



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

**(11) PI 0822236-3 B1**



**(22) Data do Depósito: 30/12/2008**

**(45) Data de Concessão: 04/02/2020**

---

**(54) Título:** SISTEMA DE CODIFICAÇÃO DE ÁUDIO, DECODIFICADOR DE ÁUDIO, MÉTODO DE CODIFICAÇÃO DE ÁUDIO E MÉTODO DE DECODIFICAÇÃO DE ÁUDIO

**(51) Int.Cl.:** G10L 19/00.

**(30) Prioridade Unionista:** 04/01/2008 SE 0800032-5; 24/05/2008 US 61/055,978; 24/05/2008 EP 08 009530.0.

**(73) Titular(es):** DOLBY INTERNATIONAL AB.

**(72) Inventor(es):** PER HENRIK HEDELIN; PONTUS JAN CARLSSON; JONAS LEIF SAMUELSSON; MICHAEL SCHUG.

**(86) Pedido PCT:** PCT EP2008011144 de 30/12/2008

**(87) Publicação PCT:** WO 2009/086918 de 16/07/2009

**(85) Data do Início da Fase Nacional:** 05/07/2010

**(57) Resumo:** SISTEMA DE CODIFICAÇÃO DE ÁUDIO, DECODIFICADOR DE ÁUDIO, MÉTODO DE CODIFICAÇÃO DE ÁUDIO E MÉTODO DE DECODIFICAÇÃO DE ÁUDIO A presente invenção refere-se a um sistema de codificação de áudio inédito que pode codificar bem sinais tanto de áudio geral quanto de fala em baixas taxas de bit. Um sistema de codificação de áudio proposto compreende a unidade de predição linear para filtrar um sinal de entrada com base em um filtro adaptativo; uma unidade de transformação, para transformar um quadro do sinal de entrada filtrado em um domínio transformado; e uma unidade de quantização para quantizar o sinal do domínio transformado. A unidade de quantização decide, com base nas características do sinal inserido, codificar o sinal do domínio transformado com um quantizador baseado em modelo ou um quantizador não baseado em modelo. Preferivelmente, a decisão é baseada na duração do quadro aplicada pela unidade de transformação.

Relatório Descritivo da Patente de Invenção para "**SISTEMA DE CODIFICAÇÃO DE ÁUDIO, DECODIFICADOR DE ÁUDIO, MÉTODO DE CODIFICAÇÃO DE ÁUDIO E MÉTODO DE DECODIFICAÇÃO DE ÁUDIO**".

Campo Técnico

[0001] A presente invenção refere-se à codificação de sinais de áudio e, em particular, à codificação de qualquer sinal de áudio, não limitado nem à fala, nem à música nem a uma combinação das mesmas.

Antecedentes da Invenção

[0002] Na tecnologia anterior, há codificadores de fala especificamente desenhados para codificar sinais de fala baseando a codificação em um modelo fonte do sinal, isto é, o sistema vocal humano. Estes codificadores não podem tratar sinais de áudio arbitrários, tal como música, ou qualquer outro sinal que não de fala. Adicionalmente, na tecnologia anterior, há codificadores de música, comumente referidos como codificadores de áudio, que baseiam sua codificação em compreensões sobre o sistema auditivo humano, e não no modelo fonte do sinal. Estes codificadores podem tratar muito bem sinais arbitrários, embora em baixas taxas de bit para sinais de fala, e o codificador de fala dedicado proporciona uma qualidade de áudio superior. Portanto, até o momento, não existe estrutura de codificação geral para codificação de sinais de áudio arbitrários que também realiza um codificador de fala para fala e, também, um codificador de música para música, quando operado em baixas taxas de bit.

[0003] Assim, há uma necessidade de melhores codificador e decodificador de áudio com maior qualidade de áudio e/ou menores taxas de bit.

Sumário da Invenção

[0004] A presente invenção refere-se à eficiente codificação de

sinais de áudio arbitrários em um nível de qualidade igual ou melhor que aquele de um sistema especificamente adaptado a um sinal específico.

[0005] A presente invenção é direcionada a algoritmos do codec de áudio que contêm tanto uma codificação por predição linear (LPC) quanto uma parte codificadora por transformada que opera em um sinal LPC processado.

[0006] A presente invenção refere-se adicionalmente a uma estratégia de quantização que depende de uma duração do quadro transformado. Além do mais, é proposto um quantizador com restrição de entropia com base em modelo que emprega codificação aritmética. Além do mais, é provida a inserção de deslocamentos aleatórios em um quantizador escalar uniforme. A invenção sugere adicionalmente um quantizador baseado em modelo, por exemplo, um Quantizador com Restrição de Entropia (ECQ), que emprega codificação aritmética.

[0007] A presente invenção refere-se adicionalmente à eficiente codificação de fatores de escala na parte de codificação transformada de um codificador de áudio pela exploração da presença de dados de LPC.

[0008] A presente invenção refere-se adicionalmente ao eficiente uso de um reservatório de bits em um codificador de áudio com um tamanho de quadro variável.

[0009] A presente invenção refere-se adicionalmente a um codificador para codificação de sinais de áudio e geração de um fluxo de bits, e a um decodificador, para decodificação do fluxo de bits e geração de um sinal de áudio reconstruído que é perceptivamente indistinguível do sinal de áudio de entrada.

[00010] Um primeiro aspecto da presente invenção refere-se à quantização em um codificador por transformada que, por exemplo, aplica uma Transformada de Cosseno Discreta Modificada (MDCT).

Preferivelmente, o quantizador proposto quantiza linhas da MDCT. Este aspecto é aplicável independentemente se o codificador usa adicionalmente uma análise da codificação por predição linear (LPC) ou predição de longo prazo adicional (LTP).

[00011] A presente invenção provê um sistema de codificação de áudio que compreende uma unidade de predição linear, para filtrar um sinal de entrada com base em um filtro adaptativo; uma unidade de transformação, para transformar um quadro do sinal de entrada filtrado em um domínio transformado; e uma unidade de quantização para quantizar o sinal do domínio transformado. A unidade de quantização decide, com base nas características do sinal de entrada, codificar o sinal do domínio transformado com um quantizador baseado em modelo ou um quantizador não baseado em modelo. Preferivelmente, a decisão é baseada na duração do quadro aplicado pela unidade de transformação. Entretanto, outros critérios dependentes do sinal de entrada para comutar a estratégia de quantização também são concebidos e estão no escopo do presente pedido.

[00012] Um outro importante aspecto da invenção é que o quantizador pode ser adaptativo. Em particular, o modelo no quantizador baseado em modelo pode ser adaptativo para se ajustar ao sinal de áudio de entrada. O modelo pode variar ao longo do tempo, por exemplo, dependendo das características do sinal de entrada. Isto permite menor distorção da quantização e, assim, maior qualidade de codificação.

[00013] De acordo com uma modalidade, a estratégia de quantização proposta é condicionada à duração do quadro. Sugere-se que a unidade de quantização possa decidir, com base na duração do quadro aplicado pela unidade de transformação, codificar o sinal do domínio transformado com um quantizador baseado em modelo ou um quantizador não baseado em modelo. Preferivelmente, a unidade de quantização é configurada para codificar um sinal do domínio trans-

formado, para um quadro com um tamanho de quadro menor que um valor limite, por meio de uma quantização com restrição de entropia com base em modelo. A quantização com base em modelo pode ser condicionada a parâmetros variados. Grandes quadros podem ser quantizados, por exemplo, por um quantizador escalar, por exemplo, com codificação por entropia Huffman, como é usado, por exemplo, no codec AAC.

[00014] O sistema de codificação de áudio pode compreender adicionalmente uma unidade de predição de longo prazo (LTP) para estimar o quadro do sinal de entrada filtrado com base em uma reconstrução de um segmento prévio do sinal de entrada filtrado, e uma unidade de combinação do sinal do domínio transformado, para combinar, no domínio transformado, a estimativa da predição de longo prazo e o sinal de entrada transformado, para gerar o sinal do domínio transformado que é inserido na unidade de quantização.

[00015] A comutação entre diferentes métodos de quantização das linhas da MDCT é um outro aspecto de uma modalidade preferida da invenção. Pelo emprego de diferentes estratégias de quantização para diferentes tamanhos de transformada, o codec pode fazer toda a quantização e codificação no domínio MDCT, sem a necessidade de ter um codificador de fala do domínio temporal específico executando em paralelo ou em série em relação ao codec do domínio transformado. A presente invenção preceitua que, para sinais tipo fala, quando houver um ganho de LTP, preferivelmente, o sinal é codificado usando uma transformada curta e um quantizador baseado em modelo. O quantizador baseado em modelo é particularmente adequado para a transformada curta, e proporciona, como será esboçado posteriormente, as vantagens de um quantizador vetorial (VQ) específico de fala do domínio temporal, ainda sendo operado no domínio MDCT, e sem nenhuma exigência de que o sinal de entrada seja um sinal de fala. Em outras

palavras, quando o quantizador baseado em modelo for usado para os segmentos de transformada curta em combinação com a LTP, a eficiência do VQ do codificador de fala do domínio temporal dedicado é retida, sem perda de generalidade e sem deixar o domínio MDCT.

[00016] Além do mais, para sinais de música mais estacionários, é preferido usar uma transformada de tamanho relativamente grande, que é comumente usada em codecs de áudio, e um esquema de quantização que pode tirar vantagem das linhas espectrais esparsas discriminadas pela grande transformada. Portanto, a presente invenção preceitua o uso deste tipo de esquema de quantização para grandes transformadas.

[00017] Assim, a comutação de estratégia de quantização em função de duração do quadro habilita o codec a reter tanto as propriedades de um codec de fala dedicado quanto às propriedades de um codec de áudio dedicado, simplesmente pela escolha do tamanho da transformada. Isto evita todos os problemas dos sistemas da tecnologia anterior, que tentam tratar sinais de fala e de áudio igualmente bem em baixas taxas, uma vez que estes sistemas, inevitavelmente, deparam com os problemas e dificuldades de combinar eficientemente a codificação do domínio temporal (o codificador de fala) com codificação do domínio de frequência (o codificador de áudio).

[00018] De acordo com um outro aspecto da invenção, a quantização usa tamanhos de etapa adaptativos. Preferivelmente, o(s) tamanho(s) da etapa de quantização para componentes do sinal do domínio transformado é(são) adaptado(s) com base em parâmetros da predição linear e/ou da predição de longo prazo. O(s) tamanho(s) da etapa de quantização pode(m) ser adicionalmente configurado(s) para depender da frequência. Em modalidades da invenção, o tamanho da etapa de quantização é determinado com base em pelo menos um de: o polinômio do filtro adaptativo, um parâmetro de controle da taxa de

codificação, um valor de ganho da predição de longo prazo e uma variância do sinal de entrada.

[00019] Preferivelmente, a unidade de quantização compreende quantizadores escalares uniformes para quantizar os componentes do sinal do domínio transformado. Cada quantizador escalar está aplicando uma quantização uniforme, por exemplo, com base em um modelo de probabilidade, em uma linha da MDCT. O modelo de probabilidade pode ser um modelo Laplaciano ou Gaussiano, ou qualquer outro modelo de probabilidade que é adequado para características de sinal. A unidade de quantização pode inserir adicionalmente um deslocamento aleatório nos quantizadores escalares uniformes. A inserção do deslocamento aleatório provê vantagens da quantização vetorial aos quantizadores escalares uniformes. De acordo com uma modalidade, os deslocamentos aleatórios são determinados com base em uma otimização de uma distorção de quantização, preferivelmente, em um domínio perceptivo e/ou em consideração do custo em termos do número de bits exigidos para codificar os índices de quantização.

[00020] A unidade de quantização pode compreender adicionalmente um codificador aritmético para codificação dos índices de quantização gerados pelos quantizadores escalares uniformes. Isto alcança uma baixa taxa de bits, que se aproxima do mínimo possível dado pela entropia de sinal.

[00021] A unidade de quantização pode compreender adicionalmente um quantizador residual para quantizar um sinal de quantização residual resultante dos quantizadores escalares uniformes, a fim de reduzir adicionalmente a distorção total. Preferivelmente, o quantizador residual é um quantizador vetorial de taxa fixa.

[00022] Múltiplos pontos de reconstrução da quantização podem ser usados na unidade de dequantização do codificador e/ou do quantizador invertido no decodificador. Por exemplo, pontos de reconstru-

ção de mínimo erro quadrático médio (MMSE) e/ou de ponto central (ponto médio) podem ser usados para reconstruir um valor quantizado com base em seu índice de quantização. Um ponto de reconstrução da quantização pode ser adicionalmente baseado em uma interpolação dinâmica entre um ponto central e um ponto MMSE, possivelmente controlado por características dos dados. Isto permite controlar a inserção de ruído e evitar furos espectrais em função da atribuição de linhas da MDCT a uma posição de quantização zero para baixas taxas de bit.

[00023] Preferivelmente, uma ponderação perceptiva no domínio transformado é aplicada durante a determinação da distorção da quantização, a fim de colocar diferentes ponderações em componentes de frequência específicos. As ponderações perceptivas podem ser eficientemente derivadas dos parâmetros de predição linear.

[00024] Um outro aspecto independente da invenção refere-se ao conceito geral de se fazer uso da coexistência dos dados LPC e SCF (Fator de Escala). Em um codificador com base em transformada, por exemplo, que aplica uma Transformada de Cosseno Discreta Modificada (MDCT), fatores de escala podem ser usados na quantização para controlar o tamanho da etapa de quantização. Na tecnologia anterior, estes fatores de escala são estimados a partir do sinal original para determinar uma curva de mascaramento. Agora, sugere-se estimar um segundo conjunto de fatores de escala com a ajuda de um filtro perceptivo ou modelo psicoacústico que é calculado a partir dos dados de LPC. Isto permite uma redução do custo para transmitir/armazenar os fatores de escala pela transmissão/armazenamento apenas da diferença entre os fatores de escala realmente aplicados e os fatores de escala estimados por LPC, em vez da transmissão/armazenamento dos fatores de escala reais. Assim, em um sistema de codificação de áudio que contém elementos da codificação de

fala, tal como, por exemplo, uma LPC, e elementos da codificação transformada, tal como uma MDCT, a presente invenção reduz o custo para transmitir informação do fator de escala necessária para a parte de codificação transformada do codec, pela exploração dos dados providos pela LPC. Percebe-se que este aspecto é independente de outros aspectos do sistema de codificação de áudio proposto e também pode ser implementado em outros sistemas de codificação de áudio.

[00025] Por exemplo, uma curva de mascaramento perceptiva pode ser estimada com base nos parâmetros do filtro adaptativo. O segundo conjunto de fatores de escala com base em predição linear pode ser determinado com base na curva de mascaramento perceptiva estimada. Então, a informação armazenada/transmitida do fator de escala é determinada com base na diferença entre os fatores de escala realmente usados na quantização e os fatores de escala que são calculados a partir da curva de mascaramento perceptiva com base em LPC. Isto remove dinâmica e redundância da informação armazenada/transmitida, de forma que menos bits sejam necessários para armazenar/transmitir os fatores de escala.

[00026] No caso em que a LPC e a MDCT não operam na mesma taxa de quadros, isto é, têm diferentes tamanhos de quadro, os fatores de escala com base em predição linear para um quadro do sinal do domínio transformado podem ser estimados com base em parâmetros interpolados de predição linear, para corresponder à janela de tempo coberta pelo quadro MDCT.

[00027] Portanto, a presente invenção provê um sistema de codificação de áudio que é baseado em um codificador por transformada, e inclui módulos de predição e modelagem fundamentais de um codificador de fala. O sistema inventivo compreende uma unidade de predição linear para filtrar um sinal de entrada com base em um filtro adap-

tativo; uma unidade de transformação, para transformar um quadro do sinal de entrada filtrado em um domínio transformado; uma unidade de quantização, para quantizar um sinal do domínio transformado; uma unidade de determinação do fator de escala, para geração de fatores de escala com base em uma curva de limite de mascaramento, para uso na unidade de quantização durante a quantização do sinal do domínio transformado; uma unidade de estimação do fator de escala da predição linear, para estimar fatores de escala com base em predição linear com base nos parâmetros do filtro adaptativo; e um codificador do fator de escala para codificar a diferença entre os fatores de escala com base na curva de limite de mascaramento e os fatores de escala com base em predição linear. Pela codificação da diferença entre os fatores de escala aplicados e os fatores de escala que podem ser determinados no decodificador com base em informação de predição linear disponível, a eficiência de codificação e armazenamento pode ser aumentada, e apenas poucos bits precisam ser armazenados/transmitidos.

[00028] Um outro aspecto específico do codificador independente da invenção refere-se ao tratamento do reservatório de bits para tamanhos de quadro variáveis. Em um sistema de codificação de áudio que pode codificar quadros de duração variável, o reservatório de bits é controlado pela distribuição dos bits disponíveis entre os quadros. Dado uma medida de dificuldade razoável para os quadros individuais e um reservatório de bits de um tamanho definido, um certo desvio de uma taxa de bits constante exigida permite uma maior qualidade total, sem uma violação das exigências de armazenamento temporário que são impostas pelo tamanho de reservatório de bits. A presente invenção amplia o conceito do uso de um reservatório de bits para um controle do reservatório de bits para um codec de áudio generalizado com tamanhos de quadro variáveis. Portanto, um sistema de codifica-

ção de áudio pode compreender uma unidade de controle do reservatório de bits, para determinação do número de bits concedidos para codificar um quadro do sinal filtrado com base na duração do quadro e em uma medida de dificuldade do quadro. Preferivelmente, a unidade de controle do reservatório de bits tem equações de controle separadas para diferentes medidas de dificuldade do quadro e/ou diferentes tamanhos de quadro. Medidas de dificuldade para diferentes tamanhos de quadro podem ser normalizadas, então, elas podem ser comparadas mais facilmente. A fim de controlar a alocação de bit para um codificador de taxa variável, preferivelmente, a unidade de controle do reservatório de bits ajusta o limite mais baixo permitido do algoritmo de controle de bit concedido no número de bits médio para o maior tamanho de quadro permitido.

[00029] Um aspecto adicional da invenção refere-se ao tratamento de um reservatório de bits em um codificador que emprega um quantizador baseado em modelo, por exemplo, um Quantizador com Restrição de Entropia (ECQ). Sugere-se minimizar a variação do tamanho da etapa ECQ. Sugere-se uma equação de controle em particular que relaciona o tamanho da etapa do quantizador com a taxa ECQ.

[00030] Preferivelmente, o filtro adaptativo para filtrar o sinal de entrada é baseado em uma análise da Codificação por Predição Linear (LPC), incluindo um filtro LPC que produz um sinal de entrada branqueado. Parâmetros LPC para o presente quadro de dados de entrada podem ser determinados por algoritmos conhecidos na tecnologia. Uma unidade de estimação do parâmetro LPC pode calcular, para o quadro de dados de entrada, qualquer representação de parâmetro LPC adequado, tais como polinômios, funções de transferência, coeficientes de reflexão, frequências de linha espectral, etc. O tipo em particular da representação do parâmetro LPC que é usada para codificação ou outro processamento depende de respectivas exigências. Ver-

sados na técnica percebem que algumas representações são mais adequadas para certas operações do que outras e, portanto, são preferidas para realizar estas operações. A unidade de predição linear pode operar em um primeiro tamanho de quadro que é fixa, por exemplo, 20 ms. A filtragem de predição linear pode operar adicionalmente em um eixo geométrico de frequência deformado para enfatizar seletivamente certas faixas de frequência, tais como baixas frequências, em relação a outras frequências.

[00031] Preferivelmente, a transformada aplicada no quadro do sinal de entrada filtrado é uma Transformada de Cosseno Discreta Modificada (MDCT) que opera em um segundo tamanho de quadro variável. O sistema de codificação de áudio pode compreender uma unidade de controle de sequência de janela, que determina, para um bloco do sinal de entrada, as durações do quadro para sobrepor janelas MDCT pela minimização de uma função do custo de codificação, preferivelmente, uma entropia perceptiva simplista, para a íntegra do bloco do sinal de entrada que inclui diversos quadros. Assim, é derivada uma segmentação ideal do bloco do sinal de entrada nas janelas MDCT com respectivas segundas durações do quadro. Em consequência, é proposta uma estrutura de codificação de domínio transformado, incluindo elementos do codificador de fala, com um quadro MDCT de duração adaptativa apenas como unidade básica para todo o processamento, exceto a LPC. Já que as durações do quadro MDCT podem tomar muitos diferentes valores, uma sequência ideal pode ser encontrada, e mudanças abruptas da duração do quadro podem ser evitadas, como é comum na tecnologia anterior, em que apenas um pequeno tamanho de janela e um grande tamanho de janela são aplicados. Além do mais, janelas transformadas transicionais com bordas agudas, da forma usada em algumas abordagens da tecnologia anterior para a transição entre pequenos e grandes tamanhos de janela,

não são necessárias.

[00032] Preferivelmente, consecutivos comprimentos de janela MDCT mudam, no máximo, em um fator de dois (2) e/ou os comprimentos de janela MDCT são valores diádicos. Mais particularmente, os comprimentos de janela MDCT podem ser partições diádicas do bloco do sinal de entrada. Portanto, a sequência de janela MDCT é limitada às sequências predeterminadas que são fáceis de codificar com um pequeno número de bits. Além do mais, a sequência de janela tem suaves transições de tamanhos de quadro, desse modo, excluindo mudanças abruptas da duração do quadro.

[00033] A unidade de controle de sequência de janela pode ser adicionalmente configurada para considerar as estimativas de predição de longo prazo, geradas pela unidade de predição de longo prazo, para durações de janela candidatas, durante a busca da sequência de durações de janela MDCT que minimiza a função do custo de codificação para o bloco do sinal de entrada. Nesta modalidade, o laço de predição de longo prazo é fechado durante a determinação das durações de janela MDCT, o que resulta em uma melhor sequência de janelas MDCT aplicada para a codificação.

[00034] O sistema de codificação de áudio pode compreender adicionalmente um codificador LPC para codificação recursiva, em uma taxa variável, de frequências espectrais de linha ou outras representações de parâmetro LPC apropriadas, geradas pela unidade de predição linear, para armazenamento e/ou transmissão a um decodificador. De acordo com uma modalidade, uma unidade de interpolação de predição linear é provida para interpolar parâmetros de predição linear gerados em uma taxa correspondente a primeira tamanho de quadro para corresponder às durações variáveis do quadro do sinal do domínio transformado.

[00035] De acordo com um aspecto da invenção, o sistema de codi-

ficação de áudio pode compreender uma unidade de modelagem perceptiva que modifica uma característica do filtro adaptativo, pelo chia-  
do e/ou oscilação de um polinômio LPC gerado pela unidade de predi-  
ção linear para um quadro LPC. O modelo perceptivo recebido pela  
modificação das características do filtro adaptativo pode ser usado  
com muitos propósitos no sistema. Por exemplo, ele pode ser aplicado  
como função de ponderação perceptiva na quantização ou na predição  
de longo prazo.

[00036] Um outro aspecto da invenção refere-se à predição de lon-  
go prazo (LTP), em particular, à predição de longo prazo no domínio  
MDCT, LTP adaptado por quadro MDCT e busca LTP ponderada por  
MDCT. Estes aspectos são aplicáveis, independente se uma análise  
LPC está presente à montante do codificador de transformada.

[00037] De acordo com uma modalidade, o sistema de codificação  
de áudio compreende adicionalmente uma unidade de quantização  
invertida e transformação invertida, para geração de uma reconstru-  
ção do domínio temporal do quadro do sinal de entrada filtrado. Além  
do mais, pode ser provido um armazenamento temporário de predi-  
ção de longo prazo para armazenar reconstruções de domínio tempo-  
ral de quadros prévios do sinal de entrada filtrado. Estas unidades  
podem ser arranjadas em um laço de realimentação da unidade de  
quantização até uma unidade de extração de predição de longo prazo  
que busca, no armazenamento temporário de predição de longo pra-  
zo, o segmento reconstruído que melhor corresponde ao presente  
quadro do sinal de entrada filtrado. Além do mais, pode ser provida  
uma unidade de estimação de ganho de predição de longo prazo que  
ajusta o ganho do segmento selecionado do armazenamento tempo-  
rário de predição de longo prazo, de forma que ele melhor corres-  
ponda ao presente quadro. Preferivelmente, a estimação da predição  
de longo prazo é subtraída do sinal de entrada transformado no do-

mínio transformado. Portanto, pode ser provida uma segunda unidade transformada para transformar o segmento selecionado no domínio transformado. O laço de predição de longo prazo pode incluir adicionalmente adicionar a estimação da predição de longo prazo no domínio transformado ao sinal de realimentação, depois da quantização invertida e antes da transformada invertida no domínio temporal. Assim, pode ser usado um esquema de predição de longo prazo adaptativo regressivo que prediz, no domínio transformado, o presente quadro do sinal de entrada filtrado com base em quadros prévios. A fim de ser mais eficiente, o esquema de predição de longo prazo pode ser adicionalmente adaptado de diferentes maneiras, da forma exposta a seguir para alguns exemplos.

[00038] De acordo com uma modalidade, a unidade de predição de longo prazo compreende um extrator de predição de longo prazo, para determinação de um valor de atraso que especifica o segmento reconstruído do sinal filtrado que melhor se adéqua ao quadro atual do sinal filtrado. Um estimador de ganho de predição de longo prazo pode estimar um valor de ganho aplicado no sinal do segmento selecionado do sinal filtrado. Preferivelmente, o valor de atraso e o valor de ganho são determinados para minimizar um critério de distorção relacionado à diferença, em um domínio perceptivo, entre a estimação da predição de longo prazo e o sinal de entrada transformado. Um polinômio de predição linear modificado pode ser aplicado como curva de ganho de equalização do domínio MDCT durante a minimização do critério de distorção.

[00039] A unidade de predição de longo prazo pode compreender uma unidade de transformação para transformar o sinal de segmentos reconstruído do armazenamento temporário da LTP no domínio transformado. Para uma eficiente implementação de uma transformada MDCT, preferivelmente, a transformada é uma Transformada de Cos-

seno Discreta tipo IV.

[00040] Um outro aspecto da invenção refere-se a um decodificador de áudio para decodificação do fluxo de bits gerado pelas modalidades do codificador exposto. Um decodificador de acordo com uma modalidade compreende uma unidade de dequantização, para dequantizar um quadro de um fluxo de bits de entrada, com base em fatores de escala; uma unidade de transformação invertida, para transformar inversamente um sinal do domínio transformado; uma unidade de predição linear, para filtrar o sinal do domínio transformado inversamente transformado; e uma unidade de decodificação de fator de escala para gerar os fatores de escala usados na dequantização com base em informação delta do fator de escala recebida, que codifica a diferença entre os fatores de escala aplicados no codificador e os fatores de escala que são gerados com base em parâmetros do filtro adaptativo. O decodificador pode compreender adicionalmente uma unidade de determinação do fator de escala, para gerar fatores de escala com base em uma curva de limite de mascaramento, que é derivada dos parâmetros de predição linear para o presente quadro. A unidade de decodificação de fator de escala pode combinar a informação delta do fator de escala recebida e os fatores de escala gerados com base em predição linear para gerar fatores de escala para entrada na unidade de dequantização.

[00041] Um decodificador de acordo com uma outra modalidade compreende uma unidade de dequantização com base em modelo, para dequantizar um quadro de um fluxo de bits de entrada; uma unidade de transformação invertida, para transformar inversamente um sinal do domínio transformado; e uma unidade de predição linear para filtrar o sinal do domínio transformado inversamente transformado. A unidade de dequantização pode compreender um dequantizador não baseado em modelo e um dequantizador baseado em modelo.

[00042] Preferivelmente, a unidade de dequantização compreende pelo menos um modelo de probabilidade adaptativo. A unidade de dequantização pode ser configurada para adaptar a dequantização em função das características de sinal transmitidas.

[00043] A unidade de dequantização pode decidir adicionalmente uma estratégia de dequantização com base nos dados de controle para o quadro decodificado. Preferivelmente, os dados de controle de dequantização são recebidos com o fluxo de bits ou são derivados dos dados recebidos. Por exemplo, a unidade de dequantização decide a estratégia de dequantização com base no tamanho da transformada do quadro.

[00044] De acordo com um outro aspecto, a unidade de dequantização compreende pontos de reconstrução adaptativos. A unidade de dequantização pode compreender dequantizadores escalares uniformes que são configurados para usar dois pontos de reconstrução de dequantização por intervalo de quantização, em particular, um ponto médio e um ponto de reconstrução MMSE.

[00045] De acordo com uma modalidade, a unidade de dequantização usa um quantizador baseado em modelo em combinação com codificação aritmética.

[00046] Além do mais, o decodificador pode compreender muitos dos aspectos supradivulgados para o codificador. No geral, o decodificador espelhará as operações do codificador, embora algumas operações sejam realizadas apenas no codificador e não tenham componentes correspondentes no decodificador. Assim, o que é divulgado para o codificador também é considerável como aplicável para o decodificador, se não declarado de outra forma.

[00047] Os aspectos expostos da invenção podem ser implementados como um dispositivo, aparelho, método ou programa de computador que opera em um dispositivo programável. Aspectos inventivos

podem ser adicionalmente incorporados em sinais, estruturas de dados e fluxos contínuos de bits.

[00048] Assim, o pedido divulga adicionalmente um método de codificação de áudio e um método de decodificação de áudio. Um método exemplar de codificação de áudio compreende as etapas de: filtrar um sinal de entrada com base em um filtro adaptativo; transformar um quadro do sinal de entrada filtrado em um domínio transformado; quantizar o sinal do domínio transformado; gerar fatores de escala, com base em uma curva de limite de mascaramento, para uso na unidade de quantização durante a quantização do sinal do domínio transformado; estimar fatores de escala com base em predição linear com base em parâmetros do filtro adaptativo; e codificar a diferença entre os fatores de escala com base na curva de limite de mascaramento e os fatores de escala com base em predição linear.

[00049] Um outro método de codificação de áudio compreende as etapas de: filtrar um sinal de entrada com base em um filtro adaptativo; transformar um quadro do sinal de entrada filtrado em um domínio transformado; e quantizar o sinal do domínio transformado; em que a unidade de quantização decide, com base em características do sinal de entrada, codificar o sinal do domínio transformado com um quantizador baseado em modelo ou um quantizador não baseado em modelo.

[00050] Um método exemplar de decodificação de áudio compreende as etapas de: dequantizar um quadro de um fluxo de bits de entrada com base em fatores de escala; transformar inversamente um sinal do domínio transformado; filtrar com predição linear o sinal do domínio transformado inversamente transformado; estimar segundos fatores de escala com base em parâmetros do filtro adaptativo; e gerar os fatores de escala usados na dequantização com base na informação da diferença do fator de escala recebida e os segundos fatores de

escala estimados.

[00051] Um outro método de codificação de áudio compreende as etapas: dequantizar um quadro de um fluxo de bits de entrada; transformar inversamente um sinal do domínio transformado; e filtrar com predição linear o sinal do domínio transformado inversamente transformado; em que a dequantização usa um quantizador não baseado em modelo e um quantizador baseado em modelo.

[00052] Estes são apenas exemplos de métodos preferidos de codificação/decodificação de áudio e programas de computador que são preceituados pelo presente pedido e que versados na técnica podem derivar da seguinte descrição das modalidades exemplares.

#### Breve Descrição dos Desenhos

[00053] A presente invenção será agora descrita a título de exemplos ilustrativos, não limitando o escopo ou o espírito da invenção, em relação aos desenhos anexos, nos quais:

[00054] a figura 1 ilustra uma modalidade preferida de um codificador e um decodificador, de acordo com a presente invenção;

[00055] a figura 2 ilustra uma vista mais detalhada do codificador e do decodificador, de acordo com a presente invenção;

[00056] a figura 3 ilustra uma outra modalidade do codificador, de acordo com a presente invenção;

[00057] a figura 4 ilustra uma modalidade preferida do codificador, de acordo com a presente invenção;

[00058] a figura 5 ilustra uma modalidade preferida do decodificador, de acordo com a presente invenção;

[00059] a figura 6 ilustra uma modalidade preferida da codificação e da decodificação das linhas da MDCT, de acordo com a presente invenção;

[00060] a figura 7 ilustra uma modalidade preferida do codificador e decodificador, e exemplos de dados de controle relevantes transmiti-

dos de um para o outro, de acordo com a presente invenção;

[00061] a figura 7a é uma outra ilustração de aspectos do codificador, de acordo com uma modalidade da invenção;

[00062] a figura 8 ilustra um exemplo de uma sequência de janela e a relação entre dados da LPC e dados da MDCT, de acordo com uma modalidade da presente invenção;

[00063] a figura 9 ilustra uma combinação de dados de fator de escala e dados da LPC, de acordo com a presente invenção;

[00064] a figura 9a ilustra uma outra modalidade da combinação de dados de fator de escala e dados da LPC, de acordo com a presente invenção;

[00065] a figura 9b ilustra um outro diagrama de bloco simplificado de um codificador e um decodificador, de acordo com a presente invenção;

[00066] a figura 10 ilustra uma modalidade preferida da tradução de polinômios LPC para uma curva de ganho da MDCT, de acordo com a presente invenção;

[00067] a figura 11 ilustra uma modalidade preferida do mapeamento dos parâmetros LPC com taxa de atualização constante para os dados da sequência de janela MDCT adaptativa, de acordo com a presente invenção;

[00068] a figura 12 ilustra uma modalidade preferida da adaptação do cálculo do filtro de ponderação perceptiva com base no tamanho da transformada e no tipo de quantizador, de acordo com a presente invenção;

[00069] a figura 13 ilustra uma modalidade preferida da adaptação do quantizador dependente da duração do quadro, de acordo com a presente invenção;

[00070] a figura 14 ilustra uma modalidade preferida da adaptação do quantizador dependente da duração do quadro, de acordo com a

presente invenção;

[00071] a figura 15 ilustra uma modalidade preferida da adaptação do tamanho da etapa de quantização em função dos dados LPC e LTP, de acordo com a presente invenção;

[00072] a figura 15a ilustra como uma curva delta é derivada dos parâmetros LPC e LTP por meio de um módulo de adaptação delta;

[00073] a figura 16 ilustra uma modalidade preferida de um quantizador baseado em modelo que utiliza deslocamentos aleatórios, de acordo com a presente invenção;

[00074] a figura 17 ilustra uma modalidade preferida de um quantizador baseado em modelo de acordo com a presente invenção;

[00075] a figura 17a ilustra uma outra modalidade preferida de um quantizador baseado em modelo de acordo com a presente invenção;

[00076] a figura 17b ilustra esquematicamente um decodificador de linhas da MDCT com base em modelo 2150 de acordo com uma modalidade da invenção;

[00077] a figura 17c ilustra esquematicamente aspectos do pré-processamento do quantizador de acordo com uma modalidade da invenção;

[00078] a figura 17d ilustra esquematicamente aspectos da computação do tamanho da etapa de acordo com uma modalidade da invenção;

[00079] a figura 17e ilustra esquematicamente um codificador com restrição de entropia com base em modelo de acordo com uma modalidade da invenção;

[00080] a figura 17f ilustra esquematicamente a operação de um quantizador escalar uniforme (USQ) de acordo com uma modalidade da invenção;

[00081] a figura 17g ilustra esquematicamente computações de probabilidade de acordo com uma modalidade da invenção;

[00082] a figura 17h ilustra esquematicamente um processo de de-quantização de acordo com uma modalidade da invenção;

[00083] a figura 18 ilustra uma modalidade preferida de um controle de reservatório de bits de acordo com a presente invenção;

[00084] a figura 18a ilustra o conceito básico de um controle do reservatório de bits;

[00085] a figura 18b ilustra o conceito de um controle do reservatório de bits para tamanhos de quadro variáveis, de acordo com a presente invenção;

[00086] a figura 18c mostra uma curva de controle exemplar para o controle do reservatório de bits de acordo com uma modalidade;

[00087] a figura 19 ilustra uma modalidade preferida do quantizador invertido usando diferentes pontos de reconstrução, de acordo com a presente invenção.

#### Descrição das Modalidades Preferidas

[00088] As modalidades descritas a seguir são meramente ilustrativas para os princípios da presente invenção para codificador e decodificador de áudio. Entende-se que modificações e variações dos arranjos e dos detalhes aqui descritos serão aparentes aos outros versados na técnica. Portanto, pretende-se que sejam limitadas apenas pelo escopo das concretizações e não pelos detalhes específicos apresentados a título de descrição e explicação das modalidades aqui expostas. Componentes similares das modalidades são numerados por números de referência similares.

[00089] Na figura 1, um codificador 101 e um decodificador 102 são visualizados. O codificador 101 toma o sinal de entrada do domínio temporal e produz um fluxo de bits 103 subsequentemente transmitido ao decodificador 102. O decodificador 102 produz uma forma de onda de saída com base no fluxo de bits recebido 103. O sinal de saída se assemelha psicoacusticamente ao sinal de entrada original.

[00090] Na figura 2, uma modalidade preferida do codificador 200 e dos decodificadores 210 é ilustrada. O sinal de entrada no codificador 200 passa através de um módulo da LPC (Codificação por Predição Linear) 201 que gera um sinal residual branqueado para um quadro LPC, com um primeiro tamanho de quadro, e os correspondentes parâmetros de predição linear. Adicionalmente, normalização de ganho pode ser incluída no módulo da LPC 201. O sinal residual do LPC é transformado no domínio de frequência por um módulo da MDCT (Transformada de Cosseno Discreta Modificada) 202 que opera em um segundo tamanho de quadro variável. No codificador 200 representado na figura 2, um módulo da LTP (Predição de Longo Prazo) 205 é incluído. LTP será elaborada em uma modalidade adicional da presente invenção. As linhas da MDCT são quantizadas 203 e também dequantizadas 204, a fim de alimentar um armazenamento temporário da LTP com uma cópia da saída decodificada, como será disponível ao decodificador 210. Em função da distorção de quantização, esta cópia é chamada de reconstrução do respectivo sinal de entrada. Na parte inferior da figura 2, o decodificador 210 é representado. O decodificador 210 toma as linhas quantizadas da MDCT, as dequantiza 211, adiciona a contribuição do módulo da LTP 214, e faz uma transformada MDCT invertida 212, seguida por um filtro de síntese da LPC 213.

[00091] Um importante aspecto da modalidade exposta é que o quadro MDCT é a única unidade básica para codificação, embora a LPC tenha seus próprios (e, em uma modalidade, constantes) duração do quadro e parâmetros LPC, que também são codificados. A modalidade começa a partir de um codificador por transformada, e introduz módulos de predição e modelagem fundamentais de um codificador de fala. Como será discutido a seguir, a duração do quadro MDCT é variável e é adaptada a um bloco do sinal de entrada pela determinação

da sequência de janela MDCT ideal para a íntegra do bloco, pela minimização uma função de custo de entropia perceptiva simplista. Isto permite escalonar para manter controle de tempo/frequência ideal. Adicionalmente, a estrutura unificada proposta evita combinações comutadas ou em camadas de diferentes paradigmas de codificação.

[00092] Na figura 3, partes do codificador 300 são esquematicamente descritas com mais detalhes. O sinal branqueado transmitido do módulo da LPC 201 no codificador da figura 2 é inserido no banco de filtro da MDCT 302. Opcionalmente, a análise da MDCT pode ser uma análise da MDCT deformada em relação ao tempo, que garante que o passo do sinal (se o sinal for periódico com um passo bem definido) seja constante em relação à janela da transformada MDCT.

[00093] Na figura 3, o módulo da LTP 310 é esboçado com mais detalhes. Ele compreende um armazenamento temporário da LTP 311 que contém amostras do domínio temporal reconstruído dos segmentos prévios do sinal de saída. Um extrator da LTP 312 encontra o melhor segmento correspondente no armazenamento temporário da LTP 311, dado o atual segmento de entrada. Um valor de ganho adequado é aplicado neste segmento pela unidade de ganho 313 antes de ele ser subtraído do segmento que é atualmente inserido no quantizador 303. Evidentemente, a fim de fazer a subtração antes da quantização, o extrator da LTP 312 também transforma o segmento de sinal escolhido no domínio MDCT. O extrator da LTP 312 busca os melhores valores de ganho e atraso que minimizam uma função de erro no domínio perceptivo durante a combinação do segmento do sinal de saída prévio reconstruído com o quadro de entrada do domínio MDCT transformado. Por exemplo, uma função de erro quadrático médio (MSE) entre o segmento reconstruído transformado do módulo da LTP 310 e o quadro de entrada transformado (isto é, o sinal residual depois da subtração) é otimizada. Esta otimização pode ser realizada em um

domínio perceptivo, em que componentes de frequência (isto é, linhas da MDCT) são ponderadas de acordo com sua importância perceptiva. O módulo da LTP 310 opera em unidades do quadro MDCT, e o codificador 300 considera um quadro MDCT residual por vez, por exemplo, para quantização no módulo de quantização 303. A busca de atraso e ganho pode ser realizada em um domínio perceptivo. Opcionalmente, a LTP pode ser seletiva de frequência, isto é, adapta o ganho e/ou atraso em relação à frequência. Uma unidade de quantização invertida 304 e uma unidade MDCT invertida 306 são representadas. A MDCT pode ser deformada em relação ao tempo, da forma explicada posteriormente.

[00094] Na figura 4, é ilustrada uma outra modalidade do codificador 400. Além da figura 3, a análise LPC 401 é incluída para esclarecimento. Uma transformada DCT-IV 414 usada para transformar um segmento de sinal selecionado no domínio MDCT é mostrada. Adicionalmente, diversas maneiras de calcular o erro mínimo para a seleção do segmento LTP são ilustradas. Além da minimização do sinal residual mostrada na figura 4 (identificada como LTP2 na figura 4), é ilustrada a minimização da diferença entre o sinal de entrada transformado e o sinal do domínio MDCT dequantizado antes de ser inversamente transformado em um sinal do domínio temporal reconstruído para armazenamento no armazenamento temporário da LTP 411 (indicada como LTP3). A minimização desta função MSE direcionará a contribuição LTP na direção de uma similaridade ideal (conforme possível) do sinal de entrada transformado e do sinal de entrada reconstruído, para armazenamento no armazenamento temporário da LTP 411. Uma outra função de erro alternativa (indicada como LTP 1) é baseada na diferença destes sinais no domínio temporal. Neste caso, o MSE entre o quadro de entrada LPC filtrado e a correspondente reconstrução do domínio temporal no armazenamento temporário da LTP 411 é mini-

mizado. O MSE é vantajosamente calculado com base na duração do quadro MDCT, que pode ser diferente da duração do quadro da LPC. Adicionalmente, os blocos do quantizador e dequantizador são substituídos pelo bloco de codificação de espectro 403 e pelos blocos de decodificação de espectro 404 ("Spec enc" e "Spec dec") que podem conter módulos adicionais, além da quantização, como será esboçado na figura 6. Novamente, a MDCT e a MDCT invertida podem ser deformadas em relação ao tempo (WMDCT, IWMDCT).

[00095] Na figura 5, um decodificador proposto 500 é ilustrado. Os dados do espectro do fluxo de bits recebido são inversamente quantizados 511 e adicionados com uma contribuição LTP provida por um extrator da LTP a partir de um armazenamento temporário da LTP 515. O extrator da LTP 516 e a unidade de ganho de LTP 517 no decodificador 500 também são ilustrados. As linhas da MDCT somadas são sintetizadas no domínio temporal por um bloco de síntese da MDCT, e o domínio temporal sinal é espectralmente modelado por um filtro de síntese da LPC 513.

[00096] Na figura 6 os blocos "Spec dec" e "Spec enc" 403, 404 da figura 4 são descritos com mais detalhes. O bloco "Spec enc" 603 ilustrado na direita da figura compreende, em uma modalidade, um módulo de análise de Predição Harmônica 610, um módulo de análise TNS (Modelagem de Ruído Temporal) 611, seguido por um módulo de escalonamento de fator de escala 612 das linhas da MDCT, e finalmente, quantização e codificação das linhas em um módulo de linhas Enc 613. O bloco "Spec Dec" do decodificador 604, ilustrado na esquerda da figura, faz o processo inverso, isto é, as linhas da MDCT recebidas são dequantizadas em um módulo de linhas Dec 620, e o escalonamento é desfeito por um módulo de escalonamento do fator de escala (SCF) 621. Síntese TNS 622 e síntese de predição Harmônica 623 são aplicadas.

[00097] Na figura 7, é esboçada uma ilustração muito geral do sistema de codificação inventivo. O codificador exemplar toma o sinal de entrada e produz um fluxo de bits que contém, entre outros dados:

- \* linhas da MDCT quantizadas;
- \* fatores de escala;
- \* representação do polinômio LPC;
- \* energia do segmento de sinal (por exemplo, variância do sinal);
- \* sequência de janela;
- \* dados da LTP.

[00098] O decodificador de acordo com a modalidade lê o fluxo de bits provido e produz um sinal de saída de áudio, que se assemelha psicoacusticamente ao sinal original.

[00099] A figura 7a é uma outra ilustração dos aspectos de um codificador 700 de acordo com uma modalidade da invenção. O codificador 700 compreende um módulo da LPC 701, um módulo da MDCT 704, um módulo da LTP 705 (mostrado somente simplificado), um módulo de quantização 703 e um módulo de quantização invertida 704 para realimentar sinais reconstruídos no módulo da LTP 705. São adicionalmente providos um módulo de estimação de passo 750, para estimar o passo do sinal de entrada, e um módulo de determinação da sequência de janela 751, para determinar a sequência de janela MDCT ideal para um maior bloco do sinal de entrada (por exemplo, 1 segundo). Nesta modalidade, a sequência de janela MDCT é determinada com base em uma abordagem de laço aberto, em que é determinada a sequência de tamanho de janelas MDCT candidatas que minimiza uma função do custo de codificação, por exemplo, uma entropia perceptiva simplista. A contribuição do módulo da LTP 705 na função do custo de codificação, que é minimizada pelo módulo de determinação da sequência de janela 751, pode ser opcionalmente considerável

durante a busca da sequência de janela MDCT ideal. Preferivelmente, para cada tamanho de janela candidato avaliado, é determinada a melhor contribuição da predição de longo prazo ao quadro MDCT correspondente ao tamanho de janela candidata, e o respectivo custo de codificação é estimado. No geral, pequenos tamanhos de quadro MDCT são mais apropriados para entrada de fala, enquanto que longas janelas transformadas com uma fina resolução espectral são preferidas para sinais de áudio.

[000100] Ponderações perceptivas ou uma função de ponderação perceptiva são determinadas com base nos parâmetros LPC calculados pelo módulo da LPC 701, o que será explicado com mais detalhes a seguir. As ponderações perceptivas são supridas no módulo da LTP 705 e no módulo de quantização 703, ambos operando no domínio MDCT, para ponderação das contribuições de erro ou distorção dos componentes de frequência de acordo com suas respectivas importâncias perceptivas. A figura 7a ilustra adicionalmente quais parâmetros de codificação são transmitidos ao decodificador, preferivelmente, por um esquema de codificação apropriado, como será discutido a seguir.

[000101] A seguir, a coexistência da LPC e de dados da MDCT, e a emulação do efeito da LPC na MDCT, tanto para ação contrária quanto omissão de filtragem real, será discutida.

[000102] De acordo com uma modalidade, o módulo LP filtra o sinal de entrada, de forma que a forma espectral do sinal seja removida, e a subsequente saída do módulo LP seja um sinal espectralmente nivelado. Isto é vantajoso, por exemplo, para a operação da LTP. Entretanto, outras partes do codec que operam no sinal espectralmente nivelado podem se beneficiar por saber qual a forma espectral do sinal original era anterior à filtragem LP. Uma vez que os módulos do codificador, depois da filtragem, operam na transformada MDCT do sinal espec-

tralmente nivelado, a presente invenção preceitua que, se necessário, a forma espectral do sinal original anterior à filtragem LP pode ser reimposta na representação MDCT do sinal espectralmente nivelado, pelo mapeamento da função de transferência do filtro LP usado (isto é, o envelope espectral do sinal original) em uma curva de ganho, ou curva de equalização, que é aplicada nas posições de frequência da representação MDCT do sinal espectralmente nivelado. Inversamente, o módulo LP pode omitir a filtragem real, e apenas estimar uma função de transferência que é subseqüentemente mapeada para uma curva de ganho, que pode ser imposta na representação MDCT do sinal, assim, removendo a necessidade da filtragem do domínio temporal do sinal de entrada.

[000103] Um proeminente aspecto de modalidades da presente invenção é que um codificador por transformada com base em MDCT é operado usando uma segmentação de janela flexível, em um sinal LPC branqueado. Isto é esboçado na figura 8, em que uma sequência de janela MDCT exemplar é dada, juntamente com o trabalho em janelas do LPC. Portanto, como fica claro a partir da figura, a LPC opera em um tamanho de quadro constante (por exemplo, 20 ms), enquanto a MDCT opera em uma sequência de janela variável (por exemplo, 4 a 128 ms). Isto permite que a escolha de durações de janela ideais para a LPC e da sequência de janela ideal seja a MDCT, independentemente.

[000104] A figura 8 ilustra adicionalmente a relação entre os dados da LPC, em particular, os parâmetros da LPC gerados em uma primeira taxa de quadros e dados MDCT, em particular, as linhas da MDCT, gerados em uma segunda taxa variável. As setas para baixo na figura simbolizam dados da LPC que são interpolados entre os quadros LPC (círculos), para corresponder aos quadros MDCT correspondentes. Por exemplo, uma função de ponderação perceptiva gerada por LPC é

interpolada para instâncias de tempo determinadas pela sequência de janela MDCT.

[000105] As setas para cima simbolizam dados de refinamento (isto é, dados de controle) usados para a codificação das linhas da MDCT. Para os quadros AAC, tipicamente, estes dados são fatores de escala e, para os quadros ECQ, tipicamente, os dados são dados de correção de variância, etc. As linhas cheias em relação às linhas tracejadas representam que dados são os dados mais "importantes" para a codificação das linhas da MDCT, dado um certo quantizador. As setas duplas para baixo simbolizam as linhas espectrais do codec.

[000106] A coexistência de dados da LPC e da MDCT no codificador pode ser explorada, por exemplo, para reduzir as exigências de bit de fatores de escala da codificação MDCT, pela consideração de uma curva de mascaramento perceptiva estimada dos parâmetros da LPC. Além do mais, ponderação perceptiva derivada da LPC pode ser usada durante a determinação da distorção de quantização. Da forma ilustrada, e como será discutido a seguir, o quantizador opera em dois modos, e gera dois tipos de quadros (quadros ECQ e quadros AAC) dependentes da duração do quadro dos dados recebidos, isto é, correspondente ao quadro MDCT ou tamanho da janela.

[000107] A figura 11 ilustra uma modalidade preferida do mapeamento dos parâmetros LPC com taxa constante para os dados da sequência de janela MDCT adaptativa. Um módulo de mapeamento LPC 1100 recebe os parâmetros LPC de acordo com a taxa de atualização LPC. Além do mais, o módulo de mapeamento LPC 1100 recebe informação na sequência de janela MDCT. Então, ele gera um mapeamento LPC-para-MDCT, por exemplo, para mapeamento dos dados psicoacústicos com base em LPC para respectivos quadros MDCT gerados na taxa de quadros MDCT variável. Por exemplo, o módulo de mapeamento LPC interpola polinômios LPC ou dados relacionados para ins-

tâncias de tempo correspondentes aos quadros MDCT para uso, por exemplo, das ponderações perceptivas no módulo da LTP ou quantizador.

[000108] Agora, especificações do modelo perceptivo com base em LPC são discutidas pela referência à figura 9. Em uma modalidade da presente invenção, o módulo da LPC 901 é adaptado para produzir um sinal de saída branco, pelo uso de predição linear, por exemplo, de ordem 16, para um sinal de taxa de amostragem de 16 kHz. Por exemplo, a saída do módulo da LPC 201 na figura 2 é o resíduo depois da estimação e filtragem do parâmetro LPC. O polinômio LPC  $A(z)$  estimado, esquematicamente visualizado na parte inferior esquerda da figura 9, pode ser chiado por um fator de expansão de largura de banda, e também oscilado, em uma implementação da invenção, pela modificação do primeiro coeficiente de reflexão do polinômio LPC correspondente. Chiado expande a largura de banda de picos na função de transferência LPC, pelo movimento dos polos do polinômio para dentro do círculo da unidade, assim, resultando em picos mais suaves. Oscilação permite tornar a função de transferência LPC mais suave, a fim de equilibrar a influência de frequências inferiores e superiores. Estas modificações tentam gerar uma curva de mascaramento perceptiva  $A'(z)$  dos parâmetros LPC estimados que estarão disponíveis tanto no lado do codificador quanto no lado do decodificador do sistema. Detalhes da manipulação do polinômio LPC são apresentados na figura 12, a seguir.

[000109] A codificação MDCT que opera no resíduo da LPC tem, em uma implementação da invenção, fatores de escala para controlar a resolução do quantizador ou os tamanhos de etapa de quantização (e, assim, o ruído introduzido pela quantização). Estes fatores de escala são estimados por um módulo de estimação do fator de escala 960 no sinal de entrada original. Por exemplo, os fatores de escala são deri-

vados de uma curva de limite de mascaramento perceptivo estimado do sinal original. Em uma modalidade, uma transformada de frequência separada (possivelmente, com uma resolução de frequência diferente) pode ser usada para determinar a curva de limite de mascaramento, mas isto não é sempre necessário. Alternativamente, a curva de limite de mascaramento é estimada a partir das linhas da MDCT geradas pelo módulo de transformada. A parte da base direita da figura 9 ilustra esquematicamente fatores de escala gerados pelo módulo de estimação do fator de escala 960, para controlar a quantização, de forma que o ruído de quantização introduzido seja limitado a distorções inaudíveis.

[000110] Se um filtro LPC for conectado à montante do módulo de transformação MDCT, um sinal branqueado é transformado no domínio MDCT. Já que este sinal tem um espectro branco, ele não é bem adequado para dele derivar uma curva de mascaramento perceptiva. Assim, uma curva de ganho de equalização do domínio MDCT, gerada para compensar o branqueamento do espectro, pode ser usada durante a estimação da curva de limite de mascaramento e/ou dos fatores de escala. Isto é em virtude de os fatores de escala precisarem ser estimados em um sinal que tem propriedades de espectro absolutas do sinal original, a fim de estimar corretamente mascaramento perceptivo. O cálculo da curva de ganho de equalização do domínio MDCT do polinômio LPC é discutido com mais detalhes em relação à figura 10 a seguir.

[000111] Uma modalidade do esquema de estimação do fator de escala esboçado exposto é esboçada na figura 9a. Nesta modalidade, o sinal de entrada é inserido no módulo LP 901 que estima o envelope espectral do sinal de entrada descrito por  $A(z)$ , e também transmite o dito polinômio como uma versão filtrada do sinal de entrada. O sinal de entrada é filtrado com o inverso de  $A(z)$ , a fim de obter um sinal espec-

tralmente branco subsequentemente usado por outras partes do codificador. O sinal filtrado  $\tilde{x}(n)$  é inserido em uma unidade de transformação MDCT 902, enquanto o polinômio  $A(z)$  é inserido em uma unidade de cálculo da curva de ganho da MDCT 970 (da forma esboçada na figura 14). A curva de ganho estimada a partir do polinômio LP é aplicada nos coeficientes ou linhas da MDCT, a fim de reter o envelope espectral do sinal de entrada original anterior à estimação do fator de escala. As linhas da MDCT com ganho ajustado são inseridas no módulo de estimação do fator de escala 960, que estima os fatores de escala para o sinal de entrada.

[000112] Usando a abordagem supraesboçada, os dados transmitidos entre o codificador e o decodificador contêm tanto o polinômio LP, a partir do qual a informação perceptiva relevante bem como um modelo de sinal podem ser derivados quando um quantizador baseado em modelo for usado, quanto os fatores de escala comumente usados em um codec transformado.

[000113] Com mais detalhes, retornando à figura 9, o módulo da LPC 901 na figura estima, a partir do sinal de entrada, um envelope espectral  $A(z)$  do sinal e deriva, a partir deste, uma representação perceptiva  $A'(z)$ . Além do mais, fatores de escala normalmente usados em codecs de áudio perceptivos com base em transformada são estimados no sinal de entrada, ou eles podem ser estimados no sinal branco produzido por um filtro LP, se a função de transferência do filtro LP for levada em consideração na estimação do fator de escala (descrito no contexto da figura 10 a seguir). Então, os fatores de escala podem ser adaptados no módulo de adaptação do fator de escala 961, dado o polinômio LP, como será esboçado a seguir, a fim de reduzir a taxa de bits exigida para transmitir fatores de escala.

[000114] Normalmente, os fatores de escala são transmitidos ao decodificador, assim como o polinômio LP. Agora, dado que ambos são

estimados a partir do sinal de entrada original e que ambos são um tanto correlacionados às propriedades de espectro absolutas do sinal de entrada original, propõe-se codificar uma representação delta entre os dois, a fim de remover qualquer redundância que pode ocorrer se ambos fossem separadamente transmitidos. De acordo com uma modalidade, esta correlação é explorada como segue. Uma vez que o polinômio LPC, quando corretamente chiado e oscilado, tenta representar uma curva de limite de mascaramento, as duas representações podem ser combinadas, de forma que os fatores de escala do codificador por transformada transmitidos representem a diferença entre os fatores de escala desejados e aqueles que podem ser derivados a partir do polinômio LPC transmitido. Portanto, o módulo de adaptação do fator de escala 961 mostrado na figura 9 calcula a diferença entre os fatores de escala desejados gerados a partir do sinal de entrada original e os fatores de escala derivados da LPC. Este aspecto retém a capacidade de ter um quantizador com base em MDCT que tem a noção dos fatores de escala comumente usados nos codificadores de transformada, em uma estrutura da LPC, operando em um resíduo da LPC, e ainda tem a possibilidade de comutar para um quantizador baseado em modelo que deriva tamanhos de etapa de quantização exclusivamente dos dados de predição linear.

[000115] Na figura 9b, é dado um diagrama de bloco simplificado do codificador e do decodificador de acordo com uma modalidade. O sinal de entrada no codificador passa através do módulo da LPC 901, que gera um sinal residual branqueado e os correspondentes parâmetros de predição linear. Adicionalmente, normalização de ganho pode ser incluída no módulo da LPC 901. O sinal residual da LPC é transformado no domínio de frequência por uma transformada MDCT 902. À direita da figura 9b, é representado o decodificador. O decodificador toma as linhas da MDCT quantizadas, as dequantiza 911 e aplica uma

transformada MDCