinado por uma frequência de timbre do dito sinal de áudio; e

combinar o dito envelope espectral de banda de frequência adjacente e o dito espectro de excitação de banda de frequência adjacente para obter um espectro de sinal de banda de frequência adjacente.

8. Dispositivo, de acordo com a reivindicação 7, **caracterizado** pelo fato de que o dito processador é configurado ainda para estimar a energia do dito sinal de áudio na dita banda de frequência adjacente.

9. Dispositivo, de acordo com a reivindicação 8, **caracterizado** pelo fato de que o dito processador é configurado ainda para combinar o dito espectro dentro da dita primeira banda de frequência e o dito espectro de sinal de banda de frequência adjacente para obter um espectro de sinal de largura de banda estendida e um sinal de largura de banda estendida correspondente.

10. Dispositivo, de acordo com a reivindicação 8, **caracterizado** pelo fato de que o dito processador é configurado ainda para mesclar o dito espectro de excitação de banda de frequência adjacente gerado por repetição

5 periódica de ao menos uma parte do dito espectro de excitação de banda de transição com um espectro de excitação de pseudo-ruído dentro da dita banda de frequência adjacente.

11. Dispositivo, de acordo com a reivindicação 10, **caracterizado** pelo fato de que o dito processador é configurado ainda para determinar uma razão de mescla, para mesclar o dito espectro de excitação de banda de frequência adjacente e o dito espectro de excitação de pseudo-ruído, com uso de um nível de voz estimado a partir do dito sinal

15 de áudio.

12. Dispositivo, de acordo com a reivindicação 11, **caracterizado** pelo fato de que o dito processador é configurado ainda para preencher quaisquer orifícios em dito espectro de excitação de banda de frequência adjacente

20 devido a orifícios correspondente em dito espectro de excitação de banda de transição com uso de dito espectro de excitação de pseudo-ruído.

