



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

(11) PI 0317954-0 B1

(22) Data do Depósito: 22/12/2003

(45) Data de Concessão: 03/01/2017



(54) Título: PROCESSO DE CODIFICAÇÃO E DECODIFICAÇÃO ÁUDIO COM TAXA VARIÁVEL

(51) Int.Cl.: G10L 19/00; H03M 7/40

(52) CPC: G10L 19/00,H03M 7/40

(30) Prioridade Unionista: 08/01/2003 FR 03 00164

(73) Titular(es): FRANCE TELECOM

(72) Inventor(es): BALAZS KOVESI; DOMINIQUE MASSALOUX

Relatório Descritivo da Patente de Invenção para **"PROCESSO DE CODIFICAÇÃO E DECODIFICAÇÃO ÁUDIO COM TAXA VARIÁVEL"**.

A presente invenção refere-se aos dispositivos de codificação e de decodificação dos sinais áudio, destinados notadamente a ocupar lugar em aplicações de transmissão ou de armazenagem dos sinais áudio (fala e/ou sons) numerados e comprimidos.

Mais particularmente, essa invenção é relativa aos sistemas de codificação áudio tendo a capacidade de fornecer taxas variadas, ainda denominados de codificação multitaxa. Esses sistemas se distinguem dos codificadores com taxa fixa por sua capacidade de modificar a taxa da codificação, eventualmente em curso de tratamento, o que é particularmente adaptado para transmissão sobre redes de acessos heterogêneos, quer se trate de redes de tipo IP, misturando acessos fixos e móveis, elevadas taxas (ADSL), baixas taxas (modems RTC, GPRS), ou fazendo intervir terminais de capacidades variáveis (móveis, PC,...).

Distinguem-se essencialmente duas categorias de codificadores multitaxa: aquela dos codificadores multitaxa "comutáveis" e aquela dos codificadores "hierárquicos".

Os codificadores multitaxa "comutáveis" se baseiam em uma estrutura de codificação pertencente a uma família tecnológica (codificação temporal, ou freqüencial, por exemplo: CELP, sinusoidal, ou por transformada), na qual uma indicação de taxa é simultaneamente fornecida ao codificador e ao decodificador. O codificador utiliza essa informação para selecionar as partes do algoritmo e as tabelas pertinentes para a taxa escolhida. O decodificador opera de maneira simétrica. Numerosas estruturas de codificação multitaxa comutáveis foram propostas para a codificação áudio. É o caso, por exemplo, dos codificadores móveis normalizados pela organização 3GPP ("3rd Generation Partnership Project"), o NB-AMR("Narrow Band Adaptive Multi-Rate", Specification Technique 3GPP TS 26090, versão 5.0.0, junho de 2002) em bandatelefônica, ou WB-AMR ("Wide Band Adaptive Multi-Rate", Specification Technique 3GPP TS 26.190, versão 5.1.0, dezembro 2001) em banda larga. Esses codificadores operam em muitas bandas lar-

gas de taxas (4,75 a 12,2 kbit/s para o NB-AMR, 6,60 a 23,85 kbits/s para o WB-AMR), com uma granularidade bastante importante (8 taxas para o NB-AMR e 9 para o WB-AMR). Todavia, o preço a pagar para essa flexibilidade é uma complexidade de estrutura bem conseqüente: para chegar a atingir

5 todas essas taxas, esses codificadores devem suportar numerosas opções diferentes, tabelas de quantificação variadas, etc. A curva dos desempenhos aumenta progressivamente com a taxa, mas a progressão não é linear e certas taxas são, por essência, melhor otimizadas do que outras.

Nos sistemas de codificação ditos "hierárquicos", ainda denomi-

10 nados "escaláveis", os dados binários provenientes da operação de codificação se repartem em camadas sucessivas. Uma camada de base, ainda denominada "núcleo", é formada dos elementos binários absolutamente necessários à decodificação do trem binário, e determinando uma qualidade mínima de decodificação.

15 As camadas seguintes permitem melhorar progressivamente a qualidade do sinal proveniente da operação de decodificação, cada nova camada levando novas informações, que, exploradas pelo decodificador, fornecem na saída um sinal de qualidade crescente.

Uma das particularidades da codificação hierárquica é a possibi-

20 lidade oferecida de intervir em qualquer nível da cadeia de transmissão ou de estocagem para suprimir uma parte do trem binário, sem ter de fornecer indicação particular ao codificador, nem ao decodificador. O decodificador utiliza as informações binárias que ele recebe e produz um sinal de qualidade correspondente.

25 O domínio das estruturas de codificação hierárquicas deu origem também a numerosos trabalhos. Certas estruturas de codificação hierárquicas funcionam a partir de um só tipo de codificador, concebido para liberar informações codificadas hierarquizadas. Quando as camadas suplementares melhoram a qualidade do sinal de saída, sem modificar a banda passante,

30 fala-se antes de tudo de "codificadores imbricados" (ver, por exemplo, R.D. Iacovo et al., Embedded CELP Coding For Variable Bit-Rate Between 6.4 and 9.6 kbit/s", Proc. ICASSP 1991, pp. 681-686). Esse tipo de codificadores

não permite todavia grandes desvios entre a mais baixa e a mais alta taxa propostos.

A hierarquia é freqüentemente utilizada para aumentar progressivamente a banda passante do sinal: o núcleo fornece um sinal em banda de base, por exemplo, telefônica (300-3400 Hz), e as camadas seguintes permitem a codificação de bandas freqüenciais suplementares (por exemplo, banda larga até 7 kHz, banda HiFi até 20 kHz ou intermediárias,...). Os codificadores em sub-bandas ou os codificadores utilizando uma transformação tempo-freqüência, tais como descrito nos documentos "Subband/transform coding using filter banks designs based on time doain aliasing cancellation: de J.P. Princen et al. (proc. IEEE ICASSP-97, pp. 2161-2164) e "High Quality Audio Transform Coding at 64 kbit/s", de y. Mahieux et al. (IEEE Trans. Commun, Vol. 42, N° 11, novembro de 1994, pp. 3010-3019) se prestam particularmente a essas operações.

Por outro lado, é freqüente aplicar uma técnica de codificação diferente para o núcleo e para o(s) módulo(s) codificando as camadas suplementares, fala-se então de diferentes estágios de codificação, cada estágio sendo constituído de um subcodificador. O subcodificador do estágio de um nível determinado poderá seja codificar partes do sinal não codificadas pelos estágios precedentes, seja codificar o resíduo de codificação do estágio precedente, o resíduo é obtido subtraindo o sinal decodificado do sinal original.

A vantagem dessas estruturas é que elas permitem descer a taxas relativamente baixas com uma qualidade suficiente, produzindo uma boa qualidade de elevada taxa. Com efeito, as técnicas aplicadas para as baixas taxas não são geralmente eficazes às elevadas taxas e inversamente.

Essas estruturas permitem utilizar duas tecnologias diferentes (por exemplo, CELP e transformada tempo-freqüência,...) são particularmente eficazes para varrer grandes faixas de taxas.

Todavia, as estruturas de codificação hierárquicas propostas na técnica anterior definem precisamente a posição atribuída a cada uma das

camadas intermediárias. Cada camada corresponde à codificação de certos parâmetros, e a granularidade do trem binário hierárquico depende da taxa atribuída a esses parâmetros (tipicamente uma camada pode conter da ordem de algumas dezenas de bits por trama, uma trama de sinal sendo constituída de um certo número de amostras do sinal em uma duração determinada, o exemplo descrito mais adiante, considerando-se uma trama de 960 amostras, correspondendo a 60 ms de sinal).

Além disso, quando a banda passante dos sinais decodificados pode variar segundo nível das camadas de elementos binários, a modificação da taxa em linha pode produzir artefatos incômodos, quando da escuta.

A presente invenção tem notadamente por finalidade propor uma solução de codificação multitaxa que previne os inconvenientes citados no caso da utilização das codificações comutáveis e hierárquicas existentes.

A invenção propões assim um processo de codificação de uma trama de sinal áudio numérico em uma seqüência binária de saída, no qual um número máximo N_{max} de bits de codificação é definido para um conjunto de parâmetros calculáveis a partir da trama de sinal, composto de um primeiro e de um segundo subconjuntos. O processo proposto compreende as seguintes etapas:

- calculam-se os parâmetros do primeiro subconjunto e codificam-se esses parâmetros sobre um número N_0 de bits de codificação, tal como $N_0 < N_{max}$;
- determina-se uma alocação de $N_{max} - N_0$ bits de codificação para os parâmetros do segundo subconjunto; e
- classificam-se os $N_{max} - N_0$ bits de codificação alocados nos parâmetros do segundo subconjunto em uma ordem determinada.

A alocação e/ou a ordem de classificação dos $N_{max} - N_0$ bits de codificação são determinados em função dos parâmetros codificados do primeiro subconjunto. O processo de codificação compreende, além disso, as seguintes etapas em resposta à indicação de um número N de bits da seqüência binária de saída disponíveis para a codificação desse conjunto de parâmetros, com $N_0 < N < N_{max}$:

- selecionam-se os parâmetros do segundo subconjunto nos quais são alocados os $N-N_0$ bits de codificação classificados os primeiros nessa ordem;

- calculam-se os parâmetros selecionados do segundo subconjunto; e codificam-se esses parâmetros para produzir os $N-N_0$ bits de codificação classificados primeiros; e

- inserem-se na seqüência de saída os N_0 bits de codificação do primeiro subconjunto, assim como os $N-N_0$ bits de codificação dos parâmetros selecionados do segundo subconjunto.

10 O processo, de acordo com a invenção, permite definir uma codificação multitaxa, que funcionará pelo menos uma faixa correspondente para cada trama a um número de bits que vai de N_0 a $N_{\max}$.

Pode-se assim considerar que a noção de taxas preestabelecidas que é ligada às codificações comutáveis e hierárquicas existentes é substituída por uma noção de cursor, permitindo fazer variar livremente a taxa entre um valor mínimo (podendo eventualmente corresponder a um número de bits N inferior a N_0) e um valor máximo (correspondente a $N_{\max}$). Esses valores extremos são potencialmente afastados. O processo oferece bons desempenhos em termos de eficácia de codificação, independentemente da taxa escolhida.

Vantajosamente, o número de N de bits da seqüência binária de saída é estritamente inferior a $N_{\max}$. O codificador tem então de notável que a alocação dos bits empregada não faz referência à taxa de saída efetiva do codificador, mas a um outro número $N_{\max}$ convencionado como decodificador.

É, todavia, possível fixar $N_{\max} = N$ em função da taxa instantânea disponível sobre um canal; de transmissão. A seqüência de saída desse codificador multitaxa comutável poderá ser tratada por um decodificador que não receberia a totalidade da seqüência, desde quando é capaz de encontrar a estrutura dos bits de codificação do segundo subconjunto, graças ao conhecimento de $N_{\max}$.

Um outro caso em que se pode ter $N = N_{\max}$ é aquele da arma-

zenagem de dados áudio à taxa de codificação máxima. Quando de uma leitura de N' bits desse conteúdo armazenado com taxa mais baixa, o decodificador será capaz de encontrar a estrutura dos bits de codificação do segundo subconjunto desde quando $N' \geq N_0$.

5 A ordem de classificação dos bits de codificação alocados nos parâmetros do segundo subconjunto pode ser uma ordem preestabelecida.

Em um modo de realização preferido, a ordem de classificação dos bits de codificação alocados nos parâmetros do segundo subconjunto é variável. Pode notadamente ser uma ordem de importância decrescente determinada em função de pelo menos os parâmetros codificados do primeiro subconjunto. Assim, o decodificador que receberá uma sequência binária de N' bits para a trama, com $N_0 \leq N' \leq N \leq N_{\max}$, poderá deduzir essa ordem dos N_0 bits recebidos para a codificação do primeiro subconjunto.

10

A alocação dos $N_{\max} - N_0$ bits na codificação dos parâmetros do segundo subconjunto pode ser realizada de forma fixa (nesse caso, a ordem de classificação desses bits será função pelo menos dos parâmetros codificados do primeiro subconjunto).

15

Em um modo de realização preferido, a alocação dos $N_{\max} - N_0$ bits na codificação dos parâmetros do segundo subconjunto é uma função dos parâmetros codificados do primeiro subconjunto.

20

Vantajosamente, essa ordem de classificação dos bits de codificação alocados nos parâmetros do segundo subconjunto é determinado com o auxílio de pelo menos um critério psicoacústico em função dos parâmetros codificados do primeiro subconjunto.

25 Os parâmetros do segundo subconjunto podem ser reportar a bandas espectrais do sinal. Nesse caso, o processo compreende vantajosamente uma etapa de estimativa de um envoltório espectral do sinal codificado a partir dos parâmetros codificados do primeiro subconjunto e uma etapa de cálculo de uma curva de mascaramento freqüencial, aplicando um modelo de percepção auditiva com envoltório espectral estimado, e o critério psicoacústico faz referência ao nível do envoltório espectral estimado em relação à curva de mascaramento em cada banda espectral.

30

Em um modo de aplicação, ordenam-se os bits de codificação na seqüência de saída de tal modo que os N_0 bits de codificação do primeiro subconjunto precedem os $N - N_0$ bits de codificação dos parâmetros selecionados do segundo subconjunto e que os bits de codificação respectivos dos parâmetros selecionados do segundo subconjunto aí aparecem na ordem determinada para esses bits de codificação. Isto permite, no caso de uma seqüência binária ser truncada, receber a parte mais importante.

O número N pode variar de uma trama a outra, notadamente em função, por exemplo, da capacidade disponível da fonte de transmissão.

10 A codificação áudio com multitaxas, segundo a presente invenção poderá ser utilizada segundo um modo comutável ou hierárquico muito flexível, já que um número qualquer de bits a transmitir escolhido livremente entre N_0 e $N_{\max}$ pode ser selecionado a qualquer momento, isto é, trama por trama.

15 A codificação dos parâmetros do primeiro subconjunto pode ser com taxa variável, o que faz variar o número N_0 de uma trama a outra. Isto permite ajustar ao máximo a repartição dos bits em função das tramas a codificar.

Em um modo de aplicação, o primeiro subconjunto compreende parâmetros calculados por um núcleo codificador. Vantajosamente, o núcleo codificador tem uma banda de freqüências de funcionamento inferior à banda passante do sinal a codificar, e o primeiro subconjunto compreende, além disso, níveis energéticos do sinal áudio associados a faixas de freqüências superiores à faixa de funcionamento do núcleo codificador. Esse tipo estrutura é aquele de um codificador hierárquico com dois níveis, que libera, por exemplo, via o núcleo codificador, um sinal; codificado de uma qualidade julgado suficiente e que, em função da posição disponível, completa a codificação feita pelo núcleo codificador por informações suplementares provenientes do processo decodificação, de acordo com a invenção.

30 De preferência, ordenam-se então os bits de codificação do primeiro subconjunto na seqüência de saída de tal modo que os bits de codificação dos parâmetros calculados pelo núcleo codificador sejam imediata-

mente seguidos pelos bits de codificação dos níveis energéticos associados às faixas de frequências superiores. Isto assegura uma mesma banda passante às tramas sucessivamente codificadas, desde quando o decodificador recebe suficientemente de bits para dispor das informações do núcleo codificador e níveis energéticos codificados associados às faixas de frequências superiores.

Em um modo de utilização, estima-se um sinal de diferença entre o sinal a codificar e um sinal de síntese derivado dos parâmetros codificados produzidos pelo núcleo codificador, e o primeiro subconjunto compreende, além disso, níveis energéticos do sinal de diferença associados a faixas de frequências incluídas na faixa de funcionamento do núcleo codificador.

Um segundo aspecto da invenção se reporta a um processo de decodificação de uma seqüência binária de entrada para sintetizar um sinal áudio numérico correspondente à decodificação de uma trama codificada segundo o processo de codificação da invenção. Segundo esse processo, um número máximo N_{max} de bits de codificação é definido para um conjunto de parâmetros de descrição de uma trama de sinal, composto de um primeiro e de um segundo subconjunto. A seqüência de entrada compreende, para uma trama de sinal, um número N' de bits de codificação do conjunto de parâmetros, com $N' \leq N_{max}$. O processo de decomposição, de acordo com a invenção, compreende as seguintes etapas:

- extrai-se, desses N' bits da seqüência de entrada, um número N_0 de bits de codificação dos parâmetros do primeiro subconjunto, se $N_0 < N'$; se recupera os parâmetros do primeiro subconjunto sobre a base desses N_0 bits de codificação extraídos;
- determina-se uma alocação de $N_{max} - N_0$ bits de codificação para os parâmetros do segundo subconjunto; e
- classificam-se os $N_{max} - N_0$ bits de codificação alocados nos parâmetros do segundo subconjunto em uma ordem determinada.

A alocação e/ou a ordem de classificação dos $N_{max} - N_0$ bits de codificação são determinadas em função dos parâmetros recuperados do

primeiro subconjunto. O processo de decodificação compreende, além disso, as seguintes etapas:

- selecionam-se os parâmetros do segundo subconjunto nos quais são alocados os N' - N_0 bits de codificação classificados os primeiros
5 nessa ordem;
- extraem-se, desses N' bits da sequência de entrada, N' - N_0 bits de codificação dos parâmetros selecionados do segundo subconjunto;
- recuperam-se os parâmetros selecionados do segundo subconjunto sobre a base desses N' - N_0 bits de codificação extraídos; e
10 - sintetiza-se a trama de sinal, utilizando-se os parâmetros recuperados dos primeiro e segundo subconjuntos.

Esse processo de decodificação é vantajosamente associado a métodos de regeneração dos parâmetros que faltam devido à truncatura da sequência de $N_{\max}$ bits produzida, virtualmente ou não, pelo codificador.

15 Um terceiro aspecto da invenção se reporta a um codificador áudio, compreendendo meios de tratamento de sinal numérico ajustados para aplicar um processo de codificação, de acordo com a invenção.

Um outro aspecto da invenção se reporta a um decodificador áudio, compreendendo meios de tratamento de sinal numérico ajustados
20 para aplicar um processo de decodificação, de acordo com a invenção.

Outras particularidades e vantagens da presente invenção aparecerão na descrição dada a seguir de exemplos de realização não limitativos, com referência aos desenhos anexados, nos quais:

- a figura 1 representa um esquema sinóptico de um exemplo de
25 codificador áudio, de acordo com a invenção;
- a figura 2 representa uma sequência binária de saída de N bits em um modo de realização da invenção; e
- a figura 3 representa um esquema sinóptico de um decodificador áudio, de acordo com a invenção.

30 O codificador representado na figura 1 tem uma estrutura hierárquica com dois estágios de codificação. Um primeiro estágio de codificação 1 consiste, por exemplo, em um núcleo codificador em banda telefônica

(300-3400 Hz) de tipo CELP. Esse codificador é, no exemplo, considerado um codificador G.723.1 normalizado pelo ITU-T ("International Telecommunication Union") em modo fixo com 6,4 kbit/s. Ele calcula parâmetros G.723.1 de acordo com a norma e os quantifica por meio de 192 bits de codificação

5 P1 por trama de 30 ms.

O segundo estágio de codificação 2, permitindo aumentar a banda passante em direção à banda larga (50-7000 Hz), opera sobre o resíduo de codificação E do primeiro estágio, fornecido por um subtraidor 3 no esquema da figura 1. Um módulo de sincronização de sinais 4 retarda a tra-

10 ma de sinal áudio S do tempo gasto pelo tratamento do núcleo codificador 1. Sua saída é endereçada ao subtraidor 3 que lhe subtrai o sinal sintético S' igual à saída do núcleo decodificador operando sobre a base dos parâmetros quantificados, tais como representados pelos bits de saída P1 do núcleo codificador. Assim como é usual, o codificador 1 incorpora um decodificador

15 local que fornece S'.

O sinal áudio a codificar S tem, por exemplo, uma banda passante de 7 kHz, sendo mostrado em 16 kHz. Uma trama consiste, por exemplo, em 960 amostras, seja 60 ms de sinal ou duas tramas elementares do núcleo codificador G.723.1. Conforme esse último opere sobre sinais mos-

20 trados em 8 kHz, o sinal S é subamostrado em um fator 2 à entrada do núcleo codificador 1. Da mesma forma, o sinal sintético S' é supermostrado em 16 kHz na saída do núcleo codificador 1.

A taxa do primeiro estágio 1 é de 6,4 kbits/s ($2 \times N1 = 2 \times 192 = 384$ bits por trama). Se o codificador tiver uma taxa máxima de 32 kbits/s ($N_{max} = 1920$ bits por trama), a taxa máxima do segundo estágio é de 25,6 kbits/s ($1920 - 384 = 1536$ bits por trama). O segundo estágio 2 funciona, por exemplo, sobre tramas elementares, ou subtramas, de 20 ms (320 amostras a 16 kHz).

O segundo estágio 2 compreende um módulo 5 de transformação

30 ção tempo-freqüência, por exemplo de tipo MDCT ("*Modified Discrete Cosine Transform*") ao qual é endereçado o resíduo E obtido pelo subtraidor 3. Na prática, o funcionamento dos módulos 3 e 5 representados na figura 1

pode ser realizado, efetuando-se as seguintes operações para cada subtrama de 20 ms:

- transformação MDCT do sinal de entrada S retardada pelo módulo 4, que fornece 320 coeficientes MDCT. O espectro sendo limitado em 7225 Hz, só os 289 primeiros coeficientes MDCT são diferentes de 0;
- transformação MDCT do sinal de sintético S' . Como se trata do espectro do sinal de banda telefônica, só os 139 primeiros coeficientes MDCT são diferentes de 0 (até 3450 Hz); e
- cálculo do espectro de diferença entre os espectros precedentes.

O espectro resultante é distribuído em várias bandas de larguras diferentes por um módulo 6. A título de exemplo, a faixa passante do codec G.723.1 pode ser subdividida em 21 bandas, enquanto que as frequências mais elevadas são repartidas em 11 bandas suplementares. Nessas 11 bandas suplementares, o resíduo E é idêntico ao sinal de entrada S .

