



República Federativa do Brasil
Ministério do Desenvolvimento, Indústria
e do Comércio Exterior
Instituto Nacional da Propriedade Industrial

(21) PI 0718300-3 A2



* B R P I 0 7 1 8 3 0 0 A 2 *

(22) Data de Depósito: 24/10/2007
(43) Data da Publicação: 07/01/2014
(RPI 2244)

(51) Int.Cl.:
G01L 19/08
G01L 19/12

(54) Título: MÉTODO E DISPOSITIVO PARA
CODIFICAR QUADROS DE TRANSIÇÃO EM SINAIS
DE FALA.

(57) Resumo:

(30) Prioridade Unionista: 24/10/2006 US 60/853,749

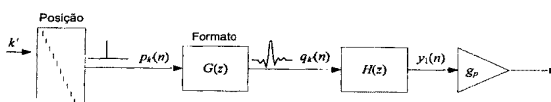
(73) Titular(es): Voiceage Corporation

(72) Inventor(es): Milan Jelinek, Redwan Salami, Vaclav Eksler

(74) Procurador(es): Dannemann, Siemsen, Bigler &
Ipanema Moreira

(86) Pedido Internacional: PCT CA2007001896 de
24/10/2007

(87) Publicação Internacional: WO 2008/049221 de
02/05/2008



Relatório Descritivo da Patente de Invenção para **"MÉTODO E DISPOSITIVO PARA CODIFICAR QUADROS DE TRANSIÇÃO EM SINAIS DE FALA"**.

Campo da Invenção

5 A presente invenção refere-se a uma técnica para codificar digitalmente um sinal sonoro, por exemplo, um sinal de fala ou de áudio, em vista de transmitir e sintetizar o referido sinal de som.

Mais especificamente, mas não exclusivamente, a presente invenção refere-se a um método e dispositivo para codificar quadros de transi-
10 ção e quadros seguindo a transição em um sinal de som, por exemplo um sinal de fala ou de áudio, de modo a reduzir a propagação de erro no decodificador em caso de apagamento do quadro e/ou para aumentar a eficiência da codificação principalmente no início de segmentos sonoros (quadros de início). Em particular, o método e dispositivo substituem o livro de codifica-
15 ção adaptativo tipicamente usado em codificadores previsivos por um livro de codificação, por exemplo, de formatos de impulsos glóticos em quadros de transição e em quadros seguindo a transição. O livro de codificação de formato glótico pode ser um livro de codificação fixo independente da excitação passada pela qual, uma vez que o apagamento do quadro passou, o
20 codificador e o decodificador usam a mesma excitação de modo que convergência para síntese de sinal limpo é relativamente rápida. Na codificação do quadro de início em CELP tradicional, o amortecedor de excitação passada é atualizado usando a excitação do tipo ruído do quadro anterior mudo ou inativo que é bastante diferente da excitação atual. Por outro lado, a técnica proposta pode construir a parte periódica da excitação com bastante
25 precisão.

Antecedentes

Um codificador de fala converte um sinal de fala em uma corrente de bits digitais que é transmitida por um canal de comunicação ou armazenada em um meio de armazenamento. O sinal de fala é digitalizado, o qual é
30 amostrado e quantificado com usualmente 16-bits por amostra. O codificador de fala é dotado do papel de representar as referidas amostras digitais com

um menor número de bits e ainda manter uma boa qualidade de fala subjetiva. O decodificador de fala ou sintetizador opera na corrente de bits transmitida ou armazenada e converte a mesma de volta em um sinal de fala.

A codificação Code-Excited Previsão linear (*CELP*) é uma das
5 melhores técnicas anteriores para alcançar um bom compromisso entre qualidade subjetiva e coeficiente de bit. A referida técnica de codificação forma a base de diversos padrões de codificação de fala tanto em aplicações sem fio e com fio. Em Codificação CELP, o sinal de fala amostrado é processado em blocos sucessivos de amostras M usualmente chamada de quadros, onde M
10 é um número predeterminado correspondendo tipicamente a 10-30 ms. Um filtro de previsão linear (*LP*) é computado e transmitido a cada quadro. A computação do filtro LP tipicamente precisa de um "análise antecipada", um segmento de fala de 5-15 ms a partir do quadro subsequente. O quadro de amostra- M é dividido em blocos menores chamada de subquadros. Usualmente o número de subquadros é três ou quatro resultando em 4-10 ms
15 subquadros. Em cada subquadro, um sinal de excitação é usualmente obtido a partir de dois componentes, a excitação passada e a inovadora, excitação do livro de codificação fixo. O componente formado a partir da excitação passada é com frequência referido como o livro de codificação adaptativo ou
20 passo de excitação. Os parâmetros que caracterizam o sinal de excitação são codificados e transmitidos ao decodificador, onde o sinal de excitação reconstruído é usado como a entrada do filtro LP.

Codecs de fala do tipo CELP se baseiam amplamente em previsão para alcançar seu alto desempenho. A previsão usada pode ser de diferentes tipos mas usualmente compreende o uso de um livro de codificação
25 adaptativo contendo um sinal de excitação selecionado em quadros passados. O *Codificador CELP* explora a quase periodicidade dos sinais de fala sonoros ao pesquisar em uma excitação passada o segmento mais similar ao segmento sendo atualmente codificados. O mesmo sinal de excitação
30 passado é mantido também no decodificador. É então suficiente para o codificador enviar um parâmetro de retardo e um ganho para o decodificador para reconstruir o mesmo sinal de excitação que é usado no codificador. A

evolução (diferença) entre o segmento de fala anterior e o segmento de fala atualmente codificado é adicionalmente modelada usando uma inovação selecionada a partir de um livro de codificação fixo. A tecnologia CELP será descrita em maiores detalhes abaixo.

5 Um problema de forte previsão inerente a codificadores de fala com base em CELP aparece na presença de erros de transmissão (quadros ou pacotes apagados) quando o estado do codificador e do decodificador se torna fora de sincronia. Em virtude da previsão, o efeito de um quadro apagado é assim não limitado ao quadro apagado, mas continua para se propagar após
10 o apagamento, com frequência durante diversos quadros subsequentes. Naturalmente, o impacto perceptivo pode ser bastante inconveniente.

Transições a partir de segmento de fala mudo para segmento de fala sonoro (por exemplo, transição entre uma consoante ou um período de fala inativa, e uma vogal) ou transições entre dois diferentes segmentos sonoros (por exemplo, transições entre duas vogais) são os casos mais problemáticos para o encobrimento de apagamento do quadro. Quando uma transição a partir de segmento de fala mudo para segmento de fala sonoro (início sonoro) é perdido, o quadro logo antes do quadro de início sonoro é mudo ou inativo e assim nenhuma excitação periódica significativa é encontrada no amortecedor da excitação passada (livro de codificação adaptativo).
15 No codificador, a excitação periódica passada se constroi no livro de codificação adaptativo durante o quadro de início, e o quadro sonoro a seguir é codificado usando a referida excitação periódica passada. Mais técnicas de encobrimento de erro de quadro usam a informação a partir do ultimo quadro corretamente recebido para encobrir o quadro que falta. Quando o quadro de
20 início é perdido, o amortecedor decodificador de excitação passada será assim atualizado usando a excitação do tipo ruído do quadro anterior (quadro mudo ou inativo). A parte periódica da excitação é assim completamente perdida no livro de codificação adaptativo no decodificador após um início
25 sonoro perdido e a mesma pode compensar diversos quadros para o decodificador recuperar a partir da referida perda.
30

Uma situação similar ocorre em caso de transição de sonoro

perdido para sonoro. Neste caso, a excitação armazenada no livro de codificação adaptativo antes do quadro de transição ser dotado de tipicamente características bastante diferentes a partir da excitação armazenada no livro de codificação adaptativo após a transição. Mais uma vez, na medida em que o decodificador usualmente encobre o quadro perdido com o uso da informação de quadro passado, o estado do codificador e do decodificador será bastante diferente, e o sinal sintetizado pode sofrer a partir de importante distorção.

OBJETIVOS DA INVENÇÃO

Um objetivo da presente invenção é, portanto, proporcionar um método e dispositivo para codificar quadros de transição em um codificador preditivo de fala e/ou áudio de modo a aprimorar a resistência do codificador contra quadros perdidos e/ou aprimorar a eficiência da codificação.

Outro objetivo da presente invenção é eliminar propagação de erro e aumento da eficiência da codificação em Codecs com base em CELP ao substituir a pesquisa do livro de codificação adaptativo dependente interquadros por uma pesquisa do livro de codificação não-previsiva, por exemplo, de formato glótico. A referida técnica não requer retardo extra, complexidade adicional insignificante, e não aumenta o coeficiente de bit comparado a codificação CELP tradicional.

Sumário da Invenção

Mais especificamente, de acordo com um aspecto da presente invenção, é proporcionado um método de modo de transição para uso em um Codec de sinal de som do tipo preditivo para produzir uma excitação de modo de transição substituindo uma excitação de livro de codificação adaptativo em um quadro de transição e/ou um quadro seguindo a transição no sinal de som, compreendendo: proporcionar um livro de codificação de modo de transição para gerar um conjunto de vetores-código independente da excitação passada; fornecer um índice de livro de codificação ao livro de codificação de modo de transição; e gerar, por meio do livro de codificação de modo de transição e em resposta ao índice do livro de codificação, um dos vetores-código do conjunto que corresponde à excitação de modo de transi-

ção.

De acordo com um segundo aspecto da presente invenção, é proporcionado um dispositivo de modo de transição para uso em um Codec de sinal de som do tipo preditivo para produzir uma excitação de modo de transição substituindo uma excitação de livro de codificação adaptativo em um quadro de transição e/ou um quadro seguindo a transição no sinal de som, compreendendo uma entrada para receber um índice de livro de codificação e um livro de codificação de modo de transição para gerar um conjunto de vetores-código independente da excitação passada. O livro de codificação de modo de transição responde ao índice para gerar, no quadro de transição e/ou quadro seguindo a transição, um dos vetores-código do conjunto que corresponde à referida excitação de modo de transição.

De acordo com um terceiro aspecto da presente invenção, é proporcionado um método de codificação para gerar uma excitação de modo de transição substituindo uma excitação de livro de codificação adaptativo em um quadro de transição e/ou um quadro seguindo a transição em um sinal de som, compreendendo: gerar um sinal alvo de pesquisa no livro de codificação; proporcionar um livro de codificação de modo de transição para gerar um conjunto de vetores-código independente da excitação passada, os vetores-código do conjunto cada um dos quais correspondendo a uma excitação de modo de transição respectiva; pesquisar o livro de codificação de modo de transição para encontrar o vetor-código do conjunto que corresponde a uma excitação de modo de transição otimamente correspondendo ao sinal alvo de pesquisa no livro de codificação.

De acordo com um quarto aspecto da presente invenção, é proporcionado um dispositivo codificador para gerar uma excitação de modo de transição substituindo uma excitação de livro de codificação adaptativo em um quadro de transição e/ou um quadro seguindo a transição em um sinal de som, compreendendo: um gerador de um sinal alvo de pesquisa no livro de codificação; um livro de codificação de modo de transição para gerar um conjunto de vetores-código independente da excitação passada, os vetores-código do conjunto cada correspondendo a uma excitação de modo de tran-

sição respectiva; e um pesquisador do livro de codificação de modo de transição para encontrar o vetor-código do conjunto que corresponde a uma excitação de modo de transição otimamente correspondendo ao sinal alvo de pesquisa no livro de codificação.

5 De acordo com um quinto aspecto da presente invenção, é proporcionado um método de decodificação para gerar uma excitação de modo de transição substituindo uma excitação de livro de codificação adaptativo em um quadro de transição e/ou um quadro seguindo a transição em um sinal de som, compreendendo: receber um índice de livro de codificação;
10 fornecer o índice do livro de codificação to um livro de codificação de modo de transição para gerar um conjunto de vetores-código independente da excitação passada; e gerar, por meio do livro de codificação de modo de transição e em resposta ao índice do livro de codificação, um dos vetores-código do conjunto que corresponde à excitação de modo de transição.

15 De acordo com um sexto aspecto da presente invenção, é proporcionado um dispositivo decodificador para gerar uma excitação de modo de transição substituindo uma excitação de livro de codificação adaptativo em um quadro de transição e/ou um quadro seguindo a transição em um sinal de som, compreendendo uma entrada para receber um índice de livro
20 de codificação e um livro de codificação de modo de transição para gerar um conjunto de vetores-código independente da excitação passada. O livro de codificação de modo de transição responde ao índice para gerar no quadro de transição e/ou quadro seguindo a transição um dos vetores-código do conjunto que corresponde à excitação de modo de transição.

25 Os referidos objetivos, vantagens e características anteriores e outras da presente invenção se tornará mais aparente com a leitura da descrição não-restritiva a seguir de uma modalidade ilustrativa da mesma, oferecidas apenas como exemplo, com referência aos desenhos anexos.

Breve Descrição dos Desenhos

30 Nos desenhos anexos:

Figura 1a é um diagrama de bloco esquemático de um codificador com base em CELP.

Figura 1b é um diagrama de bloco esquemático de um decodificador com base em CELP.

Figura 2 é um diagrama de bloco esquemático de uma máquina de classificação de estado de quadro para encobrimento de apagamento.

5 Figura 3 é um exemplo de segmento de um sinal de fala com um quadro de transição sonoro e um quadro de início.

Figura 4 é um diagrama de bloco funcional ilustrando uma regra de classificação para selecionar quadros *TM* (Modo de transição) em inícios de fala, onde *QUADROS N_TM_* representam uma série de quadros consecutivos para evitar o uso de uma técnica de codificação *TM*, 'clas' representa uma classe de quadros, e *TIPO_SONORO* significa classes de *INÍCIO*, *SONORO* e *TRANSIÇÃO SONORA*.

10

Figura 5a é uma ilustração esquemática de um exemplo de quadro de um sinal de fala dividido em quatro (4) subquadros, mostradodo o sinal de fala no domínio de tempo.

15

Figura 5b é uma ilustração esquemática de um exemplo de quadro de um sinal de fala dividido em quatro (4) subquadros, mostradodo um sinal residual LP.

Figura 5c é uma ilustração esquemática de um exemplo de quadro de um sinal de fala dividido em quatro (4) subquadros, mostradodo um primeiro estágio de sinal de excitação construído usando a técnica de codificação *TM* no codificador.

20

Figura 6 mostra gráficos ilustrando oito impulsos glóticos com comprimento de 17 amostras usadas para a construção do livro de codificação de formato glótico, onde o eixo x denota um índice de tempo distinto e o eixo y uma amplitude do impulso.

25

Figura 7 é um diagrama de bloco esquemático de um exemplo de porção *TM* de um codificador *CELP*, onde k' representa um índice de livro de codificação de formato glótico e $G(z)$ é um filtro de forma.

Figura 8 é uma representação gráfica de uma computação de Ck' a raiz quadrada do numerador no critério da Equação (16), onde porções sombreadas de vetor/matriz são não-zero.

30

Figura 9 é uma representação gráfica da computação de Ek' , o denominador do critério de Equação (16), onde porções sombreadas do vetor/matriz são não-zero.

Figura 10 é uma representação gráfica da computação da matriz de convolução \mathbf{Z}^T ; neste exemplo o filtro de forma $G(z)$ é dotado apenas de três (3) coeficientes não-zero ($L_{1/2} = 1$).

Figura 11 é um diagrama de bloco esquemático de um exemplo de porção TM de um Decodificador CELP.

Figura 12a é um diagrama de bloco esquemático de um exemplo de estrutura de filtro $Q(z)$.

Figura 12b é um gráfico de um exemplo de modificação de vetor-código de formato glótico, onde o impulso repetido é pontilhado;

Figura 13 é um diagrama de bloco esquemático da porção TM de um codificador $CELP$ incluindo o filtro $Q(z)$.

Figura 14 é um gráfico ilustrando um vetor-código de formato glótico com construção de dois impulsos quando uma pesquisa de livro de codificação adaptativo é usada em uma parte do subquadro com uma pesquisa de livro de codificação de formato glótico.

Figura 15 é um gráfico ilustrando uma construção vetor-código de formato glótico no caso onde o segundo impulso glótico aparece na primeira posição $L_{1/2}$ do próximo subquadro.

Figura 16 é um diagrama de bloco esquemático da porção TM de um codificador usado em uma implementação de Codec EV-VBR (Coeficiente de bit variável embutido).

Figura 17a é um gráfico mostrando um exemplo de sinal de fala no domínio de tempo.

Figura 17b é um gráfico mostrando um sinal residual LP correspondendo ao sinal de fala da figura 17a.

Figura 17c é um gráfico mostradodo um primeiro estágio de sinal de excitação em condições livres de erro.

Figuras 18a - 18c são gráficos ilustrando um exemplo de comparação de construção de início, onde o gráfico da figura 18a representa o si-

nal de entrada de fala, o gráfico da figura 18b representa a saída de fala sintetizada de um Codec EV-VBR sem a técnica de codificação *TM*, e o gráfico da figura 18c representa a saída de fala sintetizada de um Codec EV-VBR com a técnica de codificação *TM*.

5 As figuras 19a - 19c são gráficos ilustrando um exemplo do efeito da técnica de codificação *TM* em caso de apagamento do quadro, onde o gráfico da figura 19a representa o sinal de entrada de fala, o gráfico da figura 19b representa a saída de fala sintetizada de um Codec EV-VBR sem a técnica de codificação *TM*, e o gráfico da figura 19c representa a saída de fala sintetizada de um Codec EV-VBR com a técnica de codificação *TM*.

10 Figura 20 é um gráfico ilustrando um exemplo do primeiro estágio de sinal de excitação em um quadro de configuração TRANSIÇÃO_1_1.

Figura 21 é um gráfico ilustrando um exemplo do primeiro estágio de sinal de excitação em um quadro de configuração TRANSIÇÃO_1_2.

15 Figura 22 é um gráfico ilustrando um exemplo do primeiro estágio de sinal de excitação em um quadro de configuração TRANSIÇÕES_1_3.

Figura 23 é um gráfico ilustrando um exemplo do primeiro estágio de sinal de excitação em um quadro de configuração TRANSIÇÃO_1_4.

20 Figura 24 é um gráfico ilustrando um exemplo do primeiro estágio de sinal de excitação em um quadro de configuração TRANSIÇÃO_2.

Figura 25 é um gráfico ilustrando um exemplo do primeiro estágio de sinal de excitação em um quadro de configuração TRANSIÇÃO_3.

Figura 26 é um gráfico ilustrando um exemplo do primeiro estágio de sinal de excitação em um quadro de configuração TRANSIÇÃO_4.

25 Figura 27 é um diagrama de bloco esquemático de um sistema de comunicação de fala ilustrando o uso de dispositivos codificador e decodificador de fala.

Descrição Detalhada

30 A modalidade ilustrativa não-restritiva da presente invenção e refere a um método e dispositivo cujo objetivo é superar a propagação de erro nas situações acima discutidas e aumentar a eficiência da codificação.

Mais especificamente, o método e dispositivo de acordo com a

modalidade ilustrativa não-restritiva da presente invenção implementa uma codificação especial, chamada de técnica de codificação de modo de transição (*TM*), de quadros de transição e quadros seguindo a transição em um sinal de som, por exemplo, um sinal de fala ou de áudio. A técnica de codificação *TM* substitui o livro de codificação adaptativo do Codec CELP por um novo livro de codificação de formatos de impulsos glóticos, daqui adiante designado como livro de codificação de formato glótico, em quadros de transição e em quadros seguindo a transição. O livro de codificação de formato glótico é um livro de codificação fixo independente da excitação passada. Consequentemente, uma vez que um apagamento do quadro estiver completo, o codificador e o decodificador usam a mesma excitação pela qual a convergência em síntese de sinal limpo é relativamente rápida.

Embora o uso da técnica de codificação *TM* em quadros em seguida da transição ajude a evitar a propagação de erro em caso do quadro de transição ser perdido, outro objetivo de usar a técnica de codificação *TM* também no quadro de transição é de aprimorar a eficiência da codificação. Por exemplo, logo antes de um princípio sonante, o livro de codificação adaptativo usualmente contém um sinal como ruído não muito eficiente para codificar o início de um segmento sonoro. A ideia por trás da técnica de codificação *TM* é assim suplementar o livro de codificação adaptativo com um melhor livro de codificação composto por versões quantificadas simplificadas de impulsos glóticos para codificar os inícios sonoros.

A técnica de codificação *TM* proposta pode ser usada em qualquer Codec do tipo *CELP* ou Codec preditivo. Como um exemplo, a técnica de codificação *TM* é implementada em um Codec candidato em *Atividade de padronização ITU-T* para um Codec de coeficiente de bit variável embutido que será referido no restante do texto como Codec EV-VBR. Embora a modalidade ilustrativa não-restritiva da presente invenção seja descrita em conexão com a estrutura de Codec EV-VBR, deve se ter em mente que os princípios e os conceitos da presente invenção não são limitados a uma aplicação do Codec EV-VBR mas a qualquer outro Codec usando codificação previsiva. Também, embora a modalidade ilustrativa não-restritiva da pre-

sente invenção seja descrita em conexão com um sinal de fala, deve se ter em mente que a presente invenção não é limitada a uma aplicação para sinais de fala mas seus princípios e conceitos podem ser aplicados a qualquer outro tipo de sinais de som incluindo sinais de áudio.

5 Um quadro de fala pode ser basicamente classificado em uma de quatro (4) classes de fala a seguir (isto será explicado em maiores detalhes na descrição a seguir):

- Quadros inativos caracterizados por a ausência de atividade de fala.
- 10 - Quadros de fala muda caracterizados por uma estrutura aperiódica e concentração de energia em direção de frequências mais altas.
- Quadros de fala sonora dotados de uma natureza clara quase periódica com energia concentrada principalmente em baixas frequências.
- Qualquer outro quadro classificado como uma transição dotada de características rapidamente variáveis.

15 No Codec EV-VBR, um modo de codificação especializado foi projetado para cada uma das classes. Pode ser geralmente determinado que os quadros inativos são processados através de geração de ruído confortável, os quadros de fala muda através de um modo de codificação mudo otimizado, os quadros de fala sonora através de um modo de codificação sonoro otimizado e todos os outros quadros são processados com uma tecnologia genérica CELP algébrica(ACELP). Na estrutura de Codec EV-VBR, a técnica de codificação *TM* é assim introduzida ainda como um outro modo de codificação no *Esquema de codificação EV-VBR* para codificar quadros de transição e quadros seguindo a transição.

25 A figura 27 é um diagrama de bloco esquemático de um sistema de comunicação de fala ilustrando o uso de codificador e decodificador de fala. O sistema de comunicação de fala suporta transmissão e reprodução de um sinal de fala através de um canal de comunicação 905. Embora o mesmo possa compreender, por exemplo, um fio, ótico ou ligação de fibra, o canal de comunicação 905 tipicamente compreende pelo menos em parte
30 uma ligação de frequência de radio. A ligação de frequência de radio com

frequência suporta comunicações de fala múltiplas, simultâneas necessitando de recursos de largura de banda compartilhada tal como pode ser encontrado com telefonia celular. Embora não mostrado, o canal de comunicação 905 pode ser substituído por um dispositivo de armazenamento em uma modalidade de dispositivo único do sistema de comunicação que registra e armazena os sinais de fala codificados para reprodução posterior.

Ainda com referência à figura 27, um microfone 901 produz um sinal da fala analógico que é fornecido a um conversor de análogo-para-digital (A/D) 902 para converter o mesmo em uma forma digital. Um codificador de fala 903 codifica o sinal de fala digital deste modo produzindo um conjunto de parâmetros de codificação que são codificados em uma forma binária e enviado a um codificador de canal 904. O codificador de canal opcional acrescenta redundância à representação binária dos parâmetros de codificação antes de transmitir os mesmos sobre o canal de comunicação 905. No lado receptor, um decodificador de canal 906 utiliza a informação redundante acima mencionada na corrente de bits recebidos para detectar e corrigir erros de canal que ocorreram na transmissão. Um decodificador de fala 907 converte a corrente de bits recebida a partir do decodificador de canal 906 de volta para um conjunto de parâmetros de codificação para criar um sinal de fala digital sintetizado. O sinal de fala digital sintetizado reconstruído no decodificador de fala 907 é convertido em uma forma análoga em um conversor de digital-para-análogo (D/A) 908 e reproduzido em uma unidade de alto-falante 909.

Curto embasamento em CELP

Um Codec de fala consiste em duas partes básicas: um codificador e um decodificador. O codificador digitaliza o sinal de áudio, escolhe um número limitado de parâmetros de codificação que representam o sinal de fala e converte os referidos parâmetros em uma corrente de bits digitais que é transmitida ao decodificador através de um canal de comunicação. O decodificador reconstrói o sinal de fala para ser o mais similar possível ao sinal de fala original. Atualmente, uma técnica difundida de codificação de fala se baseia em Previsão linear (LP), e mais especificamente em tecnolo-

gia CELP. Em codificação com base em LP, o sinal de fala é sintetizado ao se filtrar um sinal de excitação através de um filtro de síntese de todos os pólos $MA(z)$. Em *CELP*, a excitação é tipicamente composta de duas partes, um primeiro estágio de sinal de excitação é selecionado a partir de um livro de codificação adaptativo e um segundo estágio de sinal de excitação é selecionado a partir de um livro de codificação fixo. Geralmente falando, a excitação do livro de codificação adaptativo modela a parte periódica da excitação e a excitação do livro de codificação fixo é adicionada para modelar a evolução do sinal de fala.

A fala é normalmente processada por quadros de tipicamente 20 ms e os coeficientes de filtro LP são transmitidos uma vez por quadro. Em *CELP*, cada quadro é adicionalmente dividido em diversos subquadros para codificar o sinal de excitação. O comprimento do subquadro é tipicamente 5 ms.

Com referência às figuras 1a e 1b, o princípio principal por trás de *CELP* é chamado de análise-por-síntese onde emissões possíveis do decodificador são já tentadas (síntese) durante o processo de codificação (análise) e então comparadas ao sinal de fala original. A pesquisa minimiza um erro de quadrados mínimos entre o sinal de entrada de fala $s(n)$ e o sinal de fala sintetizado $s'(n)$ em um domínio perceptivamente pesado, onde o índice de tempo distinto $n = 0, 1, \dots, AM$, e N é o comprimento do subquadro. O filtro perceptivo pesado $W(z)$ explora o efeito de mascaramento de frequência e é tipicamente derivado a partir de do filtro LP. Um exemplo de filtro perceptivo pesado $W(z)$ é oferecido na Equação a seguir (1):

$$W(z) = \frac{A(z/\gamma_1)}{A(z/\gamma_2)}, \quad (1)$$

onde fatores γ_1 e γ_2 controlam a quantidade de pesagem perceptiva e mantêm a relação $0 < \gamma_2 < \gamma_1 \leq 1$. O referido filtro perceptivo pesado tradicional funciona bem para sinais *NB* (banda estreita - largura de banda de 200 - 3400 Hz). Um exemplo de filtro perceptivo pesado para sinais *WB* (banda larga - largura de banda de 50 - 7000 Hz) pode ser encontrado na Referência [1].

A corrente de bits transmitida ao decodificador contém para os quadros sonoros os parâmetros de codificação a seguir: os parâmetros quantificados do filtro de síntese LP , os índices do livro de codificação adaptativo e fixo e os ganhos das partes adaptativa e fixa.

5 Pesquisa de livro de codificação adaptativo

A pesquisa do livro de codificação adaptativo em codecs com base em CELP é realizada em domínio de fala pesado para determinar o retardo (período de passo) t e o ganho de passo g_p , e para construir a parte quase periódica de um sinal de excitação referido como vetor-código adaptativo $v(n)$. O período de passo é fortemente dependente do locutor particular e sua determinação precisa influencia fundamentalmente a qualidade da fala sintetizada.

Em um Codec EV-VBR, a procedimento de três estágios é usado para determinar o período de passo e o ganho. No primeiro estágio, três estimativas de passo de alça aberta T_{op} são computadas para cada quadro – uma estimativa para cada 10 ms meio-quadro e uma para um 10 ms de antecipação - usando o sinal de fala perceptivamente pesado $s_w(n)$ e computar a correlação normalizada. No segundo estágio, uma pesquisa de passo de alça fechada é realizada para períodos integrais em torno do período de passos de alça fechada estimado T_{op} para cada subquadro. Uma vez que um período de passo otimizado inteiro é encontrado, um terceiro estágio de pesquisa ocorre através das frações em torno daquele valor inteiro otimizado. A pesquisa de passo de alça fechada é realizada ao minimizar o erro pesado de quadrados mínimos entre a original e a fala sintetizada. Isto é alcançado ao se maximizar o termo

$$\mathfrak{J}_i = \frac{\left(\sum_{n=0}^{N-1} x_1(n) y_1(n) \right)^2}{\sum_{n=0}^{N-1} y_1(n) y_1(n)}, \quad (2)$$

onde $x_1(n)$ é o sinal alvo e o sinal de contribuição do primeiro estágio (também chamado de vetor-código adaptativo filtrado) $y_1(n)$ é computado pelo

5 sinal de convolução da excitação passada $v(n)$ no período t com a resposta de impulso $h(n)$ do filtro de síntese pesado $H(z)$

$$y_1(n) = v(n)*h(n). \quad (3)$$

5 O sinal de fala de entrada perceptivamente pesado $s_w(n)$ é obtido ao processar o sinal de entrada de fala $s(n)$ através do filtro perceptivo pesado $W(z)$. O filtro $H(z)$ é formado pela cascata do filtro de síntese LP $MA(z)$ e do filtro perceptivo pesado $W(z)$. O sinal alvo $x_1(n)$ corresponde ao sinal de fala de entrada perceptivamente pesado $s_w(n)$ após subtrair do mesmo a resposta de entrada zero do filtro $H(z)$.

10 O ganho de passo é encontrado ao minimizar o erro de quadrados mínimos entre o sinal $x_1(n)$ e o sinal de contribuição do primeiro estágio $y_1(n)$. O ganho de passo é expresso pela Equação a seguir:

$$g_p = \frac{\sum_{n=0}^{N-1} x_1(n)y_1(n)}{\sum_{n=0}^{N-1} y_1(n)y_1(n)} \quad (4)$$

15 O ganho de passo é então delimitado por $0 \leq g_p \leq 1.2$ e tipicamente juntamente quantificado com o ganho do livro de codificação fixo uma vez que a inovação é encontrada.

20 Em Codecs com base em CELP, o sinal de excitação no início do quadro atualmente processado é assim reconstruído a partir do sinal de excitação a partir do quadro anterior. O referido mecanismo é muito eficiente para segmentos sonoros do sinal de fala onde o sinal é quase periódico, e na ausência de erros de transmissão. Em caso de apagamento do quadro, o sinal de excitação a partir do quadro anterior é perdido e o respectivo livro de codificação adaptativo do codificador e decodificador não é mais o mesmo. Em quadros que se seguem ao apagamento, o decodificador então continua a sintetizar a fala usando o livro de codificação adaptativo com conteúdo incorreto. Consequentemente, um apagamento do quadro degrada a qualidade da fala sintetizada não só durante o quadro apagado, mas o mesmo pode também degradar a qualidade da fala sintetizada durante diversos quadros

subsequentes. As técnicas de encobrimento tradicionais são com frequência baseadas na repetição da forma de onda do quadro anterior corretamente transmitido, mas as referidas técnicas funcionam eficientemente apenas em partes de sinal onde as características do sinal de fala são quase estacionárias, por exemplo, em segmentos sonoros estáveis. Neste caso, a diferença entre o respectivo livro de codificação adaptativo do codificador e decodificador é com frequência relativamente pequena e a qualidade do sinal sintetizado não é muito afetada. Entretanto, se o apagamento cai em um quadro de transição, a eficiência das referidas técnicas é bastante limitada. Em sistemas de comunicação usando codecs com base em CELP, onde o coeficiente de apagamento do quadro (*FER*) é tipicamente 3% a 5%, a qualidade da fala sintetizada então cai significativamente.

Mesmo em uma transmissão limpa de canal, a eficiência do livro de codificação adaptativo é limitada em quadros de transição; o codificador CELP faz uso do livro de codificação adaptativo para explorar a periodicidade na fala que é baixa ou em falta durante transições pelas quais a eficiência da codificação é reduzida. Este é o caso de inícios sonoros em particular onde o sinal de excitação passado e o sinal de excitação otimizado para o quadro atual são correlacionados com bastante fraqueza ou não são de modo algum.

Pesquisa do Livro de codificação fixo

O objetivo da contribuição pesquisa do livro de codificação (inovação) fixo (*FCB*) em codecs com base em CELP é de minimizar o erro residual após o uso do livro de codificação adaptativo, isto é,

$$\min_k \left\{ \sum_{n=0}^{N-1} [x_2(n) - g_c \cdot y_2^{(k)}(n)]^2 \right\}, \quad (5)$$

onde g_c é o ganho do livro de codificação fixo, e sinal de contribuição do segundo estágio (também chamado do vetor-código fixado filtrado) $y_2^{(k)}(n)$ é o vetor do livro de codificação fixo $C_k(n)$ envolvido com $h(n)$. O sinal alvo $x_1(n)$ é atualizado ao subtrair a contribuição do livro de codificação adaptativo a partir do alvo do livro de codificação adaptativo para obter:

$$x_2(n) = x_1(n) - g_p y_1(n). \quad (6)$$

O livro de codificação fixo pode ser realizado, por exemplo, ao usar um livro de codificação algébrico como descrito na Referência [2]. Se c_k denota o vetor-código algébrico no índice k , então o livro de codificação algébrico é pesquisado ao se maximizar o critério a seguir:

$$\mathfrak{J}_k = \frac{(\mathbf{x}_2^T \mathbf{y}_2)^2}{\mathbf{y}_2^T \mathbf{y}_2} = \frac{(\mathbf{x}_2^T \mathbf{H} \mathbf{c}_k)^2}{\mathbf{c}_k^T \mathbf{H}^T \mathbf{H} \mathbf{c}_k} = \frac{(\mathbf{d}^T \mathbf{c}_k)^2}{\mathbf{c}_k^T \Phi \mathbf{c}_k} = \frac{(C_k)^2}{E_k} \quad (7)$$

5 onde H é matriz de convolução Toeplitz triangular inferior com diagonal $h(0)$ e diagonais inferiores $h(1), \dots, h(N-1)$. Vetor $d = H^T x_2$ é a correlação entre o sinal alvo atualizado $x_2(n)$ e $h(n)$ (também conhecido como vetor alvo filtrado traseiro retrógrado), e matriz $\Phi = H^T H$ é a matriz de correlações de $h(n)$. O sobrescrito⁷ denota a matriz ou transposição de vetor. Ambos d e Φ são usualmente computados antes da pesquisa do livro de codificação fixo. A referência [1] discute que, se a estrutura algébrica do livro de codificação fixo contém apenas poucos elementos não-zeros, a computação do critério de maximização para todos os índices possíveis k é bastante rápida. Um procedimento similar é usado em uma técnica de codificação de modo de transição (TM) como será visto abaixo.

15 Acredita-se que CELP seja de outro modo bem conhecido daqueles versados na técnica e, por este motivo, não será adicionalmente descrito na presente especificação.

Classificação de Quadro no Codec EV-VBR

20 A classificação de quadro no Codec EV-VBR é baseada em classificação VMR-WB (Variable Rate Multi-Mode Wideband) como descrito na Referência [3]. Classificação VMR-WB é realizada com a consideração da estratégia de encobrimento e recuperação. Em outras palavras, qualquer quadro é classificado de tal forma que o encobrimento pode ser otimizado se o quadro a seguir estiver faltando, ou que a recuperação pode ser ótima se o quadro anterior estiver perdido. Algumas das classes usadas para o proces-

sar o encobrimento do apagamento do quadro não precisam ser transmitidas, na medida em que as mesmas podem ser deduzidas sem ambiguidade no decodificador. Cinco classes distintas são usadas, e definidas como a seguir:

- 5 - classe *MUDO* compreende todos os quadros de fala muda e todos os quadros sem fala ativa. Um quadro sonoro compensatório pode ser também classificado como *MUDO* se seu final tende a ser mudo e o encobrimento designado para quadros mudos pode ser usado para o quadro a seguir no caso que o mesmo seja perdido.
- 10 - classe *TRANSIÇÃO DE MUDO* compreende quadros mudos com um possível início sonoro no final. O início sonoro é entretanto ainda muito curto ou não suficientemente bem construído para usar o encobrimento projetado para quadros sonoros. Um quadro de *TRANSIÇÃO DE MUDO* pode seguir apenas um quadro classificado como *MUDO* ou de *TRANSIÇÃO*
- 15 *DE MUDO*.
- classe *TRANSIÇÃO SONORA* compreende quadros sonoros com características sonoras relativamente fracas. As mesmas são tipicamente quadros sonoros com características que mudam rapidamente (transições entre vogais) ou compensações sonoras durando todo o quadro. Um
- 20 quadro de transição sonoro pode seguir apenas um quadro classificado como *TRANSIÇÃO SONORA*, *SONORO* ou *INÍCIO*.
- classe *SONORO* compreende quadros sonoros com características estáveis. Um *quadro sonoro* pode seguir apenas um quadro classificado como *TRANSIÇÃO SONORA*, *SONORO* ou *INÍCIO*.
- 25 - classe *INÍCIO* compreende todos os quadros sonoros com características estáveis em seguida de um quadro classificado como *MUDO* ou *TRANSIÇÃO DE MUDO*. Quadros classificados como *INÍCIO* correspondem a quadros sonoros de início onde o início já é suficientemente bem construído para o uso do encobrimento projetado para quadros sonoros perdidos. As
- 30 técnicas de encobrimento usadas para um apagamento do quadro em seguida de um quadro classificado como *INÍCIO* são em codecs tradicionais com base em CELP os mesmos em seguida de um quadro classificado co-

mo *SONORO*, a diferença sendo em uma estratégia de recuperação quando uma técnica especial pode ser usada para reconstruir artificialmente o início perdido. De acordo com a modalidade ilustrativa não-restritiva da presente invenção, a técnica de codificação *TM* é usada com sucesso neste caso.

5 O diagrama de classificação de estado é delineado na figura 2. A informação de classificação é transmitida usando 2 bits. Como pode ser visto a partir da figura 2, a classe de *TRANSIÇÃO DE MUDO* e a classe de *TRANSIÇÃO SONORA* podem ser agrupadas juntas na medida em que as mesmas podem ser inequivocamente diferenciadas no decodificador (um quadro de *TRANSIÇÃO DE MUDO* pode seguir apenas *MUDO* ou quadros mudos de transição, um *Quadro de transição sonoro* pode seguir apenas *INÍCIO*, *SONORO* ou *Quadros sonoros de transição*).

10

Os parâmetros a seguir são usados para a classificação: uma correlação normalizada \bar{R}'_{xy} , uma medida de inclinação espectral e'_t , um contador de estabilidade do passo pc , uma energia de quadro relativa do sinal de fala no final do quadro atual E_{rel} , e um contador de zero-cruzado zc . Como pode ser visto na análise detalhada a seguir, a computação dos referidos parâmetros utiliza uma análise antecipada. A análise antecipada permite que a evolução do sinal de fala no quadro a seguir seja estimada e, con-

15

20

sequentemente, a classificação pode ser realizada ao se levar em consideração o comportamento futuro do sinal de fala.

A correlação média normalizada \bar{R}'_{xy} é computada como uma média da correlação máxima normalizada do segundo meio-quadro e a análise antecipada usando a equação a seguir:

$$\bar{R}'_{xy} = 0.5(C_{norm}(d_1) + C_{norm}(d_2)) \quad (8)$$

25 As correlações máximas normalizadas C_{norm} são computadas como a parte da pesquisa de passo de alça aberta e corresponde às correlações máximas normalizadas de dois períodos de passos adjacentes do sinal de fala pesado.

O parâmetro de inclinação espectral e'_t contém a informação so-

bre frequência de distribuição de energia. A inclinação espectral para uma análise espectral é estimada como uma proporção entre a energia concentrada em baixas frequências e a energia concentrada em altas frequências. Aqui, a medida de inclinação usada é a média no domínio logarítmico das medidas de inclinação espectrais $e_{tilt}(0)$ e $e_{tilt}(1)$ definidas como a proporção de energia de baixa e alta frequências. Que é:

$$e'_{t=10} = 10 \log(e_{tilt}(0) e_{tilt}(1)) \quad (9)$$

O contador de estabilidade do passo pc avalia a variação do período de passo. O mesmo é computado como a seguir:

$$pc = |T_{op1} - T_{op0}| + |T_{op2} - T_{op0}| \quad (10)$$

Os valores T_{op0} , T_{op1} , e T_{op2} correspondem às estimativas de passo de alça aberta a partir do primeiro meio do quadro atual, o segundo meio do quadro atual e a análise antecipada, respectivamente.

A energia de quadro relativa E_{rel} é computada como a diferença em dB entre a medida de energia de quadro atual e a energia de fala ativa de longo prazo.

O último parâmetro é o parâmetro zero-cruzado zc computado em um segmento de 20 ms do sinal de fala. O segmento se inicia no meio do quadro atual e utiliza dois subquadros da análise antecipada. Aqui, o contador de zero-cruzado zc conta o número de vezes que o sinal de fala muda a partir de positivo para negativo durante aquele intervalo.

Para tornar a classificação mais robusta, os parâmetros de classificação são considerados juntos formando uma função de mérito f_m . Para este fim, os parâmetros de classificação são primeiro escalados entre 0 e 1 de modo que o valor de parâmetro típico para sinais de fala mudos se transforma em 0 e cada o valor de parâmetro típico para sinais de fala sonoros se transforma em 1. Uma função linear é usada entre os mesmos. A versão em escala p^s de um certo parâmetro p_x é obtida usando a equação:

$$P^s = k_p P_x + c_p \text{ limitado por } 0 \leq p^s \leq 1. \quad (11)$$

Os coeficientes de função k_p e c_p foram encontrados experimentalmente para cada um dos parâmetros de modo que a distorção do sinal em

virtude das técnicas de encobrimento e recuperação usadas na presença de erros de quadro é mínima. Os valores usados são resumidos na Tabela 1.

Tabela 1 - Parâmetros de classificação de sinal e os coeficientes de suas respectivas funções de escala.

Parâmetro	Significado	k_p	C_p
\bar{R}'_{xy}	Correlação normalizada	2,857	-1,286
$e't$	Inclinação espectral	0,04167	0
pc	Contador de estabilidade do passo	-0,07143	1,857
E_{rel}	Energia de quadro relativa	0,05	0,45
ZC	Contador de zero cruzado	-0,04	2,4

5 Então a função de mérito f^m foi definida como:

$$f_m = \frac{1}{6} (2\bar{R}'_{xy} + e't + pc + E_{rel} + zc), \quad (12)$$

Onde os sobrescritos indicam a versão em escala dos parâmetros.

Uma primeira decisão de classificação é realizada para a classe MUDO como a seguir:

Se ($local_VAD = 0$) OU ($E_{rel} < -8$) então a classe = MUDO (13)

10 onde $local_VAD$ representa detecção local de atividade de voz.

Se a condição acima (13) não for satisfeita, então a classificação prossegue usando a função de mérito f_m e seguindo as regras resumidos na Tabela 2.

Tabela 2 - Regras de classificação de sinal no codificador.

Classe de quadros anteriores	Regra	Classe de quadros atuais
INÍCIO	$f_m \geq 0,66$	SONORO
SONORO	$0,66 > f_m \geq 0,49$	TRANSIÇÃO SONORA
TRANSIÇÃO SONORA	$f_m < 0,49$	MUDO
TRANSIÇÃO DE MUDO	$f_m \geq 0,63$	INÍCIO
	$0,63 \geq f_m > 0,585$	TRANSIÇÃO DE MUDO
MUDO	$f_m \leq 0,585$	MUDO

15

A informação de classe é codificada com dois bits como explica-

do aqui acima. Apesar do fato de que a informação suplementar, que aprimora o encobrimento do apagamento do quadro, ser transmitida apenas em quadros genéricos, a classificação é realizada para cada quadro. Isto é necessário para manter a máquina de classificação de estado atualizada na medida em que a mesma utiliza a informação sobre a classe do quadro anterior. A classificação é, entretanto direta para codificar tipos dedicados ao *MUDO* ou *Quadros sonoros*. Consequentemente, quadros sonoros são sempre classificados como *SONORO* e quadros mudos são sempre classificados como *MUDO*.

10 Seleção de Quadro para classificação *TM*

Como discutido anteriormente, a técnica sendo descrita substitui o livro de codificação adaptativo em Codificadores com base em CELP por um livro de codificação de formato glótico para aprimorar a resistência ao apagamento do quadros e para aumentar a eficiência da codificação quando quadros de fala não-estacionários são processados. Isto significa que a referida técnica não constroi o sinal de excitação de primeiro estágio com o uso da excitação passada, mas seleciona o sinal de excitação de primeiro estágio a partir do livro de codificação de formato glótico. O sinal de excitação de segundo estágio (a parte de inonvação da excitação total) é ainda selecionado a partir do livro de codificação fixo CELP tradicional. Qualquer dos referidos livros de codificação não usam informação a partir dos quadros de fala anteriores (anteriormente transmitidos), deste modo eliminando a razão principal para a propagação de erro de quadro inerente aos codificadores com base em CELP.

Usando a técnica de codificação *TM* sistematicamente (para codificar todos os quadros) limitará grandemente a propagação de erro, mas a eficiência da codificação e a qualidade da fala sintetizada cairiam em condições livres de erro. Como um compromisso entre o desempenho de canal limpo do Codec e a sua resistência para erros de canal, a técnica de codificação *TM* pode ser aplicada apenas aos quadros de transição e a diversos quadros em seguida de cada quadro de transição. Para apagamento do quadro resistência, a técnica de codificação *TM* pode ser usada para qua-

dros de fala sonora em seguida de transições. Como introduzido anteriormente, as referidas transições compreendem basicamente o início sonoro e as transições entre dois diferentes sons sonantes. Para selecionar quadros pertinentes a serem codificados usando a técnica de codificação *TM*, transições são detectadas. Embora qualquer detector de transições possa ser usado, a modalidade ilustrativa não-restritiva utiliza a classificação de estrutura *EV-VBR* como descrito aqui acima.

A técnica de codificação *TM* pode ser aplicada para codificar quadros de transição (início sonoro ou transição entre dois diferentes sons sonantes) como descritos acima e diversos quadros subsequentes. O número de quadros *TM* (quadros codificados usando a técnica de codificação *TM*) é uma questão de compromisso entre o desempenho do Codec em condições de canal limpo e em condições com erros de canal. Se apenas os quadros de transição (início sonoro ou transição entre dois diferentes sons sonantes) são codificados usando a técnica de codificação *TM*, a eficiência da codificação aumenta. O referido aumento pode ser medido pelo aumento da proporção de sinal de sinal de segmento-para-ruído (*SNR*), por exemplo. O *SNR* é computado usando a equação a seguir:

$$SNR = \frac{E_{sd}}{E_e} \quad (14)$$

onde E_{sd} é a energia do sinal de entrada de fala do quadro atual e E_e é a energia do erro entre o referido sinal de fala de entrada e o sinal de fala de síntese do quadro atual.

Entretanto, usando a técnica de codificação *TM* para codificar apenas os quadros de transição não ajuda muito para resistência a erro; se o quadro de transição (início sonoro ou transição entre dois diferentes sons sonantes) for perdido, o erro se propagará na medida em que o quadro a seguir seria codificado usando o procedimento CELP padrão. Por outro lado, se o quadro precedendo o quadro de transição (início sonoro ou transição entre dois diferentes sons sonantes) for perdido, o efeito do referido quadro precedente perdido no desempenho não é fundamental mesmo sem o uso

da técnica de codificação *TM*. Em caso de transições de início sonoro, o quadro precedendo o início é provável ser mudo e a contribuição do livro de codificação adaptativo não é muito importante. Em caso de uma transição entre dois sons sonoros, o quadro antes da transição é geralmente pobremente estacionário e o estado do livro de codificação adaptativo no codificador e o decodificador é com frequência similar após o apagamento do quadro.

Para aumentar a resistência, quadros seguindo a transição (início sonoro ou transição entre dois diferentes sons sonantes) podem ser codificados usando a técnica de codificação *TM*. Se o aumento do desempenho de canal limpo não for importante, a técnica de codificação *TM* pode ser usada apenas nos quadros seguindo os quadros de transição. Basicamente, o número de quadros *TM* consecutivos depende do número de apagamento de quadros consecutivos que se quer considerar para proteção. Se apenas apagamentos isolados são considerados (isto é, um apagamento isolado do quadro de cada vez), é suficiente se codificar apenas o quadro seguindo o quadro de transição (início sonoro ou transição entre dois diferentes sons sonantes). Se o quadro de transição (início sonoro ou transição entre dois diferentes sons sonantes) for perdido, o quadro a seguir é codificado sem o uso de sinal de excitação passada e a propagação de erro é interrompida. Deve ser observado, entretanto, que se um quadro de transição (início sonoro ou transição entre dois diferentes sons sonantes) for transmitido corretamente mas o quadro a seguir for perdido, a propagação de erro não será evitada na medida em que o próximo quadro já está usando um codificador CELP clássico. Entretanto, a distorção será provavelmente limitada se pelo menos um período de passo já for construído no final da transição (início sonoro ou transição entre dois diferentes sons sonantes) como mostrado na figura 3.

Quando uma implementação da técnica de codificação *TM* é realizada em algum Codec existente e a classe do quadro atual e modo de codificação são conhecidos, o esquema a seguir para ajustar o início e o quadro a seguir para *Codificação TM* pode ser usado. Um estado de parâmetro que é a contagem dos quadros *TM* consecutivos anteriormente usa-

dos é armazenado na memória de estado do codificador. Se o valor do referido estado de parâmetro for negativo, a *codificação TM* não pode ser usada. Se o estado de parâmetro não for negativo mas menor ou igual ao número de apagamentos de quadros consecutivos a proteger, e a classe do quadro é *INÍCIO*, *SONORO* ou *TRANSIÇÃO SONORA*, o quadro é estipulado como *Quadro TM* (vide figura 4 para mais detalhes). Em outras palavras, o quadro é estipulado como *Quadro TM* se $QUADROS\ N_TM_ \geq estado > 0$, onde *QUADROS N_TM_* é uma série de quadros consecutivos para evitar o uso da técnica de codificação *TM*.

Se for esperado que as características do canal de comunicação sejam tais que mais de um quadro isolado é com frequência apagado de cada vez, isto é, o apagamento do quadros é dotado da tendência de aparecer em feixes, a melhor solução deve ser o uso de uma técnica de codificação *TM* para proteger dois ou mesmo mais apagamentos de quadro consecutivos. Entretanto, a eficiência da codificação em condições de canal limpo irá cair. Se informação de retorno sobre o canal estiver disponível no codificador, o número de quadros *TM* consecutivos deve ser tornado adaptivo para as condições de transmissão. Na modalidade ilustrativa não-restritiva da presente invenção, cerca de dois quadros *TM* seguindo um quadro de transição (início sonoro ou transição entre dois diferentes sons sonantes) são considerados, o que corresponde a uma configuração capaz de lidar com cerca de dois apagamentos de quadros consecutivos.

A decisão descrita utiliza basicamente um número fixo (seja o referido número fixado antes da transmissão ou dependente das condições de transmissão de canal) de quadros *TM* seguindo um quadro de transição (início sonoro ou transição entre dois diferentes sons sonantes). O compromisso entre o desempenho de canal limpo e a resistência a erro do quadro pode ser também baseado em uma classificação de alça fechada. Mais especificamente, no quadro que se quer proteger contra apagamento anterior do quadro ou se quer decidir se é o quadro de início, uma computação de dois possíveis modos de codificação é realizada em paralelo; o quadro é processado não só usando o modo de codificação genérico (*CELP*) mas

também a técnica de codificação *TM*. O desempenho de ambas as abordagens é então comparado usando uma medição *SNR*, por exemplo; para mais detalhes vide a seção a seguir intitulada "Desempenho de Técnica de codificação *TM* em Codec EV-VBR". Quando a diferença entre o *SNR* for o modo de codificação genérico (*CELP*) e o *SNR* for a técnica de codificação *TM* forem maior do que um determinado limiar, o modo de codificação genérico (*CELP*) é aplicado. Se a diferença entre o *SNR* for o modo de codificação genérico (*CELP*) e o *SNR* for a técnica de codificação *TM* forem menor do que um determinado limiar, a técnica de codificação *TM* é aplicada. O valor do limiar é escolhido dependendo do quão forte a proteção de apagamento do quadro e determinação de codificação de início é necessário.

Seleção de subquadro para pesquisa de livro de codificação de formato glótico

Na seção anterior, as razões e os mecanismos para selecionar quadros para codificação usando a técnica de codificação *TM* foram descritas. Agora será mostrado que é em geral mais eficiente não se usar o livro de codificação de formato glótico em todos os subquadros de modo a alcançar o melhor compromisso entre o desempenho do canal limpo em um determinado bitrato e o desempenho na presença de um apagamento nos quadros precedendo os quadros *TM*. Primeiro, a pesquisa do livro de codificação de formato glótico é importante apenas no primeiro período de passo em um quadro. O período de passos a seguir pode ser codificado usando a pesquisa mais eficiente do livro de codificação adaptativo padrão uma vez que os mesmos não usam mais a excitação de quadro passado (quando o livro de codificação adaptativo é pesquisado, a excitação é pesquisada até cerca de um período de passo no passado). Não há conseqüentemente nenhuma razão para empregar a pesquisa do livro de codificação de formato glótico em subquadros contendo nenhuma porção do primeiro período de passo de um quadro.

De modo similar, quando a pesquisa do livro de codificação de formato glótico é usada para aumentar a eficiência da codificação em quadros sonoros de início, a referida pesquisa de livro de codificação de formato glótico é usada no primeiro período de passo do segmento sonoro de parti-

da. A razão é que para o primeiro período de passo, o livro de codificação adaptativo contém um sinal como ruído (o segmento anterior foi não sonoro) e substituindo o mesmo com um impulso glótico quantificado, com frequência aumenta a eficiência da codificação. Para o período de passos a seguir, 5 entretanto, a excitação periódica já foi construída no livro de codificação adaptativo e usando o referido livro de codificação renderá melhores resultados. Por esta razão, a informação na posição de início sonoro é disponível pelo menos com resolução de subquadro.

A otimização adicional da alocação de bits se refere a quadros 10 com período de passos mais longos do que o comprimento do subquadro. Considerando que o livro de codificação de formato glótico contém formatos quantificados de impulso glótico, o livro de codificação é melhor adequado para ser usado em subquadros contendo o impulso glótico. Em outros subquadros, a sua eficiência é baixa. Considerando que o coeficiente de bit é 15 com frequência relativamente limitado em aplicações de codificação de fala e que a codificação do livro de codificação de formato glótico requer um número relativamente maior de bits para codificação de fala de baixo coeficiente de bit, a alocação de bits onde o livro de codificação de formato glótico é usado e pesquisado apenas em um subquadro por quadro foi escolhida na 20 modalidade ilustrativa não-restritiva.

Para se escolher o subquadro a ser codificado com o livro de codificação de formato glótico, o primeiro impulso glótico no sinal LP residual é procurado. O procedimento simples a seguir pode ser usado. A amostra máxima no sinal LP residual é pesquisada na faixa $[0, 0 + T_{op} + 2]$, onde T_{op} é 25 o período de passo de alça aberta para o primeiro meio-quadro e 0 corresponde ao início do quadro. Em caso de quadros sonoros de início, e se o início do início pode ser determinado de modo confiável, 0 denota o início do subquadro onde o início do início é localizado. O livro de codificação de formato glótico será então empregado no subquadro com a energia de sinal residual máxima. Ademais, a posição do máximo proporciona informação 30 onde a posição do impulso glótico pode aproximadamente ser situada e a referida situação pode ser explorada para redução de complexidade como

será discutido posteriormente. Observe que na medida em que a pesquisa do livro de codificação de formato glótico substitui apenas a pesquisa do livro de codificação adaptativo, uma pesquisa de livro de codificação fixo é realizada em cada subquadro de um quadro *TM*.

5 Os outros subquadros (não codificados com o uso do o livro de codificação de formato glótico) serão processados como a seguir. Se o subquadro usando a pesquisa de livro de codificação de formato glótico não for o primeiro subquadro no quadro, o sinal de excitação no(s) subquadro(s) precedente(s) do quadro é codificado usando apenas o livro de codificação
 10 CELP fixo; isto quer dizer que o sinal de excitação de primeiro estágio é zero. Se o subquadro do livro de codificação de formato glótico não for o último subquadro no quadro, o(s) subquadro(s) a seguir do quadro é/são processado(s) usando codificação CELP padrão (isto é, usando a pesquisa do livro de codificação fixo e adaptativa). Nas figuras 5a - 5c, a situação é mostrada
 15 para o caso onde o primeiro impulso glótico emerge no segundo subquadro. Na figura 5b, $u(n)$ é o sinal LP residual. O sinal de excitação de primeiro estágio é denotado $q_k(n)$ quando o mesmo é construído usando o livro de codificação de formato glótico, ou $v(n)$ quando o mesmo é construído usando o livro de codificação adaptativo. Neste exemplo (figura 5c), o sinal de excitação de primeiro estágio é zero no primeiro subquadro, o mesmo é um vetor-código de formato glótico no segundo subquadro e um vetor do livro de codificação adaptativo nos últimos dois subquadros.
 20

De modo a adicionalmente aumentar a eficiência da codificação e a otimizar a alocação de bits, diferentes processamentos são usados em
 25 subquadros particulares de um *quadro TM* dependente do período de passo. Quando o primeiro subquadro é escolhido como subquadro *TM*, o subquadro com o segundo impulso glótico no sinal LP residual é determinado. A referida determinação é baseada no valor de período de passo e as seguintes quatro situações podem então ocorrer. Na primeira situação, o segundo impulso glótico está no primeiro subquadro, e o segundo, terceiro e quarto subquadros são processados usando codificação CELP padrão (pesquisa de
 30 livro de codificação fixo e adaptativo). Na segunda situação, o segundo im-

pulso glótico está no segundo subquadro, e o segundo, terceiro e quarto subquadros são processado usando mais uma vez a codificação CELP padrão. Na terceira situação, o segundo impulso glótico está no terceiro subquadro. O segundo subquadro é processado usando a pesquisa do livro de codificação fixo apenas na medida em que não há impulso glótico no segundo subquadro do sinal LP residual a ser pesquisado para usar o livro de codificação adaptativo. Os terceiro e quarto subquadros são processados usando codificação CELP padrão. Na última (quarta) situação, o segundo impulso glótico está no quarto, subquadro (ou no próximo quadro), os segundo e terceiro subquadros são processados usando apenas a pesquisa do livro de codificação fixo, e o quarto, subquadro é processado usando codificação CELP padrão. Uma discussão mais detalhada é proporcionada em uma implementação exemplificativa posteriormente abaixo.

A tabela 3 mostra nomes das possíveis configurações de código e suas ocorrências estatísticas. Em outras palavras, a Tabela 3 proporciona a distribuição da primeira e da segunda ocorrência de impulso glótico em cada subquadro para quadros processados com a técnica de codificação *TM*. A Tabela 3 corresponde ao cenário onde a técnica de codificação *TM* é usada para codificar apenas o quadro de início sonoro e um quadro subsequente. O comprimento do quadro do sinal de fala neste experimento foi 20 ms, o comprimento do subquadro 5 ms e o experimento foi conduzido usando vozes de 32 homens e 32 mulheres (se não mencionado de modo diferente, o mesmo banco de dados de fala foi usado também em todos os outros experimentos mencionados na descrição a seguir).

Tabela 3 – Configurações de modo de codificação para TM e sua ocorrência quando o sinal de fala é processado.

Código de Configuração	Posição(ões) do primeiro (e segundo, se relevante) impulso global	Tipo de "livro de codificação" usado (GS = "formato glótico", A = adaptativo, F = fixo)				Quantidade [%]
		1º subfr.	2º subfr.	3º subfr.	4º subfr.	
TRANSIÇÃO_1_1		GS+F	A+F	A+F	A+F	25,5
TRANSIÇÃO_1_2		GS+F	A+F	A+F	A+F	28,4
TRANSIÇÃO_1_3		GS+F	F	A+F	A+F	16,3
TRANSIÇÃO_1_4		GS+F	F	F	A+F	3,0
TRANSIÇÃO_2		F	GS+F	A+F	A+F	21,2
TRANSIÇÃO_3		F	F	GS+F		4,6
TRANSIÇÃO_4		F	F	F	GS+F	1,0

Livro de codificação de formato glótico

Em princípio, o livro de codificação de formato glótico consiste em formatos quantificados normalizados de impulsos glóticos dispostos em uma posição específica. Consequentemente, a pesquisa do livro de codificação consiste tanto na seleção do melhor formato, e na determinação de sua melhor posição em um subquadro particular. Em sua forma mais simples, o formato de impulso glótico pode ser representado por um impulso unitário e não precisa ser quantificado. Neste caso, apenas sua posição no subquadro é determinada. Entretanto o desempenho de um referido livro de codificação simples é bastante limitado.

Por outro lado, a melhor representação seria provavelmente alcançada se o comprimento L das entradas do livro de codificação de formato glótico corresponde ao comprimento do período de passo, e se um grande número de formatos de impulsos glóticos é representado. Na medida em que o comprimento e o formato dos impulsos glóticos variam a partir de orador a orador e a partir de quadro a quadro, a complexidade e necessidades de memória para pesquisar e armazenar o referido livro de codificação seria muito extensa. Como um compromisso, o comprimento dos impulsos glóticos assim como o número dos mesmos deve ser limitado. Em uma modalidade ilustrativa não-restritiva, o livro de codificação de formato glótico é composto de oito (8) diferentes formatos de impulsos glóticos e o comprimento de cada impulso glótico é $L = 17$ amostras. Os formatos quantificados foram selecionados de modo que a máxima absoluta é em torno do meio do seu comprimento. Durante a pesquisa do livro de codificação de formato glótico, o referido meio é alinhado com o índice W , que representa a posição de impulso glótico no subquadro atual e é escolhido a partir do intervalo $[0, N - 1]$, N sendo o comprimento do subquadro. Na medida em que o comprimento das entradas do livro de codificação de 17 amostras é mais curto do que o comprimento do subquadro, as amostras restantes são ajustadas a zero.

O livro de codificação de formato glótico é projetado para representar tantos impulsos glóticos existentes quanto possível. Um processo de treinamento baseado no algoritmo *média-k* [4] foi usado; o livro de codifica-

ção de formato glótico foi treinado usando mais do que três (3) horas de sinal de fala composta de discurso de muitos oradores diferentes falando em diversos idiomas diferentes. A partir do referido banco de dados, os impulsos glóticos foram extraídos a partir do sinal LP residual e truncados nas 17 amostras em torno do valor máximo absoluto. A partir de dezesseis (16) formatos selecionados pelo algoritmo de *média-k*, o número de formatos foi adicionalmente reduzido a oito (8) formatos experimentalmente usando uma medição de qualidade l segmentário. O livro de codificação de formato glótico selecionado é mostrado na figura 6. Obviamente, outros meios podem ser usados para configurar o livro de codificação de formato glótico.

Pesquisa de livro de codificação de formato glótico

A implementação atual do livro de codificação de formato glótico pode ser realizada em diversas formas. Por exemplo, a pesquisa pode ser realizada similar à pesquisa do livro de codificação fixo em *CELP*. Neste caso o livro de codificação é construído ao dispor o centro dos formatos de impulsos glóticos em todas as posições possíveis no subquadro. Por exemplo, para um comprimento de subquadro de sessenta e quatro (64) amostras e oito (8) formatos de impulsos glóticos, um livro de codificação de formato glótico de tamanho $64 \times 8 = 512$ vetores-código é obtido. De acordo com outro exemplo, de modo similar à pesquisa do livro de codificação adaptativo, as entradas do livro de codificação podem ser sucessivamente dispostas em todas as posições potenciais em uma excitação passada e a melhor combinação de formato/posição pode ser selecionada de modo similar ao que é usado na pesquisa do livro de codificação adaptativo. Na última realização toda a repetição do ciclo de passos é automaticamente realizada através do filtro *CELP* de longo prazo e os impulsos glóticos são representados com formatos de tamanho total (diferente da primeira realização onde a truncagem de formato glótico é necessária em casos de delimitação como será discutido posteriormente).

A modalidade ilustrativa não-restritiva utiliza a configuração onde a pesquisa do livro de codificação é similar à pesquisa do livro de codificação fixo em *CELP* algébrica (*ACELP*). Na referida abordagem, para cada um dos formatos candidatos, o formato é representado como um impulso de resposta

de um filtro de forma $G(z)$. Assim os vetores-código correspondendo aos formatos de impulsos glóticos centrados em diferentes posições podem ser representados por vetores-código contendo apenas um elemento não-zero filtrado através do filtro de forma $G(z)$ (para um tamanho de subquadro N há N vetores de pulso único para potenciais posições de impulso glótico k).

Em virtude do fato da posição do impulso glótico W estar no meio do formato glótico com um comprimento ímpar de L amostras e k' é a partir da faixa de $[0, AM]$, o formato glótico deve ser truncado para o primeiro e para o último $L_{1/2} = (L - 1) / 2$ amostras. Isto será levado em consideração durante a pesquisa de pulso glótico uma vez que o mesmo torna o filtro de forma $G(z)$ um filtro não-causal.

A configuração da parte TM é mostrada na figura 7 para o codificador e na figura 11 para o decodificador. Como já mencionado, a parte TM substitui a parte do livro de codificação adaptativo do codificador/decodificador. Durante a pesquisa, a resposta de impulso do filtro de forma $G(z)$ pode ser integrada à resposta de impulso do filtro $H(z)$.

Um procedimento e pesquisa do livro de codificação correspondente para pesquisar a posição central ótima de impulso glótico k' para um determinado formato de impulso glótico proporcionado pelo filtro de forma $G(z)$ serão agora descritos. Em virtude do formato do filtro $G(z)$ ser escolhido a partir de diversos formatos candidatos (oito (8) formatos são usados em uma modalidade ilustrativa não-restritiva como ilustrado na figura 6), um procedimento de pesquisa deve ser repetido para cada formato glótico do livro de codificação de modo a encontrar formato de impulso e posição otimizados.

Para determinar os parâmetros de codificação TM , a pesquisa determina o erro de quadrados mínimos entre o vetor alvo X_1 e o vetor-código de formato glótico centrados na posição k' que é filtrado através de um filtro de síntese ponderado $H(z)$. Similar ao *CELP*, a pesquisa pode ser realizada ao se encontrar a máxima de um critério na forma:

$$\mathfrak{J}_{k'} = \frac{(\mathbf{x}_1^T \mathbf{y}_1)^2}{\mathbf{y}_1^T \mathbf{y}_1} \quad (15)$$

onde y_1 é o vetor-código de formato glótico filtrado. Considerando que q_k denota o vetor-código de formato glótico centrado na posição k' e p_k a vetor-código de posição com um (1) elemento não-zero indicando a posição k' então q_k pode ser escrito como $q_k = G \cdot p_k$, onde G é a matriz de Toeplitz que representa o formato de impulso glótico. Portanto, similar à pesquisa do livro de codificação fixo, a equação a seguir pode ser escrita:

$$\begin{aligned} \mathfrak{J}_{k'} &= \frac{(\mathbf{x}_1^T \mathbf{y}_1)^2}{\mathbf{y}_1^T \mathbf{y}_1} = \frac{(\mathbf{x}_1^T \mathbf{H} \mathbf{q}_{k'})^2}{\mathbf{q}_{k'}^T \mathbf{H}^T \mathbf{H} \mathbf{q}_{k'}} = \frac{(\mathbf{x}_1^T \mathbf{H} \mathbf{G} \mathbf{p}_{k'})^2}{\mathbf{p}_{k'}^T \mathbf{G}^T \mathbf{H}^T \mathbf{H} \mathbf{G} \mathbf{p}_{k'}} = \\ &= \frac{(\mathbf{x}_1^T \mathbf{Z} \mathbf{p}_{k'})^2}{\mathbf{p}_{k'}^T \mathbf{Z}^T \mathbf{Z} \mathbf{p}_{k'}} = \frac{(\mathbf{d}_g^T \mathbf{p}_{k'})^2}{\mathbf{p}_{k'}^T \mathbf{\Phi}_g \mathbf{p}_{k'}} = \frac{(C_{k'})^2}{E_{k'}}, \end{aligned} \quad (16)$$

onde H é matriz de convolução Toeplitz triangular inferior do filtro de síntese pesado. Como será discutido posteriormente, as fileiras da matriz \mathbf{Z}^T correspondem à versão desviada filtrada do formato de impulso glótico ou a sua representação truncada. Observe que todos os vetores neste texto são supostos vetores de coluna (matrizes $N \times 1$).

Um exemplo de matriz G na forma transposta (\top) para um comprimento de impulso de três (3) amostras e $N = 4$ teria a forma:

$$\mathbf{G}^T = \begin{bmatrix} g(0) & g(1) & 0 & 0 \\ g(-1) & g(0) & g(1) & 0 \\ 0 & g(-1) & g(0) & g(1) \\ 0 & 0 & g(-1) & g(0) \end{bmatrix} \quad (17)$$

onde $g(n)$ são as respostas dos coeficientes do impulso do filtro de forma não-causal $G(z)$. Na descrição a seguir, os coeficientes de filtro de forma não causal $G(z)$ são oferecidos pelos valores $g(n)$, para n localizado dentro da faixa $[-L_{1/2}, L_{1/2}]$. Pelo fato de que um vetor-código de posição p_k ser dotado apenas de um elemento não-zero, a computação do critério (16) é bastante simples e pode ser expressa usando a equação a seguir:

$$\mathfrak{J}_{k'} = \frac{(d_g(k'))^2}{\Phi_g(k', k')} \quad (18)$$

Como pode ser visto a partir de Equação (18), apenas a diagonal da matriz Φ_g precisa ser computada.

Uma representação gráfica de computação de o critério (18) para um vetor-código de formato glótico é mostrada nas figuras 8 e 9. Como já foi anteriormente mencionado, a equação (18) é tipicamente usada na pesquisa do livro de codificação algébrico *ACELP* ao pré-computar o vetor alvo filtrado traseiro retrógrado d_g e a matriz de correlação Φ_g . Entretanto, considerando a natureza não-causal do filtro de forma $G(z)$, isto não pode ser diretamente aplicado para as primeiras posições $L_{1/2}$. Nas referidas situações uma pesquisa mais sofisticada é usada onde alguns valores computados podem ainda ser reutilizados para manter a complexidade em um nível baixo. Isto será descrito daqui adiante.

Denota-se z_k como a fileira $(k'+1)^{th}$ da matriz \mathbf{Z}^T , onde a matriz \mathbf{Z}^T (figura 10) é computada como a seguir. Considerando a a natureza não-causal do filtro de forma $G(z)$, a matriz \mathbf{Z}^T é computada em dois estágios para minimizar a complexidade computacional. As primeiras fileiras $L_{1/2} + 1$ da referida matriz são primeiro computadas. Para a parte restante da matriz \mathbf{Z}^T (a última fileira $N - L_{1/2} - 1$ da matriz \mathbf{Z}^T), o critério (18) é usado em uma maneira similar à pesquisa do livro de codificação fixo *ACELP*.

Uma descrição detalhada de como computar a matriz \mathbf{Z}^T e o critério (18) será agora descrita.

No primeiro estágio, as primeiras fileiras $L_{1/2} + 1$ da matriz \mathbf{Z}^T que correspondem às posições k' dentro da faixa $[0, L_{1/2}]$ são computadas. Para as referidas posições um diferente formato glótico truncado é usado para cada posição k' dentro da referida faixa. Em uma primeira operação, a convolução entre a resposta de formato glótico para a posição $k' = 0$ e a resposta de impulso $h(n)$ é computada usando a equação:

$$z_0(n) = \sum_{i=0}^n g(n-i)h(i) \quad (19)$$

onde se tem a vantagem do fato de que o filtro de forma $G(z)$ é dotado apenas de $L_{1/2} + 1$ coeficientes não-zero, isto é, $g(0), g(1), \dots, g(L_{1/2})$ são coeficientes não-zero.

Na segunda operação, a convolução $z_1(n)$ entre a resposta de livro de codificação de formato glótico para a posição $k' = 1$ e a resposta de impulso $H(z)$ é computada reutilizando valores de $z_0(n)$ como a seguir (a matriz $\mathbf{Z}^T = \mathbf{G}^T \mathbf{H}^T$ é a matriz com algumas diagonais de inclinação zero negativo, mas a referida matriz \mathbf{Z}^T não é mais uma matriz Toeplitz e triangular como mostrado na figura 10):

$$z_1(0) = g(-1)h(0)$$

$$z_1(n) = z_0(n-1) + g(-1)h(n) \text{ for } n = 1, \dots, N-1 \quad (20)$$

Para as fileiras a seguir, a recursão na Equação (21) é reutilizada:

$$z_{k'}(0) = g(-k')h(0)$$

$$z_{k'}(n) = z_{k'-1}(n-1) + g(-k')h(n) \text{ for } n = 1, \dots, N-1 \quad (21)$$

A recursão (21) é repetida para todos $k' \leq L_{1/2}$. Para $k' = L_{1/2}$ o filtro de forma $G(z)$ já é dotado de L coeficientes não-zero e a fileira $(L_{1/2}+1)^o$ da matriz \mathbf{Z}^T é assim obtida pela

$$z_{L_{1/2}}(0) = g(-L_{1/2})h(0)$$

$$z_{L_{1/2}}(n) = z_{L_{1/2}-1}(n-1) + g(-L_{1/2})h(n) \text{ for } n = 1, \dots, N-1 \quad (22)$$

Neste ponto, as primeiras fileiras $L_{1/2} + 1$ da matriz \mathbf{Z}^T foram computadas. As referidas fileiras compreendem coeficientes não-zero (figura 10). Então o critério (18) pode ser computado para W dentro da faixa $[0, L_{1/2}]$ usando a equação:

$$\mathfrak{J}_{k'} = \frac{\left(\sum_{i=1}^{N-1} z_{k'}(i) \cdot x_1(i) \right)^2}{\sum_{i=0}^{N-1} z_{k'}(i) \cdot z_{k'}(i)} \quad (23)$$

No segundo estágio o resto da matriz \mathbf{Z}^T é computado e o critério (18) é avaliado para as posições k' dentro da faixa $[L_{1/2} + 1, N - 1]$. Se obtém vantagem do fato de que as fileiras $L_{1/2} + 1, \dots, N - 1$ da matriz \mathbf{Z}^T são construídas usando coeficientes da convolução $z_{L_{1/2}}(n)$ a qual já foi computada como descrito pela equação (22). A diferença é que apenas uma parte dos coeficientes é necessária para computar as referidas fileiras. Ou seja, cada fileira corresponde à fileira anterior deslocada para a direita em 1 e adicionando um zero no início:

$$z_{k'}(0) = 0$$

$$z_{k'}(n) = z_{k'-1}(n-1) \text{ for } n = 1, \dots, N-1 \quad (24)$$

Isto é repetido para k' dentro da faixa $[L_{1/2} + 1, N - 1]$.

No referido segundo estágio, o critério (18) pode ser computado em uma maneira similar àquela descrita na seção acima *Pesquisa de Livro de codificação fixo* para adicionalmente reduzir a complexidade computacional. O critério (18) é primeiro avaliado para a última posição $k' = N - 1$ (esta é a última fileira da matriz \mathbf{Z}^T). Para $k' = N - 1$ o numerador e o denominador do critério (18) são proporcionados pela Equação a seguir

$$d_g(N-1) = \sum_{i=0}^{L_{1/2}} x(N-1-L_{1/2}+i)z_{L_{1/2}}(i) \quad (25)$$

e

$$\Phi_g(N-1, N-1) = \sum_{i=0}^{L_{1/2}} z_{L_{1/2}}(i)z_{L_{1/2}}(i) \quad (26)$$

Uma vez que alguns dos coeficientes da matriz \mathbf{Z}^T são zeros (figura 10), apenas as multiplicações $L_{1/2} + 1$ (em vez das multiplicações N como usado na Equação (23)) são usadas para computar o numerador e o denominador do critério (18).

Quando se usa o exemplo da figura 10 ($L_{1/2} = 1$) o critério (18), computado usando as equações (25) e (26), pode ser simplificado como a seguir:

$$\mathfrak{J}_{N-1} = \frac{(d_g(N-1))^2}{\Phi_g(N-1, N-1)} = \frac{(x(N-2)z_{L_{1/2}}(0) + x(N-1)z_{L_{1/2}}(1))^2}{z_{L_{1/2}}(0)z_{L_{1/2}}(0) + z_{L_{1/2}}(1)z_{L_{1/2}}(1)} \quad (27)$$

5 Nas próximas etapas alguns dos valores anteriormente computados podem ser mais uma vez reutilizados para a computação do denominador. Para a posição $N-2$ o denominador do critério (18) é computado usando

$$\Phi_g(N-2, N-2) = \Phi_g(N-1, N-1) + z_{L_{1/2}}(L_{1/2}+1)z_{L_{1/2}}(L_{1/2}+1) \quad (28)$$

O numerador é computado usando a Equação (25) com o índice de soma mudado:

$$d_g(N-2) = \sum_{i=0}^{L_{1/2}+1} x(N-2-L_{1/2}+i)z_{L_{1/2}}(i) \quad (29)$$

10 Em uma maneira similar, o numerador e o denominador de critério (18) são calculados para todas as posições $k' > L_{1/2}$.

O procedimento acima descrito permite encontrar o máximo do critério (18) para vetores-código que representam o primeiro formato a partir dos impulsos glóticos. A pesquisa irá continuar usando o procedimento anteriormente descrito para todos os outros formatos de impulsos glóticos. A pesquisa do máximo de critério (18) continua como a pesquisa de livro de codificação de formato glótico para encontrar um valor máximo para o critério (18) que corresponde a um de formato glótico e um de posição k' constituindo o resultado da pesquisa.

20 É também possível se usar resolução de subamostra quando se pesquisa a posição central de pulso glótico k' isto irá, entretanto, resultar em maior complexidade. Mais especificamente, isto irá requerer o aumento do coeficiente de amostragem dos formatos de impulsos glóticos para aumentar a resolução e as diferentes versões de extração deslocadas em diferentes

resoluções. Isto é equivalente ao uso de um livro de codificação de formato glótico maior.

De modo ideal o critério (18) é computado para todas as possíveis posições de impulso glótico W . Em uma modalidade ilustrativa não-restritiva, a pesquisa é realizada apenas em uma faixa restrita em torno da posição esperada de uma posição W para adicionalmente reduzir a complexidade computacional. A referida posição esperada está na faixa $[k_{min}, k_{max}]$, $0 \leq k_{min} < k_{max} < N$, e pode ser determinada para o primeiro formato glótico a partir do sinal LP residual máximo encontrado como descrito na Seção acima *Seleção de Subquadro para a pesquisa de livro de codificação de formato glótico*. Uma pesquisa de livro de codificação de formato glótico é então realizada e a posição W é encontrada para o primeiro formato glótico. A nova faixa $[k_{min}, k_{max}]$ é ajustada para a pesquisa de segundo formato glótico como a seguir:

$$k_{max} = k' + \Delta$$

Tipicamente $\Delta = 4$. De modo similar, a Equação (30) é usada para definir a faixa de pesquisa para o terceiro terceiro formato em torno da posição selecionada do segundo formato e assim por diante.

No exemplo a seguir, é suposto que a faixa de pesquisa inicial é $[N - 15, N - 7]$, $L = 17$ e $N = 64$. A pesquisa se inicia com a computação do valor $z_{1/2}^l(n)$. Então o critério (18) para a posição $k' = N - 7$ é avaliado usando

$$\mathfrak{S}_{N-7} = \frac{(d_g(N-7))^2}{\Phi_g(N-7, N-7)} = \frac{\left(\sum_{i=0}^{L_{1/2}+6} x(N-7-L_{1/2}+i)z_{L_{1/2}}(i) \right)^2}{\sum_{i=0}^{L_{1/2}+7} z_{L_{1/2}}(i)z_{L_{1/2}}(i)} \quad (31)$$

Para computar o critério para a posição $k' = N - 8$, o denominador é repetidamente computado como:

$$\Phi_g(N-8, N-8) = \Phi_g(N-7, N-7) + z_{L_{1/2}}(L_{1/2}+8)z_{L_{1/2}}(L_{1/2}+8) \quad (32)$$

De mesmo modo, o denominador é computado para todas as posições restantes até que $k' = N - 15$. O numerador de critério (18) é computado para cada posição dentro da faixa $[N - 15, N - 7]$ separadamente em uma maneira similar a Equação (29) usando:

$$d_g(k') = \sum_{i=0}^{N-k'+L_{1/2}-1} x(k' - L_{1/2} + i) z_{L_{1/2}}(i) \quad (33)$$

O último parâmetro a ser determinado na pesquisa de livro de codificação de formato glótico é o ganho g_p que pode ser computado como na Equação (4) com a diferença de que o mesmo não é ligado como na pesquisa do livro de codificação adaptativo. A razão é que o vetor-código de formato glótico filtrado é construído usando formatos glóticos normalizados quantificados com energia bastante diferente a partir da energia dos impulsos de sinal de excitação atuais.

Os índices relacionados à posição do impulso glótico e o formato glótico são transmitidos ao decodificador. A reconstrução do vetor-código de formato glótico filtrado no decodificador é mostrada na figura 11. Deve ser observado que o comprimento de período de passo não precisa mais ser transmitido em uma pesquisa de subquadro do livro de codificação de formato glótico com a exceção quando o subquadro contiver mais do que um impulso glótico como será discutido daqui adiante.

15 Mais Impulsos glóticos em Um Subquadro

Há situações onde o período de passo do sinal de fala é mais curto do que o comprimento do subquadro e neste caso o subquadro pode conter mais do que um impulso glótico (especialmente na configuração TRANSIÇÃO_1_1). Neste caso é necessário se modelar todos os impulsos glóticos. Considerando as limitações do comprimento de período de passo e o comprimento do subquadro, um subquadro não pode conter mais do que dois impulsos glóticos na referida modalidade ilustrativa não-restritiva.

As referidas situações podem ser solucionadas por duas diferentes abordagens. A primeira e mais simples soluciona as referidas situações por meio de um procedimento similar a um aumento de periodicidade (aguçamento do passo) usado em *AMR-WB* (Adaptive Multi-Rate Wideband) como descrito na Referência [1], onde o impulso é basicamente repetido com o período de passo usando um filtro linear. Como ilustrado na figura 12a, o vetor-código de formato glótico $q_k(n)$ é assim processado através de filtro de

repetição adaptativo de forma:

$$Q(z) = \frac{1}{1 - \alpha \cdot z^{-T_0}} \quad (34)$$

O período de passo T_0 pode ser determinado, por exemplo, pela abordagem de pesquisa de passo padrão de alça fechada. O parâmetro α proporciona a energia do segundo impulso e, em uma modalidade ilustrativa não-restritiva, foi ajustado a $\alpha = 0,85$. A referida técnica acrescenta o impulso glótico que falta em uma posição correta no vetor-código de formato glótico. Isto é ilustrado como o impulso pontilhado na figura 12 b. A referida situação aparece quando a soma da posição central de impulso glótico W e o período de passo T_0 é menor do que o comprimento do subquadro N , isto é, $k' + T_0 < N$. Mas também em situações onde a soma da posição de impulso k' e período de passo excede o comprimento do subquadro, o valor de período de passo é também usado para construir o vetor de código fixo quando o aguçamento do passo no livro de codificação algébrico é usado.

O filtro de repetição $Q(z)$ é inserido na parte TM do Codec entre os filtros $G(z)$ e $H(z)$, como mostrado no diagrama de bloco da figura 13 para o codificador. A mesma mudança é realizada no decodificador. De modo similar ao aguçamento do passo, a resposta de impulso do filtro de repetição $Q(z)$ pode ser adicionada à resposta de impulso de $G(z)$ e $H(z)$ antes da pesquisa do livro de codificação de modo que ambos os impulsos são considerados durante a pesquisa embora mantendo a complexidade da pesquisa a um nível baixo.

Outra abordagem para construir o vetor-código de formato glótico com dois impulsos glóticos em um subquadro é de usar uma pesquisa de livro de codificação adaptativo em uma parte do subquadro. As primeiras T_0 amostras do vetor-código de formato glótico $q_k(n)$ são construídas usando a pesquisa do livro de codificação de formato glótico e então as outras amostras no subquadro são construídas usando a pesquisa adaptativa como mostrado na figura 14. A referida abordagem é mais complexa, mas mais precisa.

Para se aumentar adicionalmente a eficiência da codificação, o

procedimento acima descrito pode ser usado mesmo se o segundo impulso glótico aparece em uma das primeiras posições $L_{1/2}$ do próximo subquadro (figura 15). Nesta situação, isto é, quando k' e T_o sustentam $N \leq (k' + T_o) < (N + L_{1/2})$ apenas poucas amostras (menos do que $L_{1/2} + 1$) do formato glótico são usadas no final do subquadro atual. A referida abordagem é usada em uma modalidade ilustrativa não-restritiva. A referida abordagem apresenta uma limitação pelo fato de que o valor de período de passo transmitida nas referidas situações é limitado ao $T_o < N$ (isto é uma questão de codificação efetiva), embora de modo ideal seu valor deva ser limitado a $T_o \leq N + L_{1/2}$. Portanto se o segundo impulso glótico aparece no início do próximo subquadro, o procedimento de repetição não pode ser usado para algumas das primeiras posições $L_{1/2}$ do impulso glótico k' do primeiro impulso glótico.

Implementação da técnica de codificação *TM* em Codec *EV-VBR*

A técnica de codificação *TM* de acordo com a modalidade não-restritiva ilustrativa foi implementada no Codec *EV-VBR*. *EV-VBR* utiliza a frequência de amostragem interna de 12,8 kHz e o comprimento do quadro de 20 ms. Cada quadro é dividido em quatro subquadros de $N = 64$ amostras. O procedimento de classificação *EV-VBR* foi adaptado para selecionar quadros a serem codificados usando a técnica de codificação *TM*. Nesta implementação, a contribuição de ganho do livro de codificação de formato glótico é quantificada em duas etapas como ilustrado na figura 16, onde $G(z)$ é o filtro de forma, W é a posição do centro do formato glótico e g_m é um ganho *TM*, isto é, uma energia basicamente quantificada do vetor-código de formato glótico. O ganho de *TM* g_m é encontrado da mesma forma que o ganho de passo usando a Equação (4) apenas com a diferença de que o mesmo não é ligado. O mesmo é então quantificado por meio de um quantificador de escala de 3-bits e um bit por sinal é usado. O vetor-código de formato glótico é então escalado usando o referido ganho g_m . Após ambas as contribuições ao sinal de excitação filtrado (sinais de contribuição de primeiro e segundo estágio, isto é, a contribuição filtrada do livro de codificação de formato glótico e a contribuição filtrada do livro de codificação algébrico) ser encontradas, o ganho do sinal de excitação de primeiro estágio é adicional-

mente ajustado junto com a quantificação de ganho do sinal de excitação de segundo estágio, usando a quantificação de vetor de ganho *EV-VBR* padrão (*VQ*). Deste modo, os livros de codificação de quantificação de ganho *EV-VBR* projetados para os modos de codificação genérico ou sonoro podem também ser usados em *Codificação TM*. Evidentemente, se encontra inserido no âmbito da presente invenção se realizar a quantificação de ganho usando outros métodos diferentes.

A pesquisa de posição central de impulso glótico k' deve ser teoricamente produzida para todas as posições no subquadro, isto é, dentro da faixa $[0, N - 1]$. No entanto como já mencionado, a referida pesquisa é computacionalmente intensiva considerando o número de formatos glóticos a serem tentados e, na prática, a mesma pode ser realizada apenas no intervalo de diversas amostras a cerca da posição do valor máximo absoluto no sinal LP residual. O intervalo de pesquisa pode ser ajustado em ± 4 amostras em torno da posição do primeiro impulso glótico máximo no sinal LP residual no quadro atual. Deste modo, a complexidade de processamento é aproximadamente a mesma que para a codificação *EV-VBR* genérica usando a pesquisa de livro de codificação fixo e adaptativo.

Os parâmetros transmitidos relacionados à técnica de codificação *TM* são relacionados na Tabela 4 com o número de bits correspondente. O parâmetro T_o , que é usado para determinar o filtro $Q(z)$ ou realizar pesquisa adaptativa para o segundo impulso glótico em caso de dois impulsos em um subquadro, é transmitida quando $T_o \leq N$. Os parâmetros restantes usados para um quadro *TM*, mas comum com o processador genérico ACELP, não são mostrados aqui (bits de identificação de quadros, parâmetros LP, retardo de passo para a excitação adaptativa, excitação de livro de codificação fixo, primeiro e segundo estágios de ganhos do livro de codificação). Quando os parâmetros *TM* são adicionados à corrente de bits, o número de bits originalmente alocados a outros parâmetros *EV-VBR* é reduzido de modo a manter um coeficiente constante de bits. Os referidos bits podem ser reduzidos, por exemplo, a partir dos bits de excitação do livro de codificação fixo assim como a partir da quantificação de ganho.

Tabela 4 - Parâmetros na corrente de bit transmitida para os subquadros codificados usando a TM.

<i>Marcação</i>	<i>Significado</i>	<i>Número de bits</i>
<i>ID</i>	Identificação da configuração	1 - 4
Formato	Formato de impulso glótico	3
k'	Posição do centro de impulso glótico	6
g^{TM}	Ganho <i>TM</i>	3
Sinal (g_m)	Sinal de ganho <i>TM</i>	1
T_o	Período de passo de alça fechada (se aplicável)	5

As tabelas de alocação de bits usadas em *EV-VBR* são mostradas aqui abaixo. Deve ser lembrado que, quando a pesquisa do livro de codificação de formato glótico não é aplicada ao primeiro subquadro, apenas o livro de codificação fixo e seu ganho são transmitidos para codificar o sinal de excitação em subquadros precedendo o subquadro do livro de codificação de formato glótico. A mesma situação ocorre para configurações TRANSIÇÃO_1_3 e TRANSIÇÃO_1_4. Nesses casos é possível se manter o mesmo tamanho ou tamanhos ainda maiores do livro de codificação fixo para todos os subquadros como na Codificação genérica original *CELP*.

Desempenho de Técnica *TM* em Codec *EV-VBR*

Nesta seção alguns exemplos do desempenho de técnica *TM* de codificação na implementação do Codec *EV-VBR* são apresentados. Na figura 17 um exemplo do impacto da técnica de codificação *TM* é mostrado para a condição de canal limpo. A figura 17a mostra o sinal de entrada de fala, a figura 17b mostra o sinal LP residual e a figura 17c mostra o sinal de excitação de primeiro estágio onde a técnica de codificação *TM* é usada nos primeiros três (3) quadros. Como esperado, a diferença entre o sinal residual e o sinal de excitação de primeiro estágio é mais pronunciado no início de cada quadro. Em direção ao final do quadro, o sinal de excitação de primeiro estágio corresponde mais proximamente ao sinal residual pelo fato de que a pesquisa do livro de codificação adaptativo padrão é usada.

Tabelas 5 e 6 resumem alguns exemplos do desempenho da técnica de codificação *TM* medida usando valores SNR.

No primeiro exemplo (Tabela 5) a técnica *TM* foi implementada em Codec com uma frequência de amostragem de núcleo (interno) $F_s = 8$ kHz (isto é, um comprimento de subquadro $N = 40$ amostras), de livro de codificação de formato glótico com dezesseis (16) formatos de comprimento dezesseis (17) amostras foram usadas, e sinais de entrada de banda estreita foram testados. A partir de Tabela 5 pode ser visto que a codificação de quadros sonoros de início usando a técnica de codificação *TM* aumenta a qualidade de sinal de fala de saída (vide valores segmentários e segmentários pesados *SNR* para 1 e 2 quadros *TM*). Ademais o aumento da *SNR* pode ser observado se o quadro de início sonoro e um quadro consecutivo forem codificados usando a técnica de codificação *TM*. Entretanto, se mais de um quadro em seguida do quadro de início sonoro for também codificado usando a técnica de codificação *TM*, os valores *SNR* diminuem. O *SNR* pesado é o *SNR* pesado pela energia do quadro normalizada pelo comprimento do quadro, em dB.

Tabela 5 – comparação da medição *SNR* do impacto da técnica de codificação *TM* em sinais NB.

<i>Número de quadros TM</i>	<i>Pesado segmentário</i>	<i>Segmentário SNR [dB]</i>	<i>SNR [dB]</i>
0 (sem Codificação <i>TM</i>)	10,85	10,20	12,05
1 (<i>TM</i> no quadro de início)	10,88	10,48	11,03
2 (<i>TM</i> em quadro de início + 1 quadro)	10,90	10,49	11,04
3 (<i>TM</i> em quadro de início + 2 quadros)	10,80	10,41	10,92

A Tabela 6 resume um exemplo do desempenho do Codec *EV-VBR* Codec com núcleo de frequência de amostragem (interno) $F_s = 12.8$ kHz, *WB* sinal de fala de entrada e de livro de codificação de formato glótico com oito (8) formatos de comprimento dezesseis (17) amostras. Em sua maioria pelo fato do comprimento mais longo de subquadro N , os valores *SNRs* mostram alguma degradação para canal limpo quando a técnica de codificação *TM* é usada, mesmo se a mesma for usada em apenas um quadro. Isto é ocasionado em sua maioria em virtude do comprimento limitado dos impulsos de formato glótico. Em comparação com o exemplo *NB*, mais valores zero são apresentados no primeiro estágio de sinal de excitação no subqua-

dro. O benefício de usar a técnica de codificação *TM* neste exemplo está na proteção *FE* (apagamento do quadro).

Tabela 6 – Comparação das medições SNR do impacto da Técnica de codificação TM em sinais WB.

Número de quadros <i>TM</i> -codificados	segmentário pesado SNR [dB]	segmentário SNR [dB]	SNR [dB]
0 (sem Codificação <i>TM</i>)	7,52	7,21	8,61
1 (<i>TM</i> em quadro de início)	7,51	7,21	8,59
1 (<i>TM</i> em quadro após quadro de início)	7,49	7,19	8,55
2 (<i>TM</i> em quadro de início + 1 quadro)	7,48	7,17	8,55
2 (<i>TM</i> em 2 quadros após quadro de início)	7,38	7,10	8,35
3 (<i>TM</i> em quadro de início + 2 quadros)	7,36	7,08	8,31

5 Deve também ser observado que mesmo quando a técnica de codificação *TM* é usada em um quadro após o quadro apagado, há ainda pouca diferença diferença entre a fala sintetizada em canal limpo e canal ruidoso. Isto é porque os estados internos do codificador e do decodificador não dependem apenas do sinal de excitação passado, mas também em muitos outros parâmetros (por exemplo, memórias de filtro, *ISF* (frequências espectrais de imitância) memórias quantificadoras, ...). é possível se testar a variante quando a *Codificação TM* otimizada por quantificação de parâmetros LP sem memória é usada e todos os estados internos são reajustados para os quadros *TM*. Deste modo, todas as memórias que o Codec EV-VBR utiliza no modo de codificação genérico padrão foram reajustadas para garantir que os estados internos de decodificador após um apagamento do quadro são os mesmos que seus estados em condições livres de erro. No entanto a qualidade da fala em condições livres de erro cai significativamente para esta variante. Consequentemente, há um compromisso a ser feito entre o

10

15

20

alto desempenho em condições livres de erro e a resistência a quadros ou pacotes apagados quando não são feitos reajustes de memória adicional.

A Tabela 7 resume o item de complexidade computacional da técnica de codificação *TM*. No pior dos casos a técnica de codificação *TM* aumenta a complexidade no codificador por 1.8 WMOPS (Milhões de Opera-

ções Pesadas Por Segundo). A complexidade no decodificador permanece aproximadamente a mesma.

Tabela 7 - complexidade da técnica de codificação TM (pior caso e valores médios).

Configuração	Codificador WMOPS		Decodificador WMOPS	
	Max	Média	Max	Média
Original (sem Codificação TM)	36,531	34,699	7,053	5,278
Técnica TM de codificação usado	38,346	34,743	7,055	5,281

5 As figuras a seguir ilustram o desempenho da técnica de codificação TM para a modelagem do quadro de início sonoro (figuras 18a - 18c) e para a mitigação da propagação de erro de quadro (figuras 19a -19c). A técnica de codificação TM é usada apenas em um quadro de cada vez neste exemplo. Um segmento do sinal de entrada de fala (figuras 18a e 19a), o

10 sinal de fala sintetizado de saída correspondente processado pelo decodificador EV-VBR sem a técnica de codificação TM como ilustrado nas figuras 18b e 19b, e o sinal de fala de saída sintetizado processado usando o decodificador EV-VBR padrão com técnica de codificação TM (figuras 18c e 19c) são mostrados. Os benefícios da técnica de codificação TM podem ser ob-

15 servados não só na modelagem do quadro de início sonoro (segundo quadro da figura 18) mas também na limitação da propagação de erro de quadro (quarto e quinto quadros da figura 19).

A técnica de encobrimento do apagamento do quadro usada no decodificador EV-VBR é baseada no uso de um retardo extra de decodificador de 20 ms de comprimento (correspondendo a um comprimento do quadro).

20 Quer dizer que se um quadro está faltando, o mesmo é encoberto com o conhecimento parâmetros de quadros futuros. Suponhamos que três (3) quadros consecutivos que são estipulados como $m-1$, m e $m+1$ e adicionalmente suponhamos a situação quando o quadro m está faltando. Então uma

25 interpolação do último quadro corretamente recebido $m-1$ e o quadro seguinte corretamente recebido $m+1$ pode ser computada em vista de determinar os parâmetros de Codec, incluindo em particular mas não exclusivamente os coeficientes de filtro LP (representados por ISFs - Frequências espectrais de

imitância), período de passo de alça fechada T_o , passo e ganhos de livro de codificação fixo. A interpolação ajuda a estimar os parâmetros de quadro perdido mais precisamente para os segmentos sonoros estáveis. Entretanto, com frequência a mesma falha para segmentos de transição quando os parâmetros de Codec variam rapidamente. Para lidar com este problema, o valor absoluto do período de passo pode ser transmitido em cada *quadro TM* mesmo no caso em que o mesmo *nba0* é usado para a construção do primeiro estágio de excitação no quadro atual $m+1$. Isto é válido especialmente para as configurações TRANSIÇÃO_1_4 e TRANSIÇÃO_4.

Outros parâmetros transmitidos no *Quadro TM* são os *ISFs* do quadro precedente. Em codificadores do tipo CELP, os parâmetros *ISF* são geralmente interpolados entre os quadros *ISFs* anteriores e o quadro *ISFs* atual para cada subquadro. Isto garante uma evolução suave do filtro de síntese LP a partir de um subquadro para outro. Em caso de um apagamento do quadro, os *ISFs* do quadro precedente ao apagamento do quadro são usualmente usados para uma interpolação no quadro em seguida do apagamento, em vez dos *ISFs* de quadro apagado. Entretanto durante segmentos de transição, os *ISFs* variam rapidamente e os últimos *ISFs* de bom quadro devem ser bastante diferentes a partir dos *ISFs* do quadro apagado que está faltando. Substituindo os *ISFs* do quadro que falta pelos *ISFs* do quadro anterior pode-se assim ocasionar importantes artefatos. Se os *ISFs* de quadro passado podem ser transmitidos, os mesmos podem ser usados para interpolação de *ISF* no *quadro TM* no caso em que o quadro anterior for apagado. Posteriormente, diferentes estimativas de coeficientes LP usadas para a interpolação *ISF* quando o quadro precedente um *Quadro TM* está faltando serão descritas.

A implementação final da técnica de codificação *TM* ao Codec EV-VBR supõe que apenas um quadro após início/quadro de transição é codificado usando *TM*. Deste modo, cerca de 6,3% de quadros de fala ativa são selecionados para codificação e decodificação *TM*.

Outra categoria de testes se focou no aumento da eficiência da codificação. A classificação foi realizada na pesquisa de alça fechada quan-

do duas variantes - com e sem a técnica de codificação *TM* – foram computadas lado a lado no codificador e uma variante com um *SNR* mais elevado foi escolhido como um sinal de saída.

Os resultados para o Codec EV-VBR com coeficiente de bit de 8 kbps são resumidos na Tabela 8. No caso *WB*, 28% de quadros de fala ativa foi classificado para codificar usando a técnica de codificação *TM* e um aumento de 0,203 dB em *SNR* segmentário foi alcançado. No caso *NB*, 25% de quadros de fala ativa foi classificado para codificar usando a técnica de codificação *TM* e um aumento de mesmo 0,300 dB em *SNR* segmentário foi alcançado. Infelizmente este teste objetivo não foi confirmado pelos testes de listagem subjetivos que reportaram nenhuma preferência entre Codec com e sem a técnica de codificação *TM*. Embora não haja degradação da qualidade da fala e o número total de quadros *TM* ser quatro (4) vezes mais alto em comparação com a classificação de alça aberta que resulta em proteção FE muito maior, a referida classificação e as classificações de resultados similares não são usadas em uma implementação de Codec EV-VBR em virtude da maior complexidade.

Tabela 8 - SNR Segmentário e comparação de medição de SNR entre Codec com e sem técnica de codificação TM implementada quando classificação de alça fechada é usada.

<i>Número de quadros TM-codificados</i>	<i>SNR segmentário [dB]</i>	<i>SNR [dB]</i>
Codec sem <i>TM</i> , sinal <i>WB</i>	7,34	8,89
Codec com <i>TM</i> , sinal <i>WB</i>	7,54	9,04
Codec sem <i>TM</i> , sinal <i>NB</i>	7,58	10,62
Codec com <i>TM</i> , sinal <i>NB</i>	7,88	10,97

Tabelas de alocação de Bits para técnica de codificação *TM* em Codec EV-VBR

A técnica de codificação *TM* foi implementada em um Codec EV-VBR candidato para padronização *ITU-T*. A Tabela 9 a seguir mostra as tabelas de alocação de bits do modo genérico original e todas as configurações de modo de *Codificação TM* que foram introduzidas aqui acima. As referidas configurações são usadas no Codec EV-VBR.

Tabela 9 - Tabelas de Alocação de bits para modo de codificação genérico e para todas as configurações TM como usadas no Codec EV-VBR (ID representa identificação da configuração, ISFs para Frequências espectrais de imitação e FCB para Livro de codificação fixo, subfr. é subquadro).

a) Genérico
Parâmetros de bits
2 tipo codificador
1 NB/WB
36 ISFs
3 estimativa de energia
8 passos de primeiro subquadro
5 ganhos de primeiro subquadro
5 passos do segundo subquadro
5 ganhos do segundo subquadro
8 passos do terceiro subquadro
5 ganhos do terceiro subquadro
5 passos do quarto subquadro
5 ganhos do quarto subquadro
12 FCB do primeiro subquadro
20 FCB do segundo subquadro
20 FCB do terceiro subquadro
20 FCB do quarto subquadro
160 bits no total

b) TRANSIÇÃO 1_1
Parâmetros de bits
2 tipo codificador
1 NB/WB
36 ISFs
3 estimativa de energia
1 ID <i>TM</i> do subquadro
5 passo do primeiro subquadro
3 formato de <i>TM</i>
6 posição de <i>TM</i>
1 sinal de ganho de <i>TM</i>
3 valor de ganho de <i>TM</i>
5 ganhos de primeiro subquadro
5 passo de segundo subquadro
5 ganhos do segundo subquadro
5 passo do terceiro subquadro
5 ganhos do terceiro subquadro
5 passo do quarto subquadro
5 ganhos do quarto subquadro
20 FCB do primeiro subquadro
20 FCB do segundo subquadro
12 FCB do terceiro subquadro
12 FCB do quarto subquadro
160 bits no total

c) TRANSIÇÃO_1_2
Parâmetros de bits
2 tipo codificador
1 NB/WB
36 ISFs
3 estimativa de energia
1 ID <i>TM</i> do subquadro
1 ID <i>TM</i> do subquadro
3 formato de <i>TM</i>
6 posição de <i>TM</i>
1 sinal de ganho de <i>TM</i>
3 valor de ganho de <i>TM</i>
5 ganhos de primeiro subquadro
1 ID2 de <i>TM</i> do subquadro
1 ID2 de <i>TM</i> do subquadro
7 passo do segundo subquadro
5 ganhos do segundo subquadro
5 passo do terceiro subquadro
5 ganhos do terceiro subquadro
5 passo do quarto subquadro
5 ganhos do quarto subquadro
20 FCB do primeiro subquadro
20 FCB do segundo subquadro
12 FCB do terceiro subquadro
12 FCB do quarto subquadro
160 bits no total

d) TRANSIÇÃO 1_3
Parâmetros de bits
2 tipo codificador
1 NB/WB
36 ISFs
3 estimativa de energia
1 ID <i>TM</i> do subquadro
1 ID <i>TM</i> do subquadro
3 formato de <i>TM</i>
6 posição de <i>TM</i>
1 sinal de ganho de <i>TM</i>
3 valor de ganho de <i>TM</i>
5 ganhos de primeiro subquadro
1 ID2 de subquadro de <i>TM</i>
1 ID2 de subquadro de <i>TM</i>
3 ganho de segundo subquadro
7 passo de terceiro subquadro
5 ganhos de terceiro subquadro
4 passo de quarto subquadro
5 ganhos de quarto subquadro
20 FCB de primeiro subquadro
12 FCB de segundo subquadro
20 FCB de terceiro subquadro
20 FCB de quarto subquadro
160 bits no total

e) TRANSIÇÃO_1_4
Parâmetros de bits
2 tipo codificador
1 NB/WB
36 ISFs
3 estimativa de energia
1 ID <i>TM</i> do subquadro
1 ID <i>TM</i> do subquadro
3 formato de <i>TM</i>
6 posição de <i>TM</i>
1 sinal de ganho de <i>TM</i>
3 valor de ganho de <i>TM</i>
5 ganhos de primeiro subquadro
1 ID2 de subquadro de <i>TM</i>
3 ganho de segundo subquadro
2 ganho de terceiro subquadro
7 passo de quarto subquadro
5 ganhos de quarto subquadro
20 FCB de primeiro subquadro
20 FCB de segundo subquadro
20 FCB de terceiro subquadro
20 FCB de quarto subquadro
160 bits no total

f) TRANSIÇÃO_2
Parâmetros de bits
2 tipo codificador
1 NB/WB
36 ISFs
3 estimativa de energia
1 ID <i>TM</i> do subquadro
1 ID <i>TM</i> do subquadro
1 ID <i>TM</i> do subquadro
2 ganho de primeiro subquadro
3 formato de <i>TM</i>
6 posição de <i>TM</i>
1 sinal de ganho de <i>TM</i>
3 valor de ganho de <i>TM</i>
5 ganhos de segundo subquadro
8 passo de terceiro subquadro
5 ganhos de terceiro subquadro
5 passo de quarto subquadro
5 ganhos de quarto subquadro
20 FCB de primeiro subquadro
20 FCB de segundo subquadro
12 FCB de terceiro subquadro
20 FCB de quarto subquadro
160 bits no total

g) TRANSIÇÃO_3
Parâmetros de bits
2 tipo codificador
1 NB/WB
36 ISFs
3 estimativa de energia
1 ID <i>TM</i> do subquadro
1 ID <i>TM</i> do subquadro
1 ID <i>TM</i> do subquadro
1 ID <i>TM</i> do subquadro
3 ganho de primeiro subquadro
3 ganho de segundo subquadro
5 passo de terceiro subquadro
3 formato de <i>TM</i>
6 posição de <i>TM</i>
1 sinal de ganho de <i>TM</i>
3 valor de ganho de <i>TM</i>
5 ganhos de terceiro subquadro
8 passo de quarto subquadro
5 ganhos de quarto subquadro
12 FCB de primeiro subquadro
20 FCB de segundo subquadro
20 FCB de terceiro subquadro
20 FCB de quarto subquadro
160 bits no total

h) TRANSIÇÃO_4
Parâmetros de bits
2 tipo codificador
1 NB/WB
36 ISFs
3 estimativa de energia
1 ID <i>TM</i> do subquadro
1 ID <i>TM</i> do subquadro
1 ID <i>TM</i> do subquadro
1 ID <i>TM</i> do subquadro
3 ganho de primeiro subquadro
2 ganho de segundo subquadro
3 ganho de terceiro subquadro
8 passo de quarto subquadro
3 formato de <i>TM</i>
6 posição de <i>TM</i>
1 sinal de ganho de <i>TM</i>
3 valor de ganho de <i>TM</i>
5 ganhos de quarto subquadro
20 FCB de primeiro subquadro
20 FCB de segundo subquadro
20 FCB de terceiro subquadro
20 FCB de quarto subquadro
160 bits no total

- Há uma excessão para a configuração TRANSIÇÃO_2 na Tabela 9. A referida Tabela de alocação de bits pode ser usada apenas na situação quando se decide usar a técnica de codificação *TM* nos quadros em seguida do quadro de início sonoro apenas (o quadro de início sonoro é codificado usando o modo de codificação genérico e apenas um quadro em seguida do quadro de início sonoro é codificado usando a técnica de codificação *TM*). Nesta situação, o período de passo T_o é $T_o \geq N$ no segundo sub-
- 5

quadro e não há necessidade de transmitir o referido parâmetro no segundo subquadro. Mas se a técnica de codificação *TM* é usada também no quadro de início sonoro, a situação a seguir pode ocorrer. O período de passo é mais curto do que N , mas o início sonoro pode iniciar apenas no segundo subquadro (por exemplo, o primeiro subquadro ainda contendo sinal mudo). Neste caso o período de passo T_o deve ser transmitido. Nesta situação uma Tabela de alocação de bits diferente é usada, parâmetro T_o é transmitido no segundo subquadro usando cinco (5) bits e em um subquadro um livro de codificação fixo mais curto é usado (vide Tabela 10). A mesma situação aparece também para a configuração TRANSIÇÃO_3. Entretanto, o período de passo é transmitido aqui de qualquer modo na implementação presente, não-limitativa (se o quadro de início for codificado usando a técnica de codificação *TM* ou não) porque não há bom uso dos bits salvos para outra codificação de parâmetro.

Outras alocações de bits podem ser usadas em diferentes configurações de modo de transição. Por exemplo, mais bits podem ser alocados ao livro de codificação fixo nos subquadros contendo pulsos glóticos. Por exemplo, no modo de TRANSIÇÃO_3, a *FCB* com doze (12) bits pode ser usada no segundo subquadro e vinte e oito (28) bits no terceiro subquadro. Evidentemente, outras *FCBs* diferentes de 12- e 20-bits podem ser usadas em diferentes implementações de codificação.

Tabela 10- Tabela de Alocação de bits para configuração de TRANSIÇÃO_2 se TM é usado também no quadro de início.

TRANSIÇÃO_2a
Parâmetros de bits
2 tipo codificador
1 NB/WB
36 ISFs
3 estimativa de energia
1 ID <i>TM</i> do subquadro
1 ID <i>TM</i> do subquadro
1 ID <i>TM</i> do subquadro

3 ganho de primeiro subquadro
5 passo de segundo subquadro
3 formato de subquadro
6 posição de <i>TM</i>
1 sinal de ganho de <i>TM</i>
3 valor de ganho de <i>TM</i>
5 ganhos de segundo subquadro
8 passo de terceiro subquadro
5 ganhos de terceiro subquadro
5 passo de quarto subquadro
5 ganhos de quarto subquadro
20 FCB de primeiro subquadro
20 FCB de segundo subquadro
12 FCB de terceiro subquadro
12 FCB de quarto subquadro
158 bits no total

Se houver largura de banda disponível, um aumento adicional pode ser alcançado ao se transmitir mais informação para melhor proteção de apagamento do quadro (*FE*). O Codec *VMR-WB* é um exemplo de um Codec que utiliza alguma porção de bits de proteção *FE*. Por exemplo, quatorze (14) bits de proteção por quadro são usados na Codificação do tipo de coeficiente amplo genérico em *VMR-WB* em Rate-Set II. Os referidos bits representam a classificação de quadro (2 bits), energia de fala sintetizada (6 bits) e posição de pulso glótico (6 bits). O pulso glótico é inserido artificialmente no decodificador quando um quadro de princípio sonante é perdido.

Os referidos bits de proteção *FER* não são muito importantes para a construção da excitação no *Quadro TM* pelo fato de que a técnica de codificação *TM* não faz uso do sinal de excitação passada; a técnica de codificação *TM* constroi o sinal de excitação usando parâmetros transmitidos no quadro atual (*TM*). Os referidos bits podem ser, entretanto, empregados para a transmissão de outros parâmetros. Em um exemplo de implementação, os referidos bits podem ser usados para transmitir no *Quadro TM* atual os parâme-

tros *ISF* do quadro anterior; entretanto doze (12) bits em vez de trinta e seis (36) bits estão disponíveis). Os referidos *ISFs* são usados para uma reconstrução mais precisa de coeficientes de filtro *LP* em caso de apagamento do quadro.

5 No Codec EV-VBR o ajuste de parâmetros *LP* é computado centrado no quarto subquadro, enquanto o primeiro, segundo, e terceiro subquadros usam uma interpolação linear dos parâmetros do filtro *LP* entre o quadro atual e o quadro anterior. A interpolação é realizada nos *ISPs* (pares de Imitância Espectral). Considerando que $\mathbf{q}_4^{(m)}$ seja o vetor *ISP* vetor no
10 quarto subquadro do quadro, e $\mathbf{q}_4^{(m-1)}$ o vetor *ISP* no quarto subquadro do quadro passado $m-1$. Os vetores *ISP* interpolados nos primeiro, segundo, e terceiro subquadros são considerados pelas equações:

$$\begin{aligned}\mathbf{q}_1^{(m)} &= 0.55\mathbf{q}_4^{(m-1)} + 0.45\mathbf{q}_4^{(m)}, \\ \mathbf{q}_2^{(m)} &= 0.2\mathbf{q}_4^{(m-1)} + 0.8\mathbf{q}_4^{(m)},\end{aligned}\quad (35)$$

A referida interpolação entretanto, não é diretamente adequada para a técnica de codificação *TM* em caso de apagamento do quadro anterior. Quando o quadro precedente o *Quadro TM* está faltando, se pode supor
15 que o último quadro corretamente recebido é mudo. É mais eficiente nesta situação para reconstruir vetor *ISF* para o quadro que falta com diferentes constantes de interpolação e não importa se há alguma informação *ISFs* a partir de bits de proteção FER disponíveis ou não. Em general, a interpolação está usando os *ISPs* do quadro anterior mais pesadamente. Os vetores
20 *ISP* para o quadro que falta m pode ser considerando no decodificador, por exemplo, ao se usar as equações a seguir:

$$\begin{aligned}\mathbf{q}_1^{(m)} &= \mathbf{q}_4^{(m-1)}, \\ \mathbf{q}_2^{(m)} &= \mathbf{q}_4^{(m-1)}, \\ \mathbf{q}_3^{(m)} &= 0.7\mathbf{q}_4^{(m-1)} + 0.3\mathbf{q}_4^{(m)}, \\ \mathbf{q}_4^{(m)} &= 0.1\mathbf{q}_4^{(m-1)} + 0.9\mathbf{q}_4^{(m)}.\end{aligned}\quad (36)$$

O quadro *TM* $m+1$ seguinte corretamente recebido então utiliza

coeficientes de interpolação LP descritos pelas equações (35). Também os coeficientes de interpolação nas Equações (36) são considerados como um exemplo não-limitante. Os coeficientes finais podem ser diferentes e adicionalmente é desejável se usar um conjunto de coeficientes de interpolação quando alguma informação ISF a partir do quadro anterior está disponível e outro conjunto quando a informação ISF a partir do quadro anterior não está disponível (isto é, não há bits de proteção de apagamento do quadro na corrente de bits).

Período de passo e codificação de ganho em quadros TM em Codec EV-VBR

O valor do período de passo T_o é transmitido para cada subquadro no modo de codificação genérico usado no Codec EV-VBR. No primeiro e terceiro subquadros, uma codificação de 8-bits é usada embora o valor de período de passo seja transferido com resolução fracional (14 para T_o na faixa $[T_{min}, 91 \text{ Vz}]$) ou inteiro (para T_o na faixa $[92, T_{max}]$). No segundo e quarto subquadros, uma pesquisa delta é usada e o valor de período de passo sempre com resolução fracional é codificado com cinco (5) bits. Pesquisa Delta significa uma pesquisa dentro da faixa $[T_{Op} - 8, T_{Op} + 7_{1/2}]$, onde T_{Op} é o inteiro mais próximo ao período de passo fracional do subquadro anterior (primeiro ou terceiro). Os valores do período de passo são limitados no Codec EV-VBR para valores dentro da faixa $[T_{min}, T_{max}]$, onde $T_{min} = 34$ e $T_{max} = 231$.

O ganho de passo g_p e o o ganho do livro de codificação fixo g_c são codificados no Codec EV-VBR em princípio da mesma maneira que em $AMR-WB+$ Codec [5]. Primeiro uma estimativa de uma energia de livro de codificação fixo não-previsiva escalada é calculada para todos os subquadros em um quadro e quantificado com três (3) bits uma vez por quadro (vide a estimativa de parâmetros de energia na Tabela 9). Então o ganho de passo g_p e o ganho do livro de codificação fixo g_c são vetores quantificados e codificados em uma etapa usando cinco (5) bits para cada subquadro.

A energia estimada do livro de codificação fixo é computada e quantificada como a seguir. Primeiro, a energia residual LP é computada em cada subquadro k usando a equação a seguir:

$$E_{res}(k) = 10 \log_{10} \left(\frac{1}{N} \sum_{n=0}^{N-1} u^2(n) \right), \quad (37)$$

onde $u(n)$ é o sinal LP residual. Então a energia residual média por subquadro é encontrada através da equação a seguir:

$$\bar{E}_{res} = \frac{1}{4} \sum_{k=0}^3 E_{res}(k) \quad (38)$$

A energia do livro de codificação fixo é estimada a partir da energia residual ao remover uma estimativa da contribuição do livro de codificação adaptativo. Isto é realizado ao remover uma energia relacionada a uma correlação média normalizada obtida a partir das duas análises de passo de alça abertas realizada no quadro. A equação a seguir é usada:

$$E_s = \bar{E}_{res} - 10\bar{R} \quad (39)$$

onde R é a média das correlações de passo normalizada obtida a partir da análise de passo de alça aberta para cada meio-quadro do quadro atual. A energia do livro de codificação fixo estimada escalada não é dependente da energia do quadro anterior e assim o princípio de codificação de ganho é robusto ao apagamento de quadros.

Uma vez que a estimativa da energia do livro de codificação fixo é encontrada, o ganho de passo e o a correção do ganho do livro de codificação fixo são computados: a energia do livro de codificação fixo escalada estimada é usada para calcular o ganho do livro de codificação fixo estimado e o fator de correção γ (proporção entre os ganhos verdadeiros e os estimados do livro de codificação fixo). O valor γ é o vetor quantificado junto com o ganho de passo usando cinco (5) bits por subquadro. Para a configuração do quantificador, um método de meio k modificado [4] é usado. O ganho de passo é restrito dentro do intervalo $\langle 0; 1,2 \rangle$ durante a inicialização do livro de codificação e $\langle 0; \infty \rangle$ durante o aprimoramento iterativo do livro de codificação. Da mesma forma, o fator de correção γ é limitado por $\langle 0; 5 \rangle$ durante a inicialização e $\langle 0; \infty \rangle$ durante o aprimoramento do livro de codificação. O

algoritmo de meio k modificado procura minimizar o critério a seguir:

$$E = g_p^2 \mathbf{y}_1^T \mathbf{y}_1 - 2g_p \mathbf{x}_1^T \mathbf{y}_1 + g_c^2 \mathbf{y}_2^T \mathbf{y}_2 - 2g_c \mathbf{x}_1^T \mathbf{y}_2 + 2g_p g_c \mathbf{y}_1^T \mathbf{y}_2 \quad (40)$$

Quando se usa a técnica de codificação *TM*, a transmissão do período de passo e ambos passo e ganhos de livro de codificação fixo podem não ser necessários para os subquadros onde não há importante impulso glótico, e apenas a contribuição do livro de codificação fixo pode ser computada.

O a seguir é uma lista e a descrição de todas as configurações *TM*:

Configuração de TRANSIÇÃO 1 1 (figura 20) - Nesta configuração um ou dois primeiros impulsos glóticos aparecem no primeiro subquadro que é processado usando a pesquisa do livro de codificação de formato glótico. Isto significa que o valor de período de passo no primeiro subquadro pode ser dotado de um valor máximo menor do que o comprimento do subquadro, isto é, $T_{min} < T_o < N$. Com a resolução inteira pode ser codificado com cinco (5) bits. O período de passos nos próximos subquadros é encontrado usando a pesquisa delta de 5-bits com uma resolução fracional.

Esta é a configuração mais exigente de bit da técnica de codificação *TM*, isto é, quando o livro de codificação de formato glótico é usado no primeiro subquadro e o período de passo T_o é transmitido para determinação de filtro $Q(z)$, ou para a pesquisa do livro de codificação adaptativo na parte do primeiro subquadro. Esta configuração utiliza no primeiro subquadro o procedimento como descrito acima. A referida configuração é usada no Codec EV-VBR também quando apenas um impulso glótico aparece no primeiro subquadro. Aqui o período de passo T_o sustenta $T_o < N$ e é usado para aumento de periodicidade [1] na pesquisa do livro de codificação fixo.

Configuração de TRANSIÇÃO 1 2 (figura 21) - Quando a configuração *TRANSIÇÃO_1_2* é usada, o primeiro subquadro é processado usando a pesquisa do livro de codificação de formato glótico. O período de passo não é necessário e todos os subquadros seguintes são processados usando a pesquisa do livro de codificação adaptativo. Pelo fato de que o se-

gundo subquadro é conhecido por conter o segundo impulso glótico, o valor máximo de período de passo sustenta $T_o < 2/V - 1$. O referido valor máximo pode ser adicionalmente reduzido graças ao conhecimento da posição do impulso glótico k' . O valor de período de passo no segundo subquadro é então codificado usando sete (7) bits com uma resolução fracional em toda a faixa. No terceiro e quarto subquadros, pesquisa delta usando cinco (5) bits é usada com a resolução fracional.

Configuração de TRANSIÇÃO_1_3 (figura 22) - Quando a configuração de TRANSIÇÃO_1_3 é usada o primeiro subquadro é processado usando a pesquisa do livro de codificação de formato glótico mais uma vez sem uso do período de passo. Em virtude do segundo subquadro do sinal LP residual não conter impulso glótico e a pesquisa adaptativa ser inútil, o sinal de excitação de primeiro estágio é substituído por zeros no segundo subquadro. Os parâmetros do livro de codificação adaptativo (T_0 e g_p) não são transmitidos no segundo subquadro e os bits salvos são usados para o aumento do tamanho FCB no terceiro subquadro. Em virtude do segundo subquadro conter um mínimo de informação útil, apenas FCB de 12-bits é usado e o FCB de 20-bits é usado no quarto subquadro. O sinal de excitação de primeiro estágio no terceiro subquadro é construído usando a pesquisa do livro de codificação adaptativo com o valor máximo de período de passo ($3.N - 1 - k'$) e valor mínimo ($2.N - k'$); assim apenas uma codificação de 7-bits do período de passo com resolução fracional sobre toda a faixa é usada. O quarto subquadro é processado usando a pesquisa adaptativa mais uma vez com codificação de pesquisa delta de 5-bits do valor de período de passo.

No segundo subquadro apenas o ganho do livro de codificação fixo g_c é transmitido. Conseqüentemente, apenas dois (2) ou três (3) bits são necessários para a quantificação de ganho em vez do quantificador de 5-bits usado no subquadro com codificação tradicional *ACELP* (isto é, quando ganhos g_p e g_c são transmitidos). Isto é válido também para todas as configurações a seguir. A decisão de se o quantificador de ganho deve usar dois (2) ou três (3) bits é tomada de modo a se adaptar ao número de bits disponível no quadro.

Configuração de TRANSIÇÃO_1_4 (figura 23) - Quando a configuração TRANSIÇÃO_1_4 é usada, o primeiro subquadro é processado usando a pesquisa do livro de codificação de formato glótico. Mais uma vez, o período de passo não precisa ser transmitido. Mas em virtude do sinal LP residual não conter impulso glótico no segundo e também no terceiro subquadro, a pesquisa do livro de codificação adaptativo é inútil para os referidos dois subquadros. Mais uma vez, o sinal de excitação de primeiro estágio nos referidos subquadros é substituído por zeros e os bits salvos são usados para o aumento do tamanho FCB de modo que todos os subquadros podem se beneficiar do uso de $FCBs$ de 20-bits. O valor de período de passo é transmitido apenas no quarto subquadro e o seu valor mínimo é $(3.N-k')$. O valor máximo do período de passo é limitado por T_{max} . Não importa se o segundo impulso glótico aparece no quarto subquadro ou não (o segundo impulso glótico pode estar presente no próximo quadro se $k'+T_{max} \geq N$).

O valor absoluto do período de passo é usado no decodificador para o encobrimento do quadro; portanto o referido valor absoluto do período de passo é transmitido na situação quando o segundo impulso glótico aparece no próximo quadro. Quando um quadro m precedente do Quadro TM $m+1$ está faltando, o conhecimento correto do valor de período de passo a partir dos quadros $m-1$ e $m+1$ ajuda a reconstruir as partes que faltam do sinal de síntese no quadro m com sucesso.

Configuração de TRANSIÇÃO_2 (figura 24) - Quando o primeiro impulso glótico aparece no segundo subquadro e apenas quadros após quadros sonoros de início são codificados usando a técnica de codificação TM (isto é, os quadros sonoros de início são codificados com a codificação genérica legal), o período de passo é transmitido apenas no terceiro e quarto subquadros. Neste caso, apenas os parâmetros do livro de codificação fixo são transmitidos no primeiro subquadro.

O quadro mostrado na figura 24 supõe uma configuração quando TM não é usado em quadros sonoros de início. Se TM for usada também nos quadros sonoros de início, a configuração de TRANSIÇÃO_2a é usada onde o período de passo T_0 é transmitido no segundo subquadro para usar

o procedimento como descrito acima.

Configuração de TRANSIÇÃO 3 (figura 25) - Quando o primeiro impulso glótico aparece no terceiro subquadro e apenas quadros após os quadros sonoros de início são codificados usando a técnica de codificação *TM* (isto é, os quadros sonoros de início são codificados com a codificação genérica), o período de passo é transmitido apenas no quarto subquadro. Neste caso apenas os parâmetros do livro de codificação fixo são transmitidos nos primeiro e segundo subquadros.

O período de passo é ainda transmitido para o terceiro subquadro na corrente de bits. Entretanto o mesmo não é útil se a técnica de codificação *TM* não for usada para codificar os quadros sonoros de início. Este valor é útil apenas quando quadros sonoros de início são codificados usando a técnica de codificação *TM*.

Configuração de TRANSIÇÃO 4 (figura 26) - Quando o primeiro impulso glótico aparece no quarto subquadro e apenas quadros após quadros sonoros de início são codificados usando a técnica de codificação *TM* (isto é, os quadros sonoros de início são codificados com o legado da codificação genérica), o valor da informação de período de passo não é usado no referido subquadro. Entretanto o valor de período de passo é usado no encobrimento de quadro no decodificador (este valor é usado para a reconstrução do quadro que falta quando o quadro precedente ao *Quadro TM* está faltando). Assim o valor de passo é transmitido apenas no quarto subquadro e apenas os parâmetros do livro de codificação fixo são transmitidos no primeiro, segundo e terceiro subquadros (o passo de ganho g_p não é necessário). Os bits salvos permitem que o *FCB* de 20 bits seja usado em cada subquadro.

Embora a presente invenção tenha sido descrita na descrição anterior em conexão com a modalidade ilustrativa não-restritiva da mesma, a referida modalidade ilustrativa não-restritiva pode ser modificada à vontade, dentro do âmbito das reivindicações anexas, sem se desviar do âmbito e espírito da presente invenção.

Referências

- [1] B. BESSETTE, R. SALAMI, R. LEFEBVRE, M. JELINEK, J. ROTOLA-PUKKILA, J. VAINIO, H. MIKKOLA, e K. JARVINEN, "The Adaptive Multi-Rate Wideband Speech Codec (AMR-WB)", Special Issue of IEEE Transactions on Speech and Audio Processing, Vol. 10, No. 8, pp. 620 - 636, Novembro de 2002.
- [2] R. SALAMI, C. LAFLAMME, J-P. ADOUL, e D. MASSALOUX, "A toll quality 8 kb/s speech Codec for the personal communications system (PCS)", IEEE Trans, on Vehicular Technology, Vol. 43, No. 3, pp. 808 - 816, Agosto de 1994.
- [3] 3GPP2 Tech. Spec. C.S0052-A v1.0, "Source-Controlled Variable-Rate Multimode Wideband Speech Codec (VMR-WB), Service Options 62 e 63 for Spread Spectrum Systems," Apr. 2005; <http://www.3gpp2.org>
- [4] S. P. Lloyd, "Least squares quantization in PCM," IEEE Transactions on Information Theory, Vol. 28, No.2, pp. 129 - 136, Março de 1982.
- [5] 3GPP Tech. Spec. 26.290, "Adaptive Multi-Rate - Wideband (AMR-WB+) Codec; Transcoding functions," June 2005.
- [6] "Extended high-level description of the Q9 EV-VBR baseline Codec", ITU-T SG16 Tech. Cont. COM16-C199R1-E, Junho de 2007.

REIVINDICAÇÕES

1. Dispositivo de modo de transição para uso em um Codec de sinal de som do tipo preditivo para produzir uma excitação de modo de transição substituindo uma excitação de livro de codificação adaptativo em um quadro de transição e/ou um quadro seguindo a transição no sinal de som, compreendendo:

uma entrada para receber um índice de livro de codificação; e um livro de codificação de modo de transição para gerar um conjunto de vetores-código independente da excitação passada, o livro de codificação de modo de transição respondendo ao índice para gerar, no quadro de transição e/ou quadro seguindo a transição, um dos vetores-código do conjunto que corresponde à referida excitação de modo de transição.

2. Dispositivo de modo de transição, de acordo com o definido na reivindicação 1, em que o livro de codificação de modo de transição compreende um livro de codificação fixo independente da excitação passada.

3. Dispositivo de modo de transição, de acordo com o definido na reivindicação 1, em que o livro de codificação de modo de transição compreende um livro de codificação de formatos de impulsos glóticos.

4. Dispositivo de modo de transição, de acordo com o definido na reivindicação 1, em que o sinal de som compreende um sinal de fala e em que o quadro de transição é selecionado a partir do grupo que consiste em um quadro compreendendo um princípio sonante e um quadro compreendendo uma transição entre dois diferentes sons sonantes.

5. Dispositivo de modo de transição, de acordo com o definido na reivindicação 1, em que o quadro de transição e/ou o quadro seguindo a transição compreende um quadro de transição seguido de diversos quadros.

6. Dispositivo de modo de transição, de acordo com o definido na reivindicação 1, em que o quadro de transição e/ou o quadro seguindo a transição compreende pelo menos um quadro seguindo a transição.

7. Dispositivo de modo de transição, de acordo com o definido na reivindicação 1, em que o quadro de transição e/ou o quadro seguindo a transição cada um dos quais compreende uma pluralidade de subquadros, e

em que o livro de codificação de modo de transição é usado em uma primeira parte dos subquadros e um livro de codificação do tipo preditivo do Codec do tipo preditivo é usado em uma segunda parte dos subquadros.

5 8. Dispositivo de modo de transição, de acordo com o definido na reivindicação 1, em que o livro de codificação compreende um livro de codificação de formato glótico compreendendo vetores-código formados de um formato de impulso glótico dispostos em uma posição específica no vetor-código.

10 9. Dispositivo de modo de transição, de acordo com o definido na reivindicação 8, em que o livro de codificação de formato glótico inclui um número predeterminado de diferentes formatos de impulsos glóticos, e em que cada formato de impulso glótico é posicionado em uma pluralidade de diferentes posições no vetor-código para formar uma pluralidade de diferentes vetores-código do livro de codificação de formato glótico.

15 10. Dispositivo de modo de transição, de acordo com o definido na reivindicação 8, em que o livro de codificação de formato glótico compreende um gerador de vetores-código contendo apenas um elemento não-zero e um filtro de forma para processar os vetores-código contendo apenas um elemento não-zero para produzir vetores-código que representam formatos
20 de impulsos glóticos centrados em diferentes posições.

11. Dispositivo de modo de transição, de acordo com o definido na reivindicação 8, em que os impulsos de formato glótico compreendem primeira e última amostras em que um número predeterminado das primeira e última amostras é truncado.

25 12. Dispositivo codificador para gerar uma excitação de modo de transição substituindo uma excitação de livro de codificação adaptativo em um quadro de transição e/ou um quadro seguindo a transição em um sinal de som, compreendendo:

30 um gerador de um sinal alvo de pesquisa no livro de codificação;
um livro de codificação de modo de transição para gerar um conjunto de vetores-código independente da excitação passada, os vetores-código do referido conjunto cada um dos quais correspondendo a uma exci-

tação de modo de transição respectiva;

um pesquisador do livro de codificação de modo de transição para encontrar o vetor-código do referido conjunto correspondendo a uma excitação de modo de transição otimamente correspondendo ao sinal alvo de pesquisa no livro de codificação.

13. Dispositivo codificador, de acordo com o definido na reivindicação 12, em que o livro de codificação de modo de transição compreende um livro de codificação fixo independente da excitação passada.

14. Dispositivo codificador, de acordo com o definido na reivindicação 12, em que o livro de codificação de modo de transição compreende um livro de codificação de formatos de impulsos glóticos.

15. Dispositivo codificador, de acordo com o definido na reivindicação 14, em que o pesquisador aplica um critério determinado a cada formato de impulso glótico do livro de codificação de formatos de impulsos glóticos e encontra como o vetor-código otimamente correspondendo ao sinal alvo de pesquisa do livro de codificação adaptativo, o vetor-código do conjunto correspondendo ao valor máximo do referido critério.

16. Dispositivo codificador, de acordo com o definido na reivindicação 15, em que o pesquisador identifica o vetor-código encontrado por meio de parâmetros de modo de transição selecionados a partir do grupo que consiste em uma identificação de configuração de modo de transição, um formato de impulso glótico, uma posição do centro de formato de impulso glótico no vetor-código encontrado, um ganho de modo de transição, um sinal do ganho de modo de transição e um período de tempo de alça fechada.

17. Dispositivo codificador, de acordo com o definido na reivindicação 12, em que o sinal de som compreende um sinal de fala e em que o quadro de transição é selecionado a partir do grupo que consiste em um quadro compreendendo um princípio sonante e um quadro compreendendo uma transição entre dois diferentes sons sonantes.

18. Dispositivo codificador, de acordo com o definido na reivindicação 12, em que o quadro de transição e/ou o quadro seguindo a transição compreende um quadro de transição seguido de diversos quadros.

19. Dispositivo codificador, de acordo com o definido na reivindicação 12, em que o quadro de transição e/ou o quadro seguindo a transição compreende pelo menos um quadro seguindo a transição.

5 20. Dispositivo codificador, de acordo com o definido na reivindicação 12, em que o quadro de transição e/ou o quadro seguindo a transição cada um dos quais compreende uma pluralidade de subquadros, e em que o pesquisador pesquisa o livro de codificação de modo de transição em uma primeira parte dos subquadros e um livro de codificação do tipo preditivo do dispositivo codificador em uma segunda parte dos subquadros.

10 21. Dispositivo codificador, de acordo com o definido na reivindicação 12, em que o livro de codificação de modo de transição compreende um livro de codificação de formato glótico compreendendo vetores-código formados de um formato de impulso glótico dispostos em uma posição específica no vetor-código.

15 22. Dispositivo codificador, de acordo com o definido na reivindicação 21, em que o livro de codificação de formato glótico inclui um número predeterminado de diferentes formatos de impulsos glóticos, e em que cada formato de impulso glótico é posicionado em uma pluralidade de diferentes posições no vetor-código para formar uma pluralidade de diferentes vetores-código do livro de codificação de formato glótico.

20 23. Dispositivo codificador, de acordo com o definido na reivindicação 21, em que o livro de codificação de formato glótico compreende um gerador de vetores-código contendo apenas um elemento não-zero e um filtro de forma para processar os vetores-código contendo apenas um elemento não-zero para produzir vetores-código que representam formatos de impulsos glóticos centrados em diferentes posições.

25 24. Dispositivo codificador, de acordo com o definido na reivindicação 12, adicionalmente compreendendo:

30 um gerador de um sinal de inovação alvo de pesquisa no livro de codificação;

um livro de codificação de inovação para gerar um conjunto de vetores-código de inovação cada um dos quais correspondendo à excitação

de inovação respectiva;

um pesquisador do livro de codificação de inovação para encontrar o vetor-código de inovação do referido conjunto correspondendo a uma excitação de inovação otimamente correspondendo ao sinal de inovação
5 alvo de pesquisa no livro de codificação; e

um somador da excitação de modo de transição e a excitação de inovação para produzir uma excitação global para um filtro de síntese de sinal de som.

25. Dispositivo codificador, de acordo com o definido na reivindicação 24, em que o quadro de transição e/ou o quadro seguindo a transição
10 cada um dos quais compreende uma pluralidade de subquadros e em que, dependendo de em que o impulso ou impulsos glóticos estão localizados nos subquadros, o dispositivo codificador compreende meios para codificar os subquadros usando pelo menos um de livro de codificação de modo de transição, livro de codificação adaptativo e livro de codificação de inovação.
15

26. Dispositivo decodificador para gerar uma excitação de modo de transição substituindo uma excitação de livro de codificação adaptativo em um quadro de transição e/ou um quadro seguindo a transição em um sinal de som, compreendendo:

20 uma entrada para receber um índice de livro de codificação;

um livro de codificação de modo de transição para gerar um conjunto de vetores-código independente da excitação passada, o livro de codificação de modo de transição respondendo ao índice para gerar no quadro de transição e/ou quadro seguindo a transição um dos vetores-código do
25 conjunto que corresponde à excitação de modo de transição.

27. Dispositivo decodificador, de acordo com o definido na reivindicação 26, em que o livro de codificação de modo de transição compreende um livro de codificação fixo independente da excitação passada.

28. Dispositivo decodificador, de acordo com o definido na reivindicação 26, em que o livro de codificação de modo de transição compreende um livro de codificação de formatos de impulsos glóticos.
30

29. Dispositivo decodificador, de acordo com o definido na rei-

vindicação 26, em que o sinal de som compreende um sinal de fala e em que o quadro de transição é selecionado a partir do grupo que consiste em um quadro compreendendo um princípio sonante e um quadro compreendendo uma transição entre dois diferentes sons sonantes.

5 30. Dispositivo decodificador, de acordo com o definido na reivindicação 26, em que o quadro de transição e/ou o quadro seguindo a transição cada um dos quais compreende uma pluralidade de subquadros, e em que o livro de codificação de modo de transição é usado em uma primeira parte dos subquadros e o dispositivo decodificador compreende um livro de
10 codificação do tipo preditivo que é usado em uma segunda parte dos subquadros.

 31. Dispositivo decodificador, de acordo com o definido na reivindicação 26, em que o livro de codificação de modo de transição compreende um livro de codificação de formato glótico compreendendo vetores-código formados de um formato de impulso glótico dispostos em uma posição específica no vetor-código.

 32. Dispositivo decodificador, de acordo com o definido na reivindicação 31, em que o livro de codificação de formato glótico inclui um número predeterminado de diferentes formatos de impulsos glóticos, e em
20 que cada formato de impulso glótico é posicionado em uma pluralidade de diferentes posições no vetor-código para formar uma pluralidade de diferentes vetores-código do livro de codificação de formato glótico.

 33. Dispositivo decodificador, de acordo com o definido na reivindicação 31, em que o livro de codificação de formato glótico compreende
25 um gerador de vetores-código contendo apenas um elemento não-zero e um filtro de forma para processar os vetores-código contendo apenas um elemento não-zero para produzir vetores-código que representam formatos de impulsos glóticos centrados em diferentes posições.

 34. Dispositivo decodificador, de acordo com o definido na reivindicação 26, adicionalmente compreendendo:

uma entrada para receber um índice de livro de codificação de inovação;

um livro de codificação de inovação para gerar um conjunto de vetores-código de inovação, o livro de codificação de inovação respondendo ao índice de livro de codificação de inovação para gerar no quadro de transição e/ou quadro seguindo a transição um dos vetores-código de inovação do conjunto que corresponde a uma excitação de inovação; um somador da excitação de modo de transição e a excitação de inovação para produzir uma excitação global para um filtro de síntese de sinal de som.

35. Método de modo de transição para uso em um Codec de sinal de som do tipo preditivo para produzir uma excitação de modo de transição substituindo uma excitação de livro de codificação adaptativo em um quadro de transição e/ou um quadro seguindo a transição no sinal de som, compreendendo:

proporcionar um livro de codificação de modo de transição para gerar um conjunto de vetores-código independente da excitação passada;

15 fornecer um índice de livro de codificação do livro de codificação de modo de transição; e

gerar, por meio do livro de codificação de modo de transição e em resposta ao índice do livro de codificação, um dos vetores-código do conjunto que corresponde à referida excitação de modo de transição.

20 36. Método de modo de transição, de acordo com o definido na reivindicação 35, em que o livro de codificação de modo de transição compreende um livro de codificação fixo independente da excitação passada.

25 37. Método de modo de transição, de acordo com o definido na reivindicação 35, em que o livro de codificação de modo de transição compreende um livro de codificação de formatos de impulsos glóticos.

30 38. Método de modo de transição, de acordo com o definido na reivindicação 35, em que o sinal de som compreende um sinal de fala e o referido método compreende selecionar o quadro de transição a partir do grupo consistindo de um quadro compreendendo um princípio sonante e um quadro compreendendo uma transição entre dois diferentes sons sonantes.

39. Método de modo de transição, de acordo com o definido na reivindicação 35, em que o quadro de transição e/ou o quadro seguindo a

transição compreende um quadro de transição seguido de diversos quadros.

40. Método de modo de transição, de acordo com o definido na reivindicação 35, em que o quadro de transição e/ou o quadro seguindo a transição compreende pelo menos um quadro seguindo a transição.

5 41. Método método de modo de transição, de acordo com o definido na reivindicação 35, em que o quadro de transição e/ou o quadro seguindo a transição cada um dos quais compreende uma pluralidade de subquadros, e o referido método compreende usando o livro de codificação de modo de transição em uma primeira parte dos subquadros e um livro de codificação do tipo preditivo do Codec do tipo preditivo em uma segunda parte dos subquadros.

42. Método de modo de transição, de acordo com o definido na reivindicação 35, em que proporcionar um livro de codificação de modo de transição compreende proporcionar um livro de codificação de formato glótico compreendendo vetores-código formados de um formato de impulso glótico dispostos em uma posição específica no vetor-código.

43. Método de modo de transição, de acordo com o definido na reivindicação 42, em que proporcionar um livro de codificação de formato glótico compreende proporcionar um livro de codificação de formato glótico incluindo um número predeterminado de diferentes formatos de impulsos glóticos e formar no livro de codificação de formato glótico uma pluralidade de diferentes vetores-código ao posicionar cada formato de impulso glótico em uma pluralidade de diferentes posições no vetor-código.

44. Método de modo de transição, de acordo com o definido na reivindicação 42, compreendendo, no livro de codificação de formato glótico, gerar vetores-código contendo apenas um elemento não-zero e processar através de um filtro de forma os vetores-código contendo apenas um elemento não-zero para produzir vetores-código que representam formatos de impulsos glóticos centrados em diferentes posições.

30 45. Método de modo de transição, de acordo com o definido na reivindicação 42, em que os impulsos de formato glótico compreendem primeira e última amostras, o referido método compreendendo truncar um nú-

mero predeterminado das primeira e última amostras.

5 46. Método de codificação para gerar uma excitação de modo de transição substituindo uma excitação de livro de codificação adaptativo em um quadro de transição e/ou um quadro seguindo a transição em um sinal de som, compreendendo:

gerar um sinal alvo de pesquisa no livro de codificação;

10 proporcionar um livro de codificação de modo de transição para gerar um conjunto de vetores-código independente da excitação passada, os vetores-código do referido conjunto cada um dos quais correspondendo a uma excitação de modo de transição respectiva;

pesquisar o livro de codificação de modo de transição para encontrar o vetor-código do referido conjunto correspondendo a uma excitação de modo de transição otimamente correspondendo ao sinal alvo de pesquisa no livro de codificação.

15 47. Método de codificação, de acordo com o definido na reivindicação 65, em que proporcionar um livro de codificação de modo de transição compreende proporcionar um livro de codificação fixo independente da excitação passada.

20 48. Método de codificação, de acordo com o definido na reivindicação 46, em que proporcionar um livro de codificação de modo de transição compreende proporcionar um livro de codificação de formatos de impulsos glóticos.

25 49. Método de codificação, de acordo com o definido na reivindicação 48, em que pesquisar o livro de codificação de modo de transição compreende aplicar um determinado critério a cada formato de impulso glótico do livro de codificação de formatos de impulsos glóticos e encontra como o vetor-código otimamente correspondendo ao sinal alvo de pesquisa do livro de codificação adaptativo o vetor-código do conjunto que corresponde ao valor máximo do referido critério.

30 50. Método de codificação, de acordo com o definido na reivindicação 49, em que pesquisar o livro de codificação de modo de transição compreende identificar o vetor-código encontrado por meio de parâmetros

de modo de transição selecionados a partir do grupo que consiste em uma identificação de configuração de modo de transição, um formato de impulso glótico, uma posição do centro de formato de impulso glótico no vetor-código encontrado, um ganho de modo de transição, um sinal do ganho de modo de transição e um período de tempo de alça fechada.

51. Método de codificação, de acordo com o definido na reivindicação 46, em que o sinal de som compreende um sinal de fala e o referido método adicionalmente compreende selecionar o quadro de transição a partir do grupo consistindo de um quadro compreendendo um princípio sonante e um quadro compreendendo uma transição entre dois diferentes sons sonantes.

52. Método de codificação, de acordo com o definido na reivindicação 46, em que o quadro de transição e/ou o quadro seguindo a transição compreende um quadro de transição seguido de diversos quadros.

53. Método de codificação, de acordo com o definido na reivindicação 46, em que o quadro de transição e/ou o quadro seguindo a transição compreende pelo menos um quadro seguindo a transição.

54. Método de codificação, de acordo com o definido na reivindicação 46, em que o quadro de transição e/ou o quadro seguindo a transição cada um dos quais compreende uma pluralidade de subquadros, e em que pesquisar o livro de codificação de modo de transição compreende pesquisar o livro de codificação de modo de transição em uma primeira parte dos subquadros e pesquisar um livro de codificação do tipo preditivo do dispositivo codificador em uma segunda parte dos subquadros.

55. Método de codificação, de acordo com o definido na reivindicação 46, em que proporcionar um livro de codificação de modo de transição compreende proporcionar um livro de codificação de formato glótico compreendendo vetores-código formados de um formato de impulso glótico dispostos em uma posição específica no vetor-código.

56. Método de codificação, de acordo com o definido na reivindicação 55, em que proporcionar um livro de codificação de formato glótico compreende proporcionar um livro de codificação de formato glótico inclu-

do um número predeterminado de diferentes formatos de impulsos glóticos, e formar uma pluralidade de diferentes vetores-código do livro de codificação de formato glótico ao posicionar cada formato de impulso glótico em uma pluralidade de diferentes posições no vetor-código.

5 57. Método de codificação, de acordo com o definido na reivindicação 55, em que gerar no livro de codificação de formato glótico um conjunto de vetores-código independente da excitação passada compreende gerar vetores-código contendo apenas um elemento não-zero e processar através de um filtro de forma os vetores-código contendo apenas um elemento não-
10 zero para produzir vetores-código que representam formatos de impulsos glóticos centrados em diferentes posições.

58. Método de codificação, de acordo com o definido na reivindicação 46, adicionalmente compreendendo:

15 gerar um sinal de inovação alvo de pesquisa no livro de codificação;

proporcionar um livro de codificação de inovação para gerar um conjunto de vetores-código de inovação cada um dos quais correspondendo à excitação de inovação respectiva;

20 pesquisar o livro de codificação de inovação para encontrar o vetor-código de inovação do referido conjunto correspondendo a uma excitação de inovação otimamente correspondendo ao sinal de inovação alvo de pesquisa no livro de codificação; e

25 adicionar a excitação de modo de transição e a excitação de inovação para produzir uma excitação global para um filtro de síntese de sinal de som.

59. Método de codificação, de acordo com o definido na reivindicação 58, em que o quadro de transição e/ou o quadro seguindo a transição cada um dos quais compreende uma pluralidade de subquadros e em que, dependendo de em que o impulso ou impulsos glóticos estão localizados nos
30 subquadros, o método de codificação compreende codificar os subquadros usando pelo menos um do livro de codificação de modo de transição, o livro de codificação adaptativo e o livro de codificação de inovação.

60. Método de decodificação para gerar uma excitação de modo de transição substituindo uma excitação de livro de codificação adaptativo em um quadro de transição e/ou um quadro seguindo a transição em um sinal de som, compreendendo:

- 5 receber um índice de livro de codificação;
 fornecer o índice do livro de codificação to um livro de codificação de modo de transição para gerar um conjunto de vetores-código independente da excitação passada; e
 gerar, por meio do livro de codificação de modo de transição e
 10 em resposta ao índice do livro de codificação, um dos vetores-código do conjunto que corresponde à excitação de modo de transição.

61. Método de decodificação, de acordo com o definido na reivindicação 60, em que o livro de codificação de modo de transição compreende um livro de codificação fixo independente da excitação passada.

- 15 62. Método de decodificação, de acordo com o definido na reivindicação 60, compreendendo proporcionar as o livro de codificação de modo de transição um livro de codificação de formatos de impulsos glóticos.

- 20 63. Método de decodificação, de acordo com o definido na reivindicação 60, em que o sinal de som compreende um sinal de fala e em que o referido método compreende selecionar o quadro de transição a partir do grupo consistindo de um quadro compreendendo um princípio sonante e um quadro compreendendo uma transição entre dois diferentes sons sonantes.

- 25 64. Método de decodificação, de acordo com o definido na reivindicação 60, em que o quadro de transição e/ou o quadro seguindo a transição cada um dos quais compreende uma pluralidade de subquadros, e em que o referido método compreende usando o livro de codificação de modo de transição em uma primeira parte dos subquadros e um livro de codificação do tipo preditivo em uma segunda parte dos subquadros.

- 30 65. Método de decodificação, de acordo com o definido na reivindicação 60, compreendendo proporcionar como o livro de codificação de modo de transição um livro de codificação de formatos glóticos compreen-

dendo vetores-código formados a partir de um formato de impulso glótico disposto em uma posição específica no vetor-código.

5 66. Método de decodificação, de acordo com o definido na reivindicação 65, em que o livro de codificação de formato glótico inclui um número predeterminado de diferentes formatos de impulsos glóticos, e em que o referido método compreende formar uma pluralidade de diferentes vetores-código do livro de codificação de formato glótico ao posicionar cada um dos impulsos glóticos em uma pluralidade de diferentes posições no vetor-código.

10 67. Método de decodificação, de acordo com o definido na reivindicação 65, em que vetores-código do conjunto são gerados pelo livro de codificação de formato glótico ao gerar vetores-código contendo apenas um elemento não-zero e processar através do filtro de forma os vetores-código contendo apenas um elemento não-zero para produzir vetores-código que
15 representam o formato de impulsos glóticos centrados em diferentes posições.

68. Método de decodificação, de acordo com o definido na reivindicação 60, adicionalmente compreendendo:

20 proporcionar um livro de codificação de inovação para gerar um conjunto de vetores-código de inovação;

fornecer um índice de livro de codificação de inovação ao livro de codificação de inovação;

25 gerar, por meio do livro de codificação de inovação e em resposta ao índice de livro de codificação de inovação, um dos vetores-código de inovação do conjunto correspondendo à uma excitação de inovação; e

adicionar a excitação de modo de transição e a excitação de inovação para produzir uma excitação global para um filtro de síntese de sinal de som.

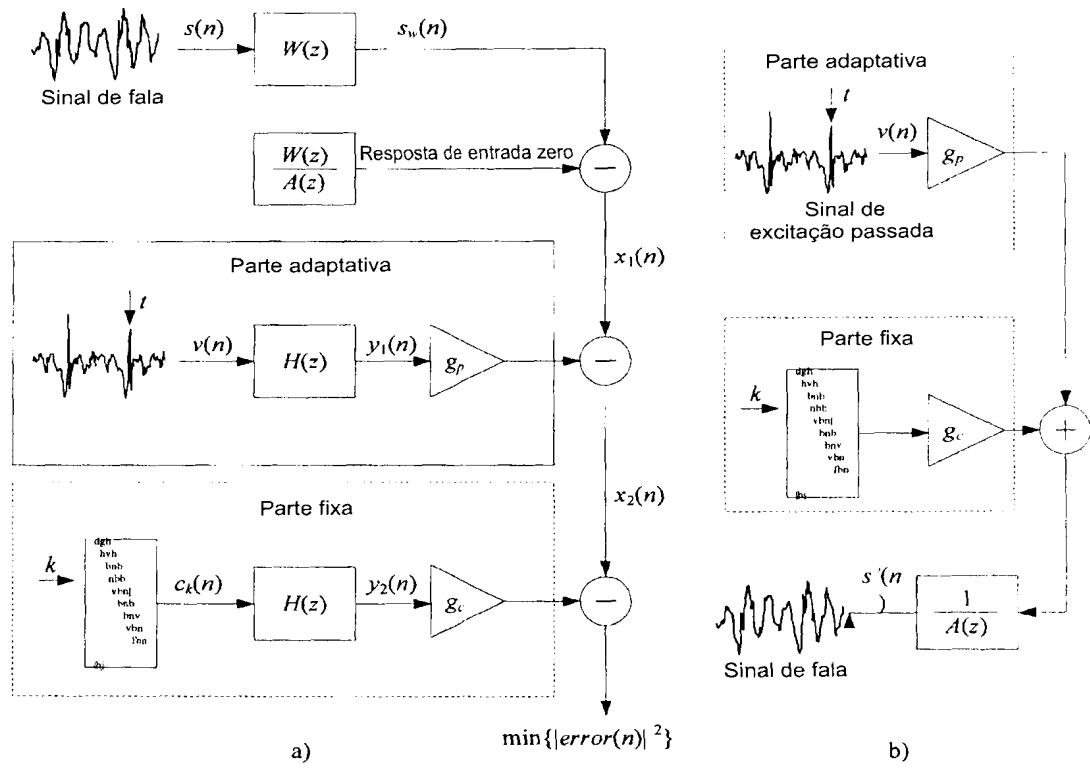


Fig 1

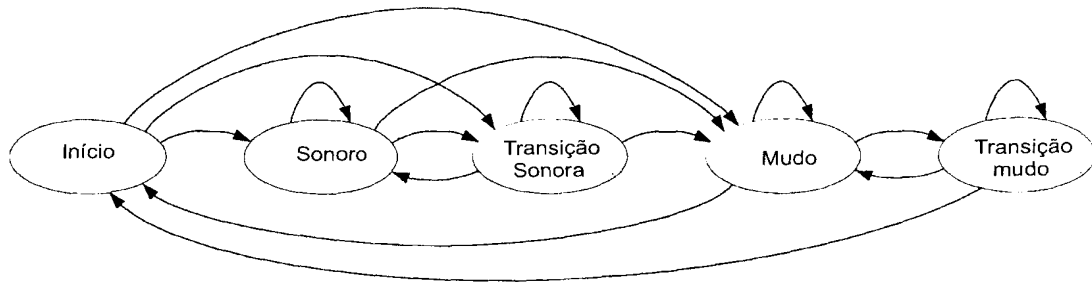


Fig 2

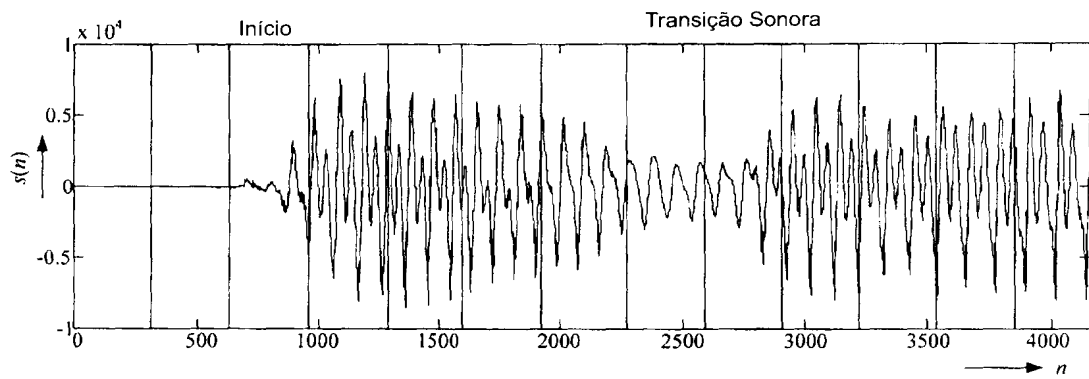


Fig 3

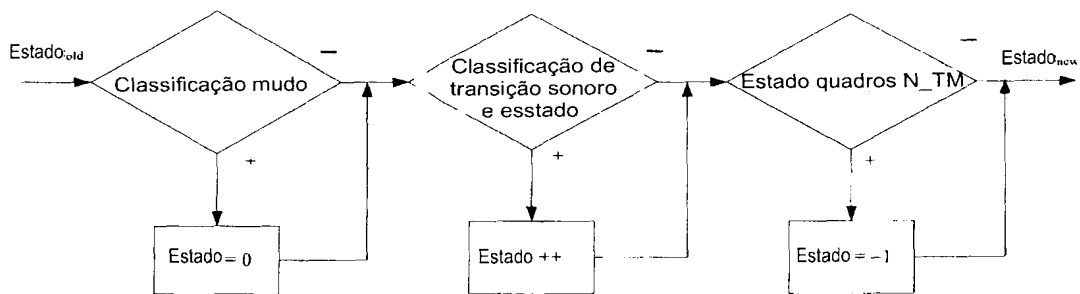


Fig 4

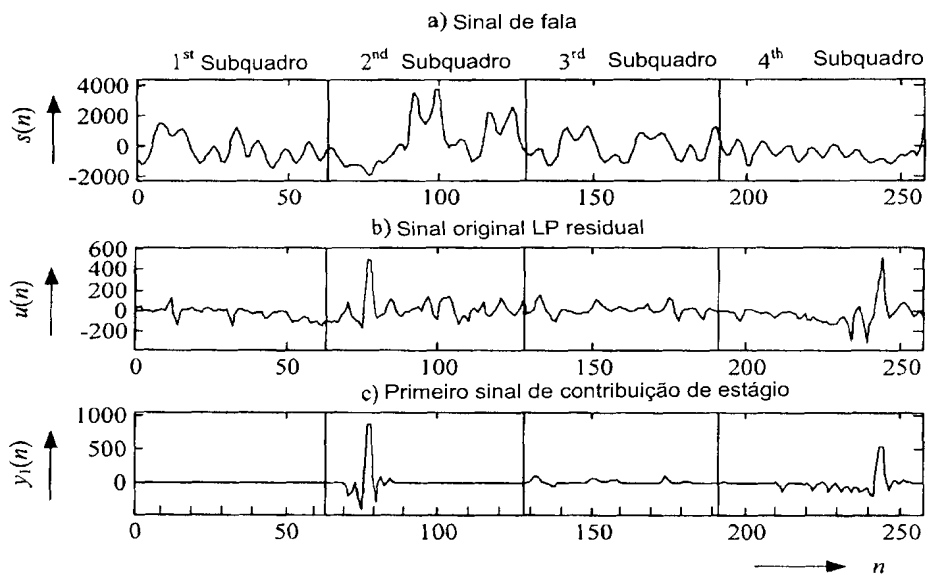


Fig 5

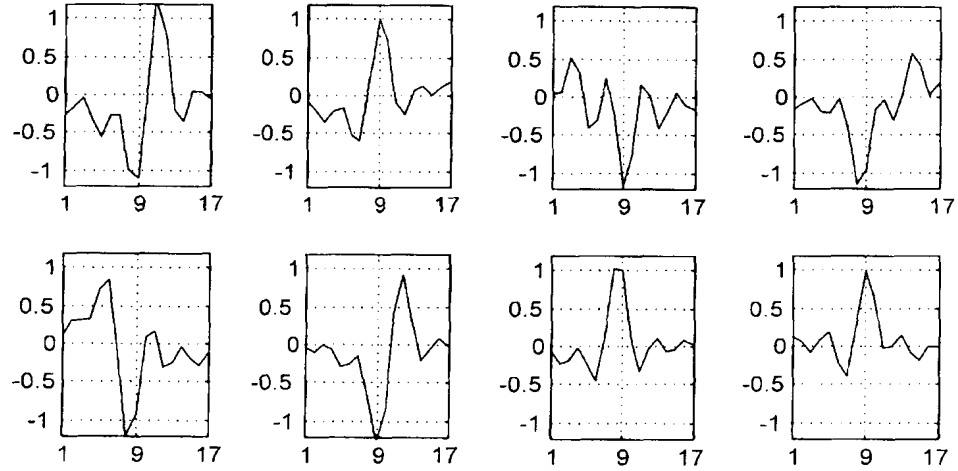


Fig 6

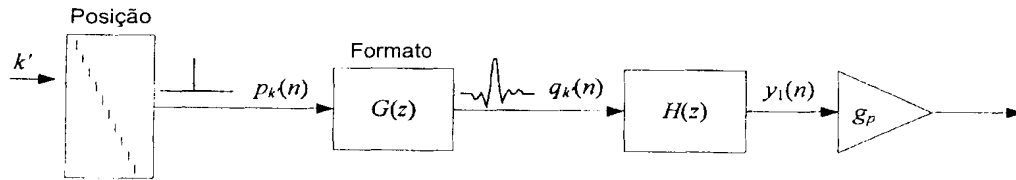


Fig 7

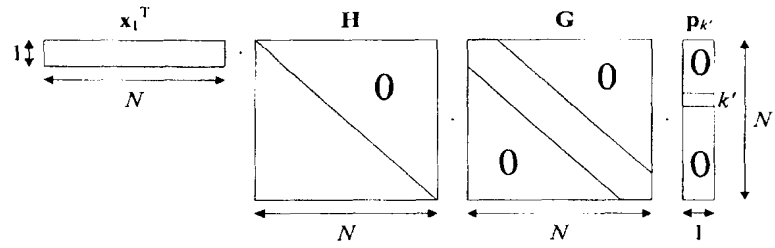


Fig 8

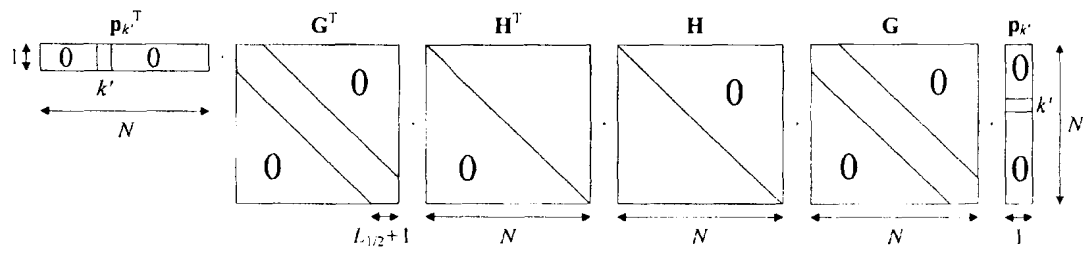


Fig 9

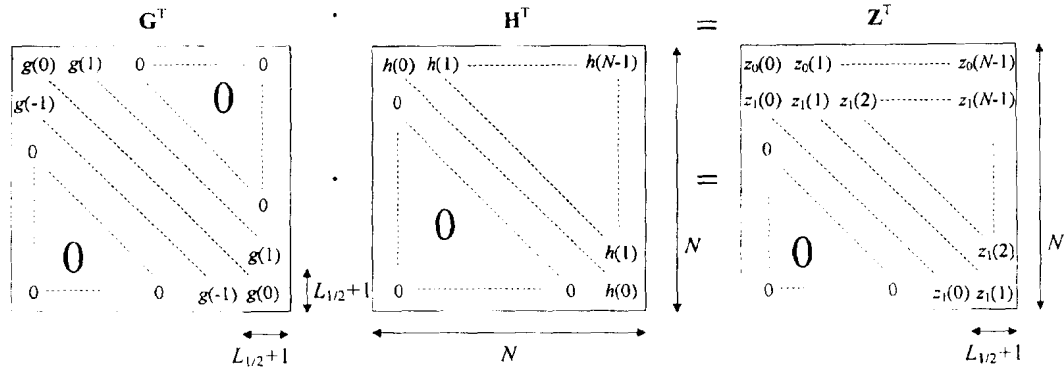


Fig 10

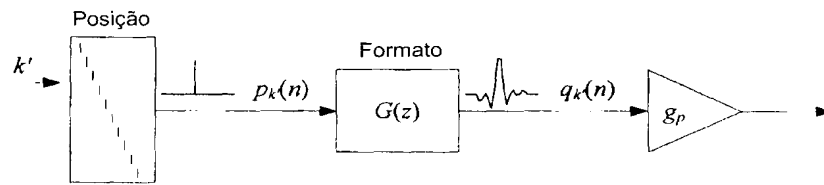
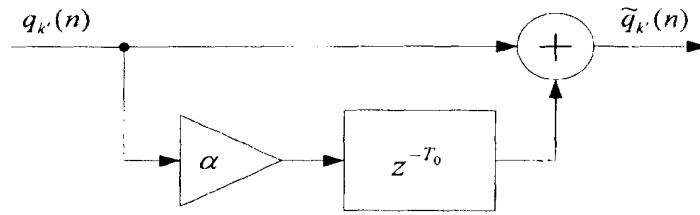
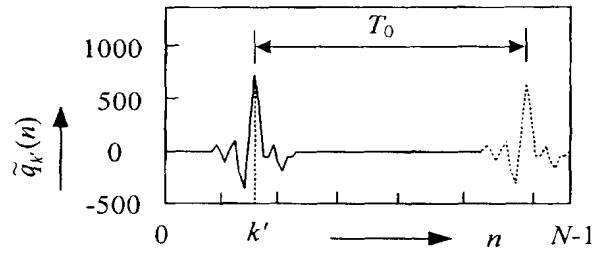


Fig 11



a)



b)

Fig 12

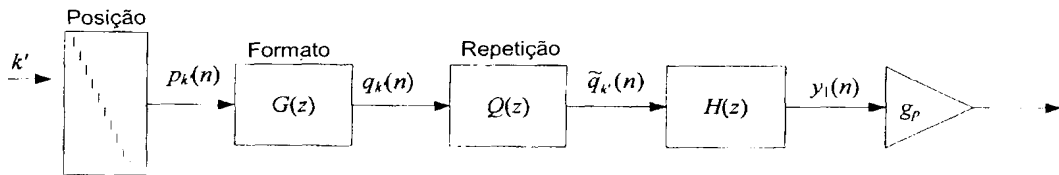


Fig 13

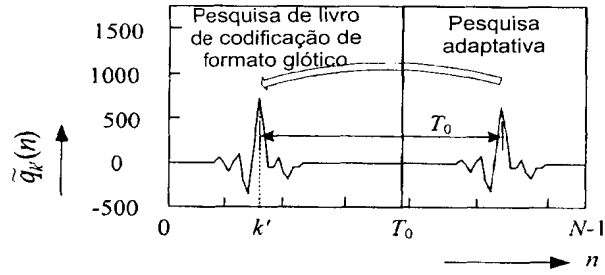


Fig 14

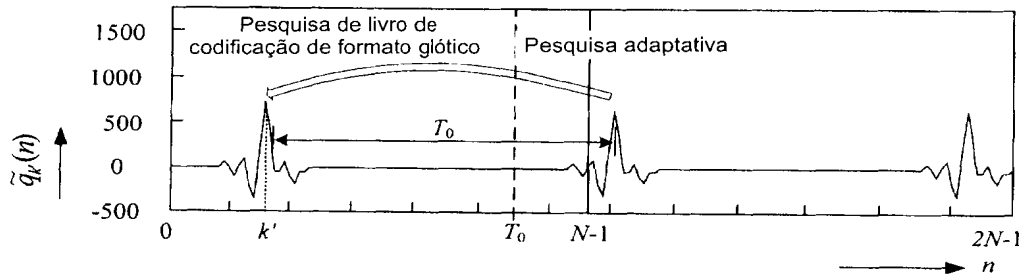


Fig 15

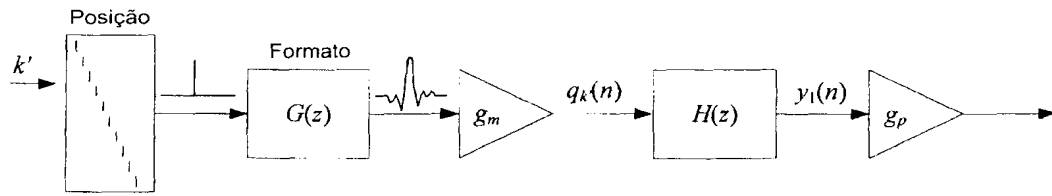


Fig 16

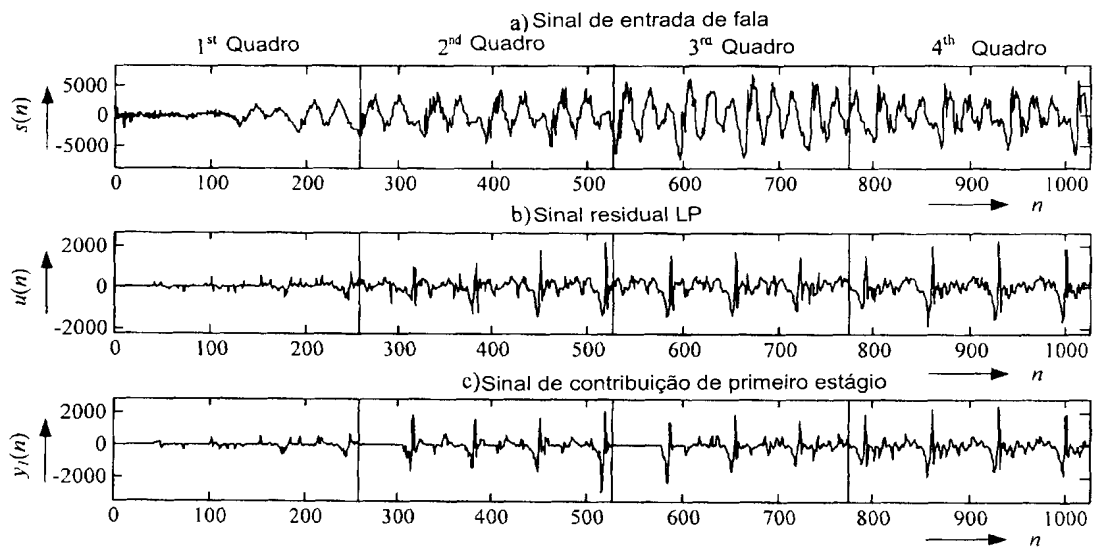


Fig 17

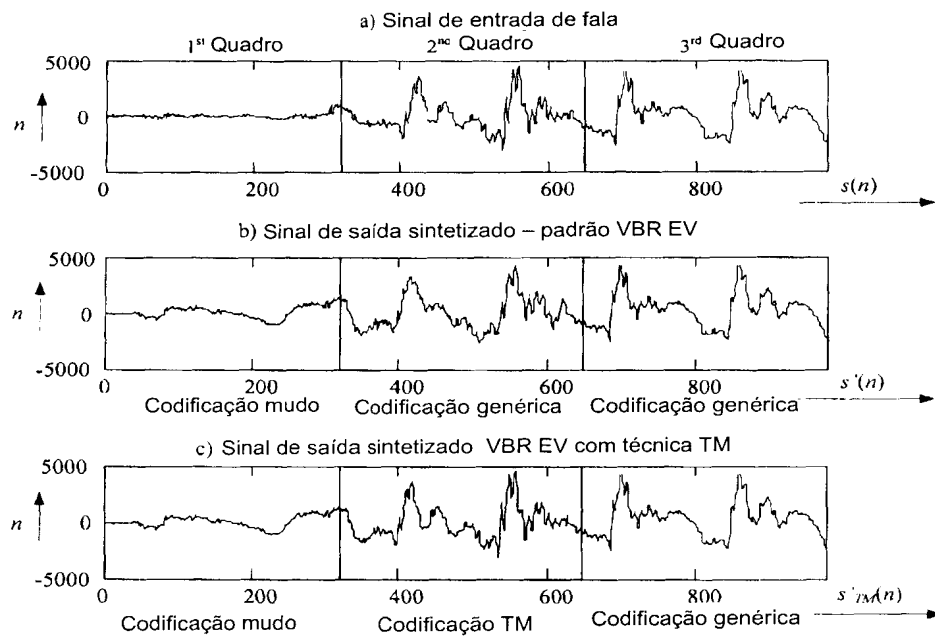
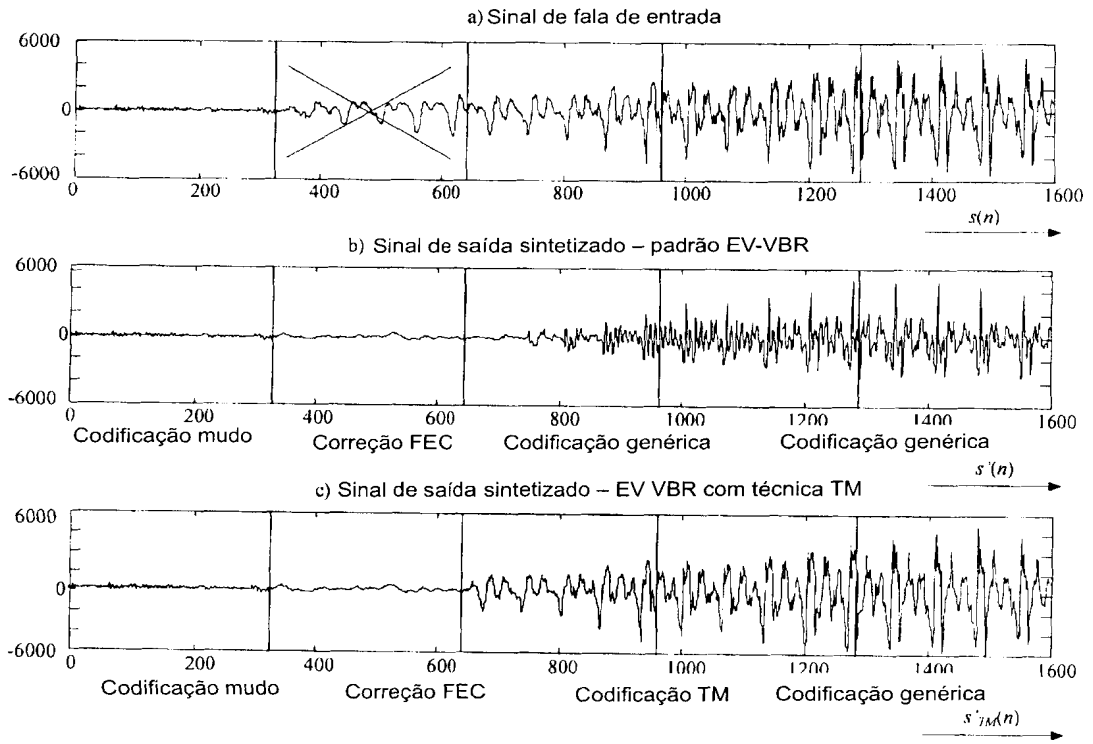


Fig 18



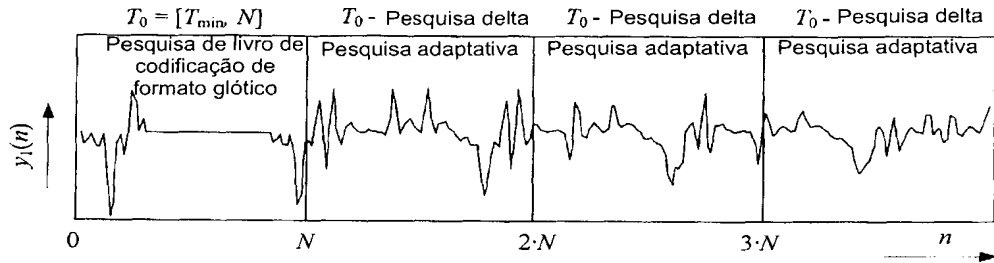


Fig 20

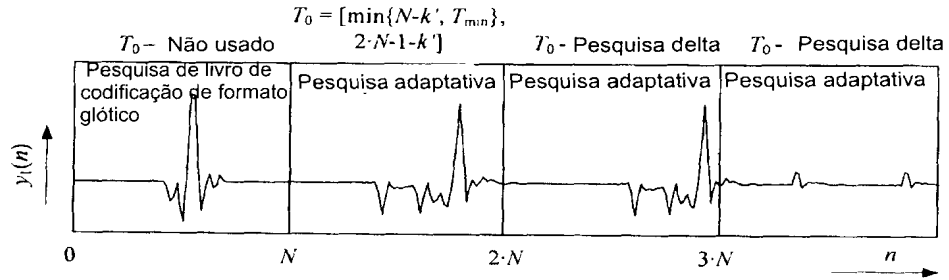


Fig 21

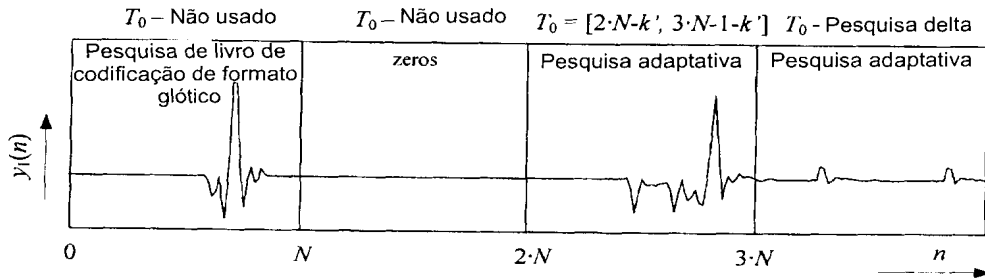


Fig 22

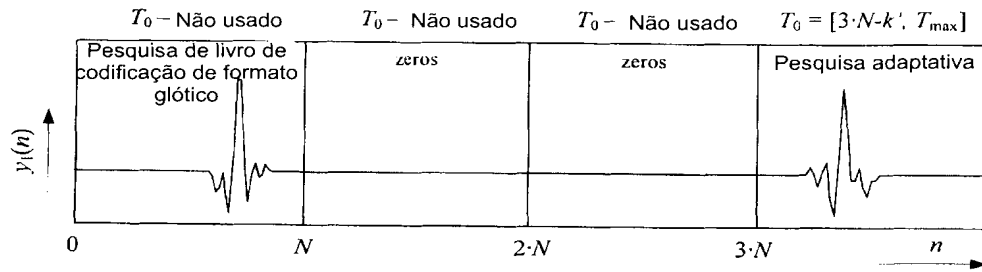


Fig 23

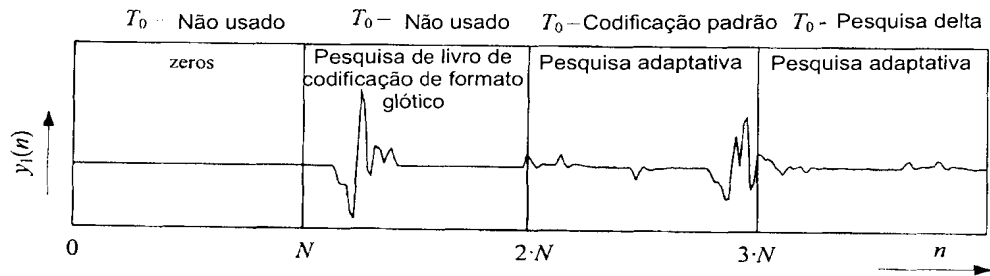


Fig 24

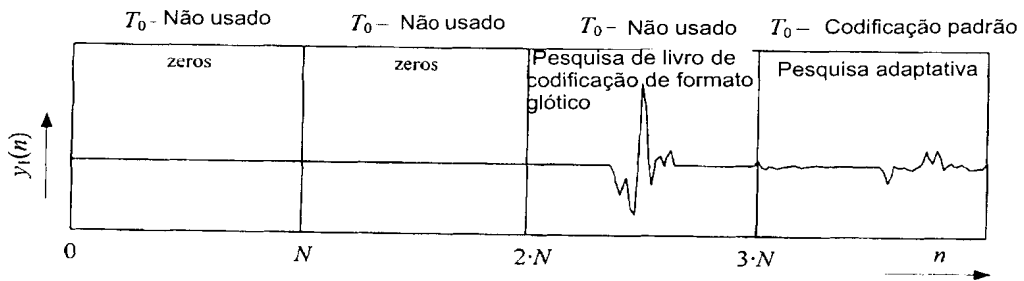


Fig 25

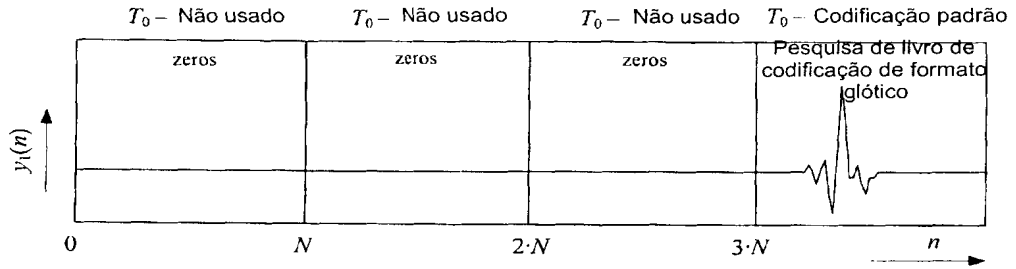


Fig 26

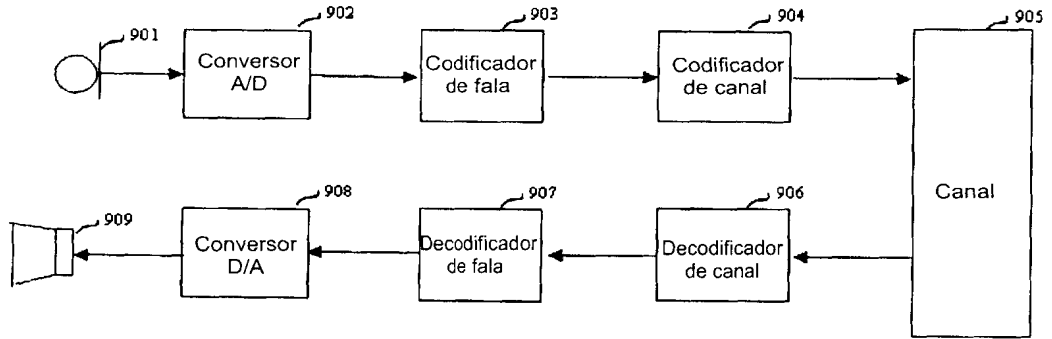


Fig 27

RESUMO

Patente de Invenção: **"MÉTODO E DISPOSITIVO PARA CODIFICAR QUADROS DE TRANSIÇÃO EM SINAIS DE FALA"**.

A presente invenção refere-se a um dispositivo de modo de transição e método para uso em um Codec de sinal de som do tipo preditivo para produzir uma excitação de modo de transição substituindo uma excitação de livro de codificação adaptativo em um quadro de transição e/ou um quadro seguindo a transição no sinal de som, compreendendo uma entrada para receber um índice de livro de codificação e um livro de codificação de modo de transição para gerar um conjunto de vetores-código independente da excitação passada. O livro de codificação de modo de transição responde ao índice para gerar, no quadro de transição e/ou quadro seguindo a transição, um dos vetores-código do conjunto que corresponde à excitação de modo de transição. É também proporcionado um dispositivo codificador e método e um dispositivo decodificador e método que usa dispositivo de modo de transição e método acima descritos.