100

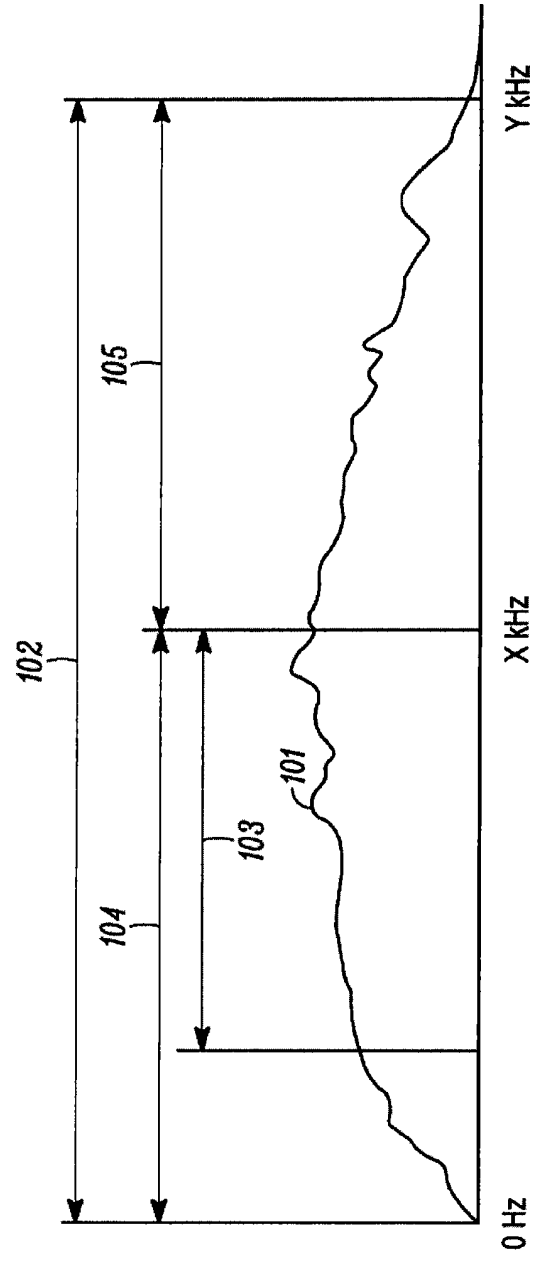
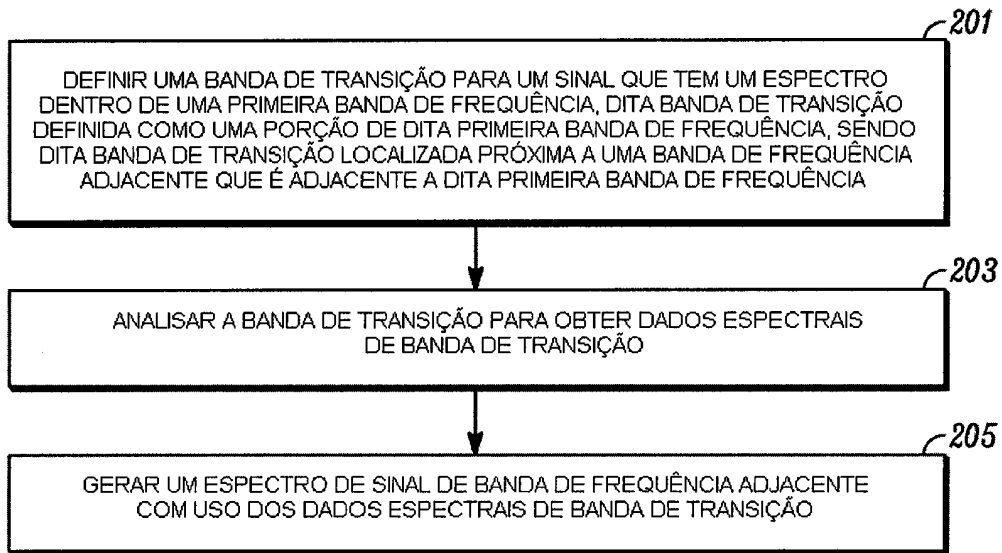
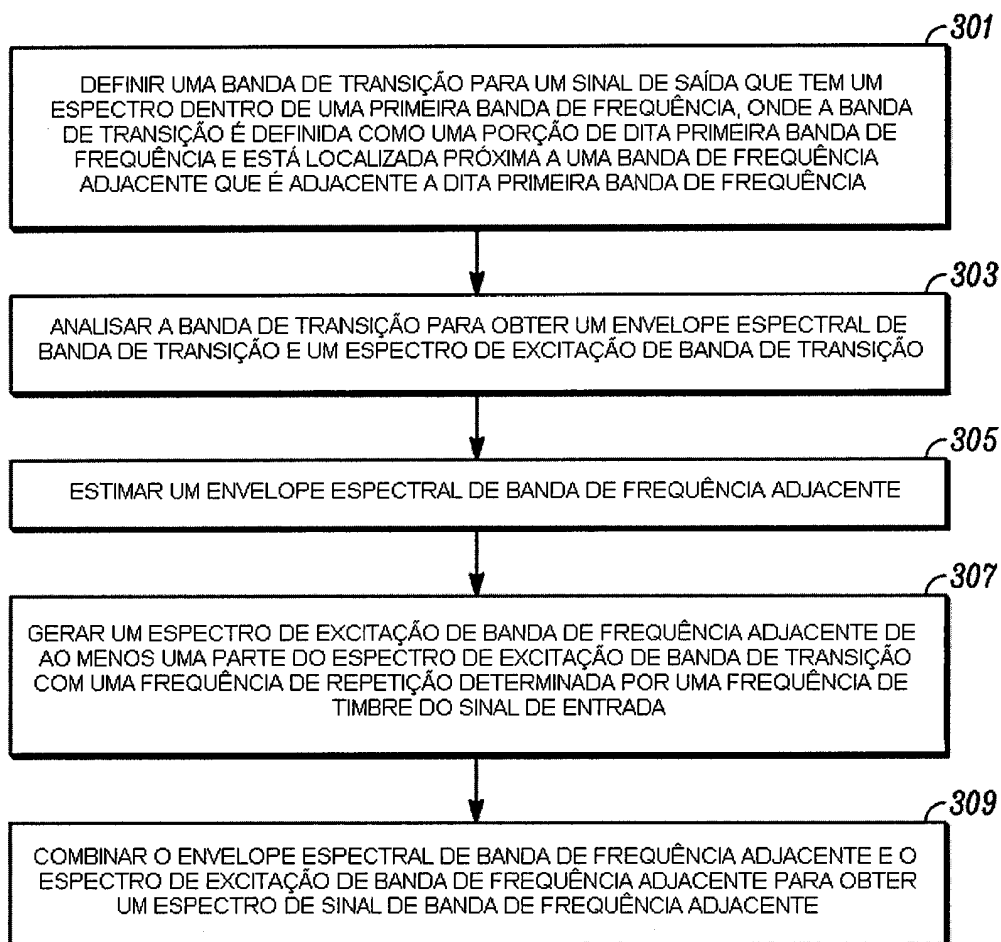


FIG. 1

*FIG. 2*

*FIG. 3*

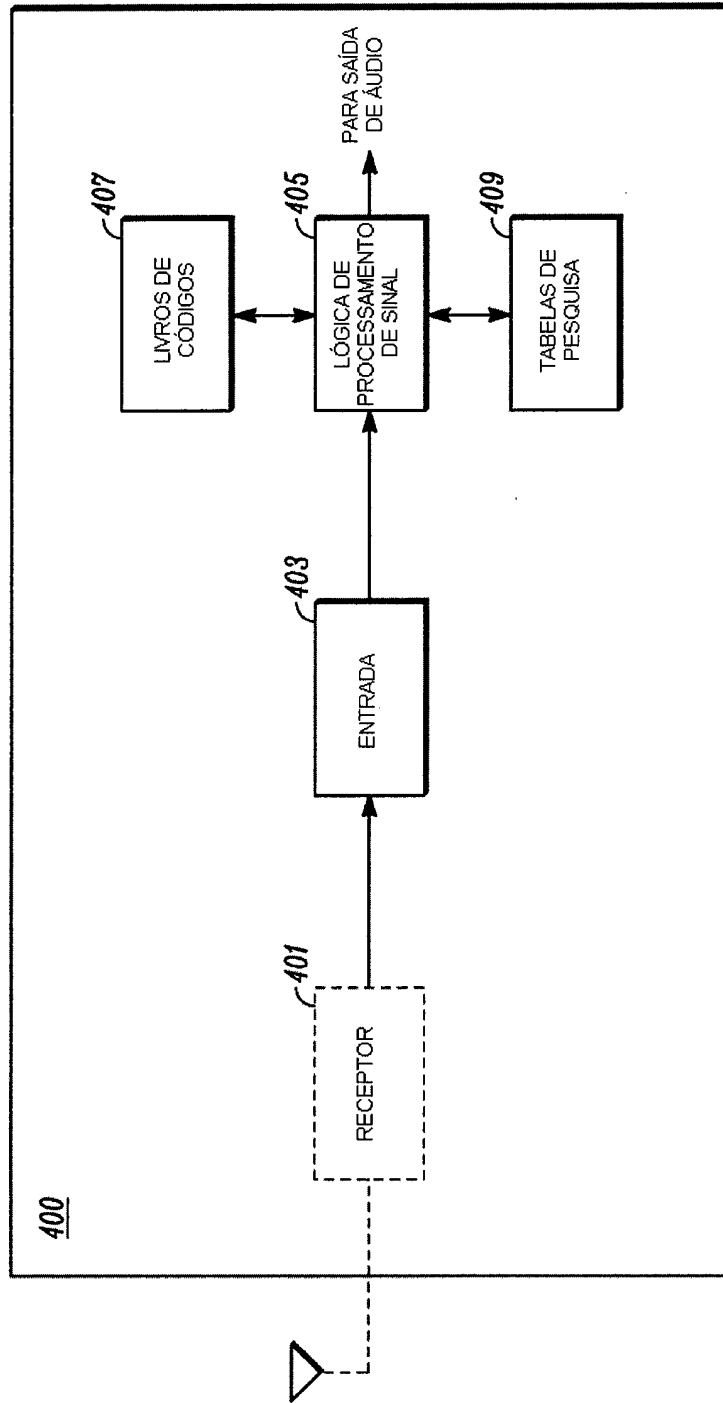


FIG. 4

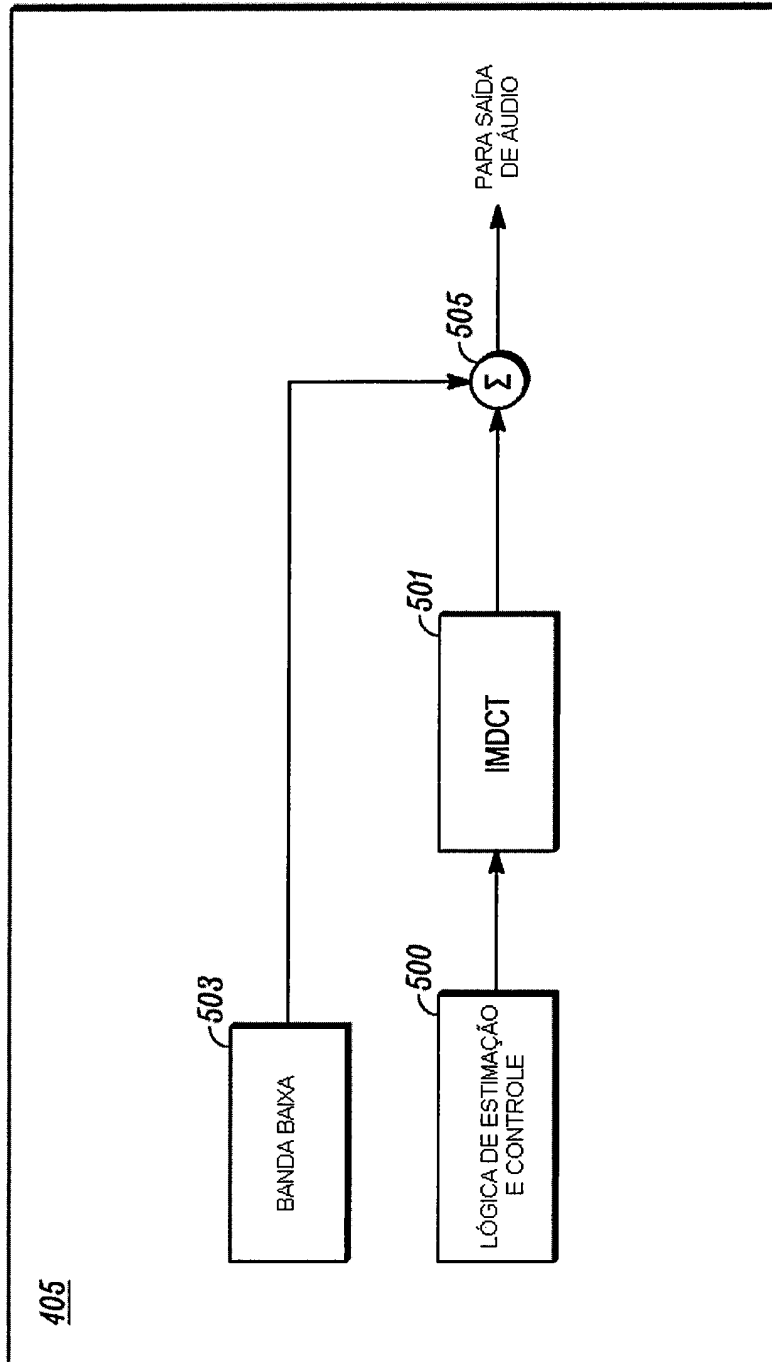


FIG. 5