Um módulo 7 efetua a codificação do envoltório espectral do resíduo E . Ele começa por calcular a energia dos coeficientes MDCT de cada banda do espectro de diferença. Essas energias são a seguir denominadas "fatores de escala". Os 32 fatores de escala constituem o envoltório espectral do sinal de diferença. O módulo 7 procede então a sua quantificação em duas partes. A primeira parte corresponde à banda telefônica (21 primeiras bandas, de 0 a 3450 Hz), a segunda às bandas altas (11 últimas bandas, de 3450 a 7225 Hz). Em cada parte, o primeiro fator de escala é quantificado em absoluto, e os seguintes em diferencial, utilizando uma codificação clássica de Huffman com taxa variável. Esses 32 fatores de escala são quantificados sobre um número variável $N2(i)$ de bits $P2$ para cada subtrama de fileira i ($i = 1, 2, 3$).

Os fatores de escalas quantificadas são anotados como FQ na figura 1. Os bits de quantificação $P1$, $P2$ do primeiro subconjunto constituído dos parâmetros quantificados do núcleo codificador 1 e dos fatores de escala quantificados FQ são em um número variável $N0 = (2 \times N1) + N2(1) + N2(2) + N2(3)$. A diferença $N_{max} - N0 = 1536 - N2(1) - N2(2) - N2(3)$ é dis-

ponível para quantificar mais finamente os espectros das bandas.

Um módulo 8 normaliza os coeficientes MDCT repartidos em bandas pelo módulo 6, dividindo-os pelos fatores de escala quantificados FQ respectivamente determinados para essas bandas. Os espectros assim normalizados são fornecidos ao módulo de quantificação 9 que utiliza um esquema de quantificação vetorial de tipo conhecido. Os bits de quantificação provenientes do módulo 9 são anotados P3 na figura 1.

Um multiplexador de saída 10 reúne os bits P1, P2 e P3 oriundos dos módulos 1, 7 e 9 para formar a sequência binária ϕ de saída do codificador.

De acordo com a invenção, o número total de bits N da sequência de saída, representando uma trama comum não é necessariamente igual a Nmax. Ele pode lhe ser inferior. Todavia, a alocação dos bits de quantificação nas bandas é feita, baseando-se no número Nmax.

No esquema da figura 1, essa alocação é feita para cada subtrama pelo módulo 12 a partir do número Nmax – N0, fatores de escala quantificados FQ e de uma curva de ocultação espectral calculada por um módulo 11.

O funcionamento deste módulo 11 é o seguinte. Ele determina inicialmente um valor aproximado do envoltório espectral original do sinal S a partir daquele do sinal de diferença, tal como quantificada pelo módulo 7, e daquele que ele determina com a mesma resolução para o sinal sintético S' resultante do núcleo codificador. Estes dois envoltórios que são também determináveis por um decodificador só dispõem dos parâmetros do primeiro subconjunto pré-citado. Assim, o envoltório espectral estimado do sinal S será também disponível no decodificador. Em seguida, o módulo 11 calcula uma curva de ocultação espectral, aplicando, de forma conhecida em si, um modelo de percepção auditiva banda por banda com envoltório espectral original estimada. Essa curva 11 dá um nível de ocultação para cada banda considerada.

O módulo 12 realiza uma alocação dinâmica dos Nmax – N0 bits restantes da sequência ϕ dentre as 3 x 32 bandas das três transformações

MDCT do sinal de diferença. Na aplicação da invenção no caso exposta, em função de um critério de importância perceptual psicoacústica, fazendo referência ao nível do envoltório espectral estimada em relação à curva de ocultação em cada banda, aloca-se em cada banda uma taxa proporcional a esse nível. Outros critérios de classificação seriam utilizáveis.

Na sequência dessa alocação de bits, o módulo 9 sabe quantos bits devem ser considerados para a quantificação de cada banda em cada subtrama.

Todavia, se $N < N_{\max}$, esses bits alocados não serão necessariamente todos utilizados. Uma ordenação dos bits que representam as bandas é feita por um módulo 13, em função de um critério de importância perceptual. O módulo 13 classifica as 3×32 bandas em uma ordem de importância decrescente que pode ser a ordem decrescente das relações sinal com ocultação (relação entre o envoltório espectral estimado e a curva de ocultação em cada banda). Essa ordem é utilizada para a construção da sequência binária ϕ , de acordo com a invenção.

Em função do número N de bits desejado na sequência ϕ para a codificação da trama corrente, determinam-se as bandas que devem ser quantificadas pelo módulo 9, selecionando as bandas classificadas as primeiras pelo módulo 13 e retendo-se para cada banda selecionada um número de bits, tais como determinado pelo módulo 12.

Depois os coeficientes MDCT de cada banda selecionada são quantificados pelo módulo 9, por exemplo com o auxílio de um quantificador vetorial, de acordo com um número de bits alocado, para produzir um número de bits total igual a $N - N_0$.

O multiplexador de saída 10 constitui a sequência binária ϕ constituída dos N primeiros bits da sequência ordenada apresentada a seguir representada na figura 2 (caso $N = N_{\max}$):

a) inicialmente os trens binários correspondentes às duas tramas G.723.1 (384 bits);

b) depois os bits $F_{22}^{(i)}, \dots, F_{32}^{(i)}$

de quantificação dos fatores de escala, para as três subtramas ($i = 1, 2, 3$),

da 22ª banda espectral (primeira banda além da banda telefônica) à 32ª banda (codificação de Huffman com taxa variável);

c) depois os bits $F_{22}^{(i)}, \dots, F_{32}^{(i)}$ de quantificação dos fatores de escala, para as três subtramas ($i = 1, 2, 3$), da primeira banda espectral à vigésima primeira banda (codificação de Huffman com taxa variável); e

d) enfim, os índices $M_{c1}, M_{c2} \dots M_{c96}$ de quantificação vetorial das 96 bandas por ordem de importância perceptual, da banda a mais importante à banda a menos importante, respeitando a ordem determinada pelo módulo 13.

O fato de colocar em primeiro lugar (a e b) os parâmetros G.723.1 e os fatores de escala das bandas altas permite conservar a mesma banda passante para o sinal restituível pelo decodificador independentemente da taxa efetiva além de um valor mínimo correspondente à recepção desses grupos a e b. Esse valor mínimo, suficiente para a codificação de Huffman dos $3 \times 11 = 33$ fatores de escala das bandas altas além da codificação G.723.1, é, por exemplo, de 8 kbits/s.

O processo de codificação acima permite uma decodificação da trama, caso o decodificador receba N' bits com $N_0 \leq N' \leq N$. Esse número N' será geralmente variável de uma trama à outra.

Um decodificador, de acordo com a invenção, correspondendo a esse exemplo, é ilustrado pela figura 3. Um demultiplexador 20 separa a sequência de bits recebidos ϕ' para daí extrair os bits de codificação P1 e P2. Os 384 bits P1 são fornecidos ao núcleo decodificador 21 de tipo G.723.1 para que este sintetize duas tramas do sinal de base S' em banda telefônica. Os bits P2 são decodificados, segundo o algoritmo de Huffman por um módulo 22 que recupera assim os fatores de escalas quantificados FQ para cada uma das 3 subtramas.

Um módulo 23 de cálculo da curva de ocultação, idêntico àquele 11 do codificador da figura 1, recebe o sinal de base S' e os fatores de escalas quantificados FQ e produz os níveis de ocultação espectral para cada uma das 96 bandas. A partir desses níveis de ocultação espectral para cada uma das 96 bandas. A partir desses níveis de ocultação, fatores de escalas

quantificados FQ e do conhecimento do número $N_{\max}$ (assim como aquela do número N_0 que se deduz da decodificação de Huffman dos bits P2 pelo módulo 22), um módulo 24 determina uma alocação de bits do mesmo modo que o módulo 12 da figura 1. Além disso, um módulo 25 procede à ordena-
 5 ção das bandas, segundo o mesmo critério de classificação que o módulo 13 descrito com referência à figura 1.

A partir das informações fornecidas pelos módulos 24 e 25, o módulo 26 extrai os bits P3 da seqüência de entrada ϕ' e sintetiza os coeficientes MDCT normalizados relativos às bandas representadas na seqüência
 10 ϕ'' . Se for o caso ($N' < N_{\max}$), os coeficientes MDCT normalizados relativos às bandas que faltam podem, além disso, ser sintetizados por interpolação ou extrapolação conforme descritos depois (módulo 27). Essas bandas que faltam podem ter sido eliminados pelo codificador devido a uma truncatura com $N < N_{\max}$, ou elas podem ter sido eliminadas no decorrer da transmis-
 15 são ($N' < N$).

Os coeficientes MDCT normalizados, sintetizados pelo módulo 26 e/ou o módulo 27, são multiplicados por seus fatores de escala quantifi-
 cados respectivos (multiplicador 28) antes de serem apresentados no mó-
 dulo 29 que efetua a transformação frequência-tempo inversa da transfor-
 20 mação MDCT operada pelo módulo 5 do codificador. O sinal temporal de correção que daí resulta é adicionado ao sinal sintético S' liberado pelo nú-
 cleo decodificador 21 (adicionador 30) para produzir o sinal áudio de saída S do decodificador.

Deve ser observado que o decodificador poderá sintetizar um
 25 sinal S , mesmo em casos em que ele não recebe os N_0 primeiros bits da seqüência.

Basta-lhe receber os $2 \times N_1$ bits correspondentes à parte a da enumeração acima, a decodificação estando então em um modo "degrada-
 do". Só esse modo degradado não utiliza a síntese MDCT para obter o sinal
 30 decodificado. Para assegurar a comutação sem ruptura entre esse modo e os outros modos, o decodificador faz três análises MDCT seguidas por três sínteses MDCT, permitindo publicar relatórios da transformação MDCT. O

sinal de saída contém um sinal de qualidade banda telefônica. Se os $2 \times N1$ primeiros bits não são mesmo recebidos, o decodificador considera a trama correspondente como oculta e pode utilizar um algoritmo conhecido de dissimulação das tramas ocultas.

5 Se o decodificador receber os $2 \times N1$ bits correspondentes à parte a mais bits da parte b (bandas altas dos três envoltórios espectrais), poderá começar a sintetizar um sinal em banda larga. Ele pode notadamente proceder conforme a seguir:

10 1) o módulo 22 recupera as partes dos três envoltórios espectrais recebidos;

 2) as bandas não recebidas têm seus fatores de escala temporariamente colocados em zero;

 3) as partes baixas dos envoltórios espectrais são calculadas a partir das análises MDCT feitas sobre o sinal obtido após a decodificação G.723.1 e o módulo 23 calcula as três curvas de ocultação sobre os envoltórios assim obtidos;

20 4) o envoltório espectral é corrigido para a regularização, evitando os orifícios devido às bandas não recebidas: os valores nulos na parte alta dos envoltórios espectrais FQ são, por exemplo, substituídos pelo centésimo do valor da curva de ocultação calculada anteriormente, de tal modo que permanecem inaudíveis. O espectro completo das bandas baixas e o envoltório espectral das bandas altas são conhecidos nesse estágio;

25 5) o módulo 27 gera então o espectro alto. A estrutura fina dessas bandas é gerada por reflexão da estrutura finas de sua proximidades conhecidas antes da ponderação pelos fatores de escala (multiplicadores 28). No caso de nenhum dos bits P3 ser recebido, a "proximidade conhecida" corresponderá ao espectro do sinal S' produzido pelo núcleo decodificador G.723.1. Sua reflexão pode consistir em recopiar o valor do espectro MDCT normalizado, com eventualmente uma atenuação de suas variações

30 6) após transformação MDCT inversa (29) e adição (30) do sinal de correção resultante no sinal de saída do núcleo decodificador, obtém-se o

sinal sintetizado em banda larga.

No caso de o decodificador receber também uma parte pelo menos do envoltório espectral baixo do sinal de diferença (parte c), ele pode ou não considerar essa informação para afinar o envoltório espectral na etapa

5 3.

Se o decodificador 10 receber suficientemente bits P3 para decodificar pelo menos os coeficientes MDCT da faixa a mais importante, classificada a primeira na parte d da sequência, então o módulo 26 recupera determinados coeficientes MDCT normalizados a partir da alocação e a ordenação indicados pelos módulos 24 e 25. Esses coeficientes MDCT não têm necessidade de serem interpolados conforme na etapa 5 acima. Para as outras bandas, o processo das etapas 1 a 6 é aplicável pelo módulo 27 do mesmo modo que anteriormente, o conhecimento dos coeficientes MDCT recebidos para certas bandas, permitindo uma interpolação mais confiável na etapa 5.

As bandas não recebidas podem variar de uma subtrama MDCT à seguinte. A proximidade conhecida de uma banda que falta pode corresponder à mesma banda em uma outra subtrama na qual ela não está ausente e/ou a uma ou várias bandas as mais próximas no domínio freqüencial no decorrer da mesma subtrama. É também possível regenerar um espectro MCDT que está ausente em uma banda para uma subtrama, fazendo uma soma ponderada de contribuições avaliadas a partir de várias bandas/subtramas da proximidade conhecida.

À medida que a taxa efetiva de N' bits por trama coloca arbitrariamente o último bit de uma trama determinada, o último parâmetro codificado transmitido pode, segundo os casos, ser transmitido completa ou parcialmente. Dois casos podem então se apresentar:

. ou a estrutura de codificação adotada permite explorar a informação parcial recebida (caso de quantificadores escalares, ou de quantificação vetorial com dicionários divididos);

. ou ela não permite e trata-se o parâmetro não inteiramente recebido como os outros parâmetros não recebidos. Nota-se que, para esse

último caso, caso a ordem dos bits varia a cada trama, o número de bits assim perdidos é variável e a seleção de N' bits produzirá em média, sobre o elemento das tramas decodificadas, uma qualidade melhor do que aquela que se obteria com um número de bits menor.

REIVINDICAÇÕES

1. Processo de codificação de uma trama de sinal áudio numérico (S) em uma seqüência binária de saída (ϕ), na qual um número máximo N_{max} de bits de codificação é definido para um conjunto de parâmetros calculáveis a partir da trama de sinal, composto de um primeiro e de um segundo subconjunto, o processo compreendendo as seguintes etapas:
 - calcular os parâmetros do primeiro subconjunto, e codificar esses parâmetros sobre um número N_0 de bits de codificação tal que $N_0 < N_{max}$;
 - 10 - determinar uma alocação de $N_{max} - N_0$ bits de codificação para os parâmetros do segundo subconjunto; e
 - classificar os $N_{max} - N_0$ bits de codificação alocados nos parâmetros do segundo subconjunto em uma ordem determinado, no qual a alocação e/ou a ordem de classificação dos $N_{max} - N_0$ bits de
 - 15 codificação é determinado em função dos parâmetros codificados do primeiro subconjunto, o processo compreendendo, além disso, as etapas seguintes em resposta à indicação de um número N de bits da seqüência binária de saída disponíveis para a codificação desse conjunto de parâmetros, com $N_0 < N \leq N_{max}$:
 - 20 . selecionar os parâmetros do segundo subconjunto aos quais são alocados os $N - N_0$ bits de codificação classificados os primeiros nessa ordem;
 - . calcular os parâmetros selecionados do segundo subconjunto, e codificar esses parâmetros para produzir esses $N - N_0$ bits de codificação
 - 25 classificados os primeiros; e
 - . inserir na seqüência de saída os N_0 bits de codificação do primeiro subconjunto, assim como os $N - N_0$ bits de decodificação dos parâmetros selecionados do segundo subconjunto.
2. Processo, de acordo com a reivindicação 1, no qual a ordem
- 30 de classificação dos bits de codificação alocados nos parâmetros do segundo subconjunto é variável de uma trama à outra.

3. Processo, de acordo com a reivindicação 1 ou 2, no qual $N <$

Nmax.

4. Processo, de acordo com qualquer uma das reivindicações precedentes, no qual a ordem de classificação dos bits de codificação alocados nos parâmetros do segundo subconjunto é uma ordem de importância decrescente determinada em função de pelo menos os parâmetros codificados do primeiro subconjunto.

5. Processo, de acordo com a reivindicação 4, no qual a ordem de classificação dos bits de codificação alocados nos parâmetros do segundo subconjunto é determinado com o auxílio de pelo menos um critério psicoacústico em função dos parâmetros codificados do primeiro subconjunto.

6. Processo, de acordo com a reivindicação 5, no qual os parâmetros do segundo subconjunto se reportam a bandas espectrais do sinal, no qual se estima um envoltório espectral do sinal codificado a partir dos parâmetros codificados do primeiro subconjunto, no qual se calcula uma curva de ocultação freqüencial aplicando um modelo de percepção auditiva ao envoltório espectral estimada, e no qual o critério psicoacústico faz referência ao nível do envoltório espectral estimado em relação à curva de ocultação em cada banda espectral.

7. Processo, de acordo com qualquer uma das reivindicações 4 a 6, no qual $N_{max} = N$.

8. Processo, de acordo com qualquer uma das reivindicações precedentes, no qual se ordenam os bits de codificação na seqüência de saída de tal modo que os N_0 bits de codificação do primeiro subconjunto precedem os $N - N_0$ bits de codificação dos parâmetros selecionados do segundo subconjunto e que os bits de codificação respectivos dos parâmetros selecionados do segundo subconjunto aí aparecem na ordem determinado para esses bits de codificação.

9. Processo, de acordo com qualquer uma das reivindicações precedentes, no qual o número N varia de uma trama à outra.

10. Processo, de acordo com qualquer uma das reivindicações precedentes, no qual a codificação dos parâmetros do primeiro subconjunto é de taxa variável, o que faz variar o número N_0 de uma trama à outra.

11. Processo, de acordo com qualquer uma das reivindicações precedentes, no qual o primeiro subconjunto compreende dos parâmetros calculados por um núcleo codificador (1).

12. Processo, de acordo com a reivindicação 11, no qual o núcleo codificador (1) tem uma banda de frequências de funcionamento inferior à banda passante do sinal a codificar, e no qual o primeiro subconjunto compreende, além disso, níveis energéticos do sinal áudio associados a bandas de frequências superiores à banda de funcionamento do núcleo codificador.

13. Processo, de acordo com cada uma das reivindicações 8 e 12, no qual se ordenam os bits de codificação do primeiro subconjunto na seqüência de saída de tal modo que os bits de codificação dos parâmetros calculados pelo núcleo codificador sejam imediatamente seguidos pelos bits de codificação dos níveis energéticos associados às bandas de frequências superiores.

14. Processo, de acordo com qualquer uma das reivindicações 11 a 13, no qual se estima um sinal de diferença entre o sinal a codificar e um sinal de síntese derivados dos parâmetros codificados produzidos pelo núcleo codificador, e no qual o primeiro subconjunto compreende, além disso, níveis energéticos do sinal de diferença associados a bandas de frequências incluídas na banda de funcionamento do núcleo codificador.

15. Processo, de acordo com a reivindicação 8 e de acordo com qualquer uma das reivindicações 12 a 14, no qual se ordena os bits de codificação do primeiro subconjunto na seqüência de saída de tal modo que os bits de codificação dos parâmetros calculados pelo núcleo codificador (1) sejam seguidos pelos bits de codificação dos níveis energéticos associados às bandas de frequências.

16. Processo de decodificação de uma seqüência binária de entrada (ϕ') para sintetizar um sinal áudio numérico (S), no qual um número máximo N_{max} de bits de codificação é definido para um conjunto de parâmetros de descrição de uma trama de sinal, composto de um primeiro e de um segundo subconjuntos, a seqüência de entrada compreendendo, para

uma trama de sinal, um número N' de bits de codificação desse conjunto de parâmetros, com $N' \leq N_{\max}$, o processo compreendendo as seguintes etapas:

- 5 - extrair desses N' bits da sequência de entrada um número N_0 de bits de codificação dos parâmetros do primeiro subconjunto, caso $N_0 < N'$;
- recuperar os parâmetros do primeiro subconjunto sobre a base desses N_0 bits de codificação extraídos;
- determinar uma alocação de $N_{\max} - N_0$ bits de codificação
- 10 para os parâmetros do segundo subconjunto; e
- classificar os $N_{\max} - N_0$ bits de codificação alocados nos parâmetros do segundo subconjunto em uma ordem determinada,
- no qual a alocação e/ou a ordem de classificação dos $N_{\max} - N_0$ bits de codificação é determinado em função dos parâmetros recuperados do primeiro subconjunto, o processo compreendendo, além disso, as
- 15 seguintes etapas:
- selecionar os parâmetros do segundo subconjunto nos quais são alocados os $N' - N_0$ bits de codificação classificados os primeiros nessa ordem;
- 20 - extrair, dessas N' bits da sequência de entrada, $N' - N_0$ bits de codificação dos parâmetros selecionados do segundo subconjunto;
- recuperar os parâmetros selecionados do segundo subconjunto sobre a base desses $N' - N_0$ bits de codificação extraídos; e
- sintetizar a trama de sinal, utilizando os parâmetros recuperados dos primeiros e segundo subconjuntos.
- 25 17. Processo, de acordo com a reivindicação 16, no qual a ordem de classificação dos bits de codificação alocados nos parâmetros do segundo subconjunto é variável de uma trama à outra.
18. Processo, de acordo com a reivindicação 16 ou 17, no qual
- 30 $N' < N_{\max}$.
19. Processo, de acordo com qualquer uma das reivindicações 16 a 18, no qual a ordem de classificação dos bits de codificação alocados

nos parâmetros do segundo subconjunto é uma ordem de importância decrescente determinado em função de pelo menos os parâmetros recuperados do primeiro subconjunto.

20. Processo, de acordo com a reivindicação 19, no qual a ordem de classificação dos bits de codificação alocados nos parâmetros do segundo subconjunto é determinado com o auxílio de pelo menos um critério psicoacústico em função dos parâmetros recuperados do primeiro subconjunto.

21. Processo, de acordo com a reivindicação 20, no qual os parâmetros do segundo subconjunto se reportam a bandas espectrais do sinal, no qual se estima um envoltório espectral do sinal, a partir dos parâmetros recuperados do primeiro subconjunto, no qual se calcula uma curva de ocultação freqüencial, aplicando-se um modelo de percepção auditiva no envoltório espectral estimado, e no qual o critério psicoacústico faz referência ao nível do envoltório espectral estimado em relação à curva de ocultação em cada banda espectral.

22. Processo, de acordo com qualquer uma das reivindicações 16 a 21, no qual os N_0 bits de codificação dos parâmetros do primeiro subconjunto são extraídos N' bits recebidos em posições da seqüência que precedem as posições de onde são extraído os $N' - N_0$ bits de codificação dos parâmetros selecionados do segundo subconjunto.

23. Processo, de acordo com qualquer uma das reivindicações 16 a 22, no qual, para sintetizar a trama de sinal, se estimam parâmetros não selecionados do segundo subconjunto por interpolação a partir de pelo menos dos parâmetros selecionados recuperados sobre a base desses $N' - N_0$ bits de codificação extraídos.

24. Processo, de acordo com qualquer uma das reivindicações 16 a 23, no qual o primeiro subconjunto compreende parâmetros de entrada de um núcleo decodificador (21).

25. Processo, de acordo com a reivindicação 24, no qual o núcleo decodificador (21) tem uma banda de freqüência de funcionamento inferior à banda passante do sinal a sintetizar, e no qual o primeiro subconjunto

compreende, além disso, níveis energéticos do sinal áudio associados a bandas de frequências superiores à banda de funcionamento do núcleo decodificador.

26. Processo, de acordo com cada uma das reivindicações 22 e 5 25, no qual os bits de codificação do primeiro subconjunto na sequência de entrada são ordenados de tal modo que os bits de codificação dos parâmetros de entrada do núcleo decodificador (21) sejam imediatamente seguidos por bits de codificação dos níveis energéticos associados às faixas de frequências superiores.

10 27. Processo, de acordo com a reivindicação 26, compreendendo as seguintes etapas, caso os N' bits da sequência de entrada (ϕ') se limitem aos bits de codificação dos parâmetros de entrada do núcleo decodificador (21) e a uma parte pelo menos dos bits de codificação dos níveis energéticos associados às faixas de frequências superiores:

15 - extrair da sequência de entrada os bits de codificação dos parâmetros de entrada do núcleo decodificador e essa parte dos bits de codificação dos níveis energéticos;

- sintetizar um sinal de base (S') no núcleo decodificador e recuperar níveis energéticos associados às faixas de frequências superiores sobre a base dos bits de codificação extraídos;

20 - calcular um espectro do sinal de base;

- afetar um nível energético a cada banda superior à qual é associados um nível energético não codificado na sequência de entrada;

25 - sintetizar componentes espectrais para cada faixa de frequências superior a partir do nível energético correspondente e do espectro do sinal de base em pelo menos uma banda desse espectro;

- aplicar uma transformação para o domínio temporal nas componentes espectrais sintetizadas para se obter um sinal de correção do sinal de base; e

30 - adicionar o sinal de base e o sinal de correção para sintetizar a trama de sinal.

28. Processo, de acordo com a reivindicação 27, no qual o nível

energético afetado em uma banda superior à qual é associado um nível energético não codificado na sequência de entrada é uma fração de um nível de ocultação perceptual calculado a partir do espectro do sinal de base e os níveis energéticos recuperados sobre a base dos bits de codificação extraí-

5 dos.

29. Processo, de acordo com qualquer uma das reivindicações 24 a 28, no qual se sintetiza um sinal de base (S') no núcleo decodificador, e no qual o primeiro subconjunto compreende, além disso, níveis energéticos de um sinal de diferença entre o sinal a sintetizar e o sinal de base associa-

10 dos a faixas de frequências inclusas na banda de funcionamento do núcleo codificador.

30. Processo, de acordo com qualquer uma das reivindicações 25, 26 e 29, no qual, para $N_0 < N' < N_{\max}$, se estimam os parâmetros não selecionados do segundo subconjunto relativos a componentes espectrais em faixas de frequências com o auxílio de um espectro calculado do sinal de

15 base e/ou dos parâmetros selecionados recuperados sobre a base desses $N' - N_0$ bits de codificação extraídos.

31. Processo,, de acordo com a reivindicação 30, no qual os parâmetros não selecionados do segundo subconjunto em uma faixa de fre-

20 quências são estimados com o auxílio de uma proximidade espectral dessa banda, determinado sobre a base dos N' bits de codificação da sequência de entrada.

32. Processo, de acordo com a reivindicação 22 e de acordo com qualquer uma das reivindicações 25 a 31, no qual os bits de codificação dos parâmetros de entrada do núcleo decodificador (21) são extraídos dos

25 N' bits recebidos em posições da sequência que precedem as posições de onde são extraídos os bits de codificação dos níveis energéticos associados às faixas de frequências.

33. Processo, de acordo com qualquer uma das reivindicações 16 a 32, no qual o número N' varia de uma trama à outra.

30

34. Processo, de acordo com qualquer uma das reivindicações 16 a 33, no qual o número N_0 varia de uma trama à outra.

35. Codificador áudio, compreendendo meios de tratamento de sinal numérico ajustados para aplicar um processo decodificação, como definido em qualquer uma das reivindicações 1 a 15.

5 36. Decodificador áudio, compreendendo meios de tratamento de sinal numérico ajustados para aplicar um processo de codificação, como definido em qualquer uma das reivindicações 16 a 34.

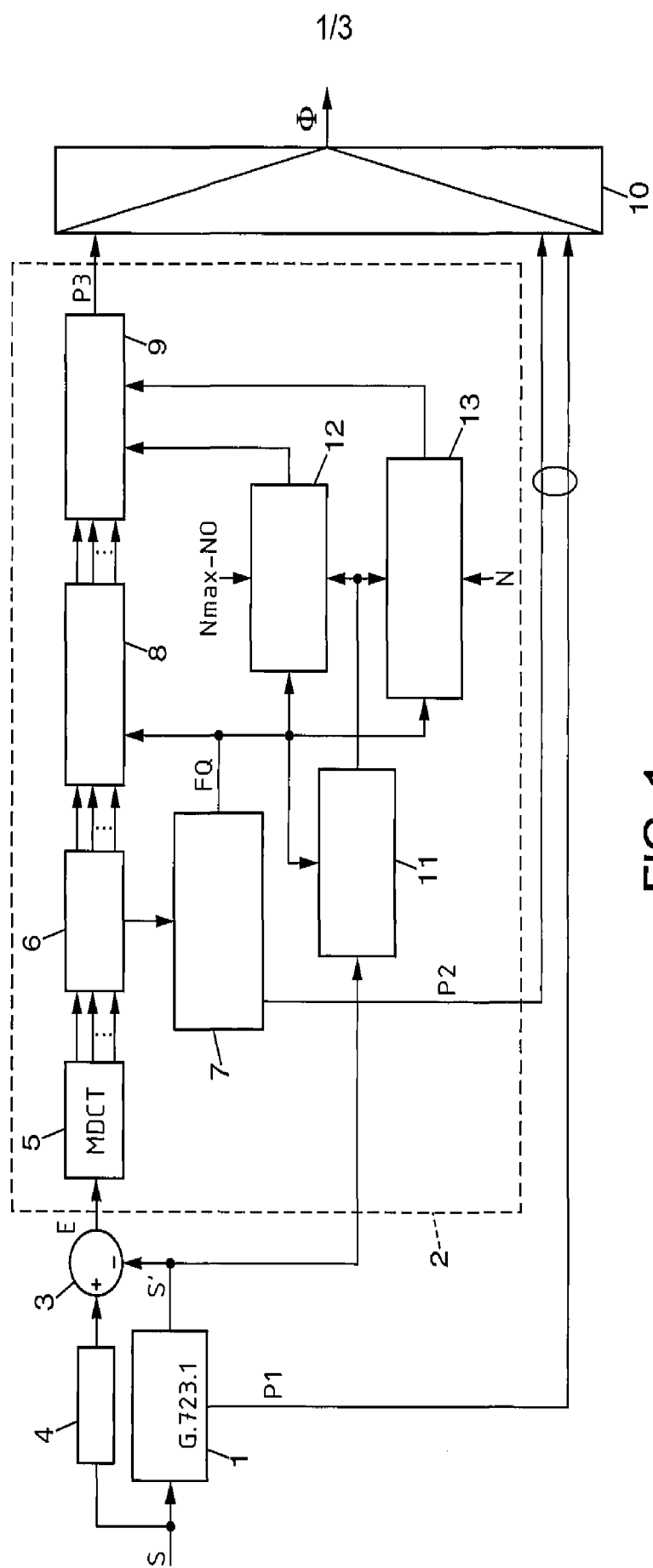


FIG. 1

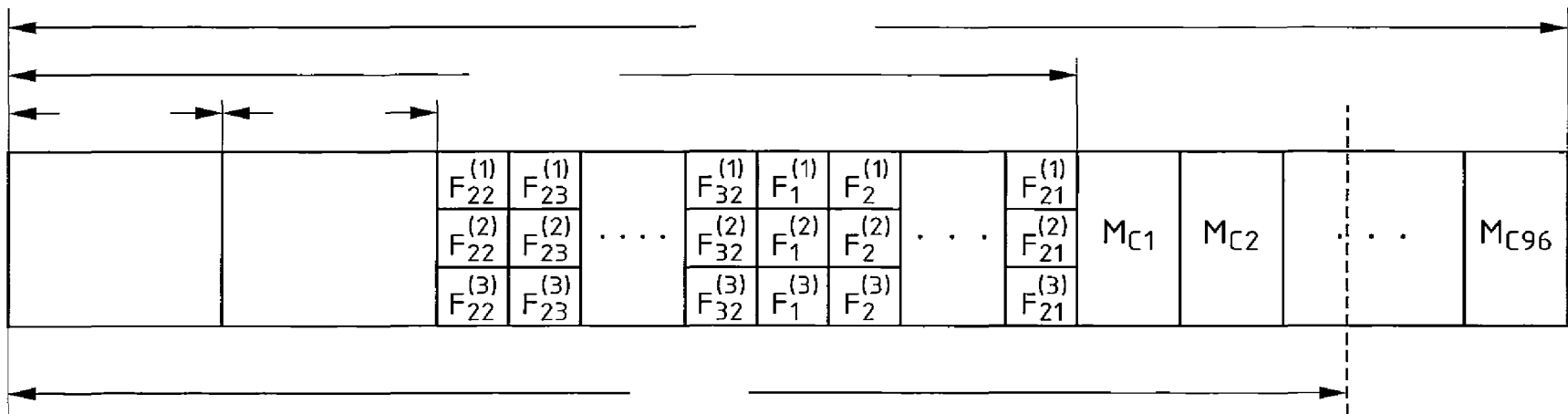


FIG. 2

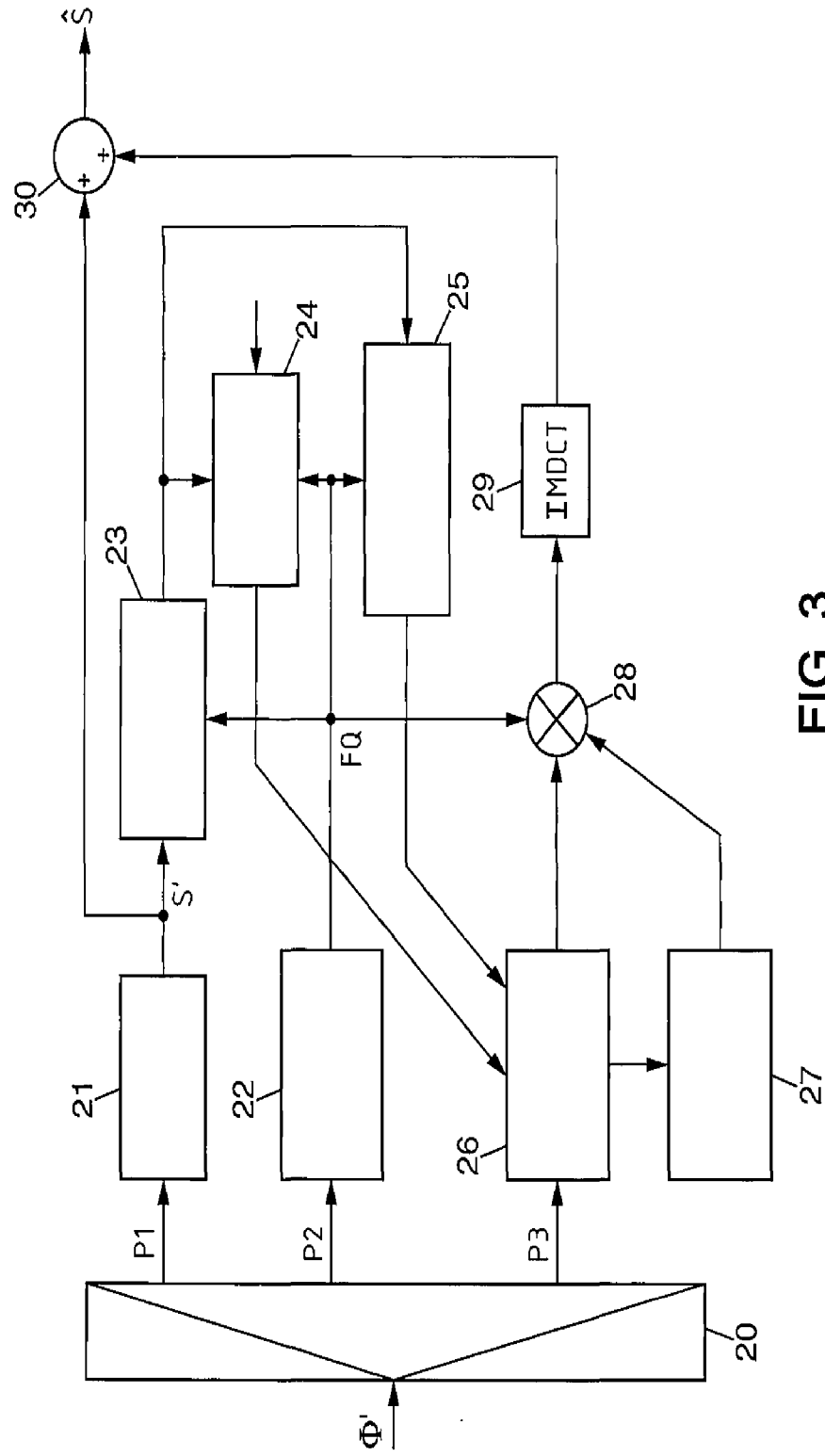


FIG. 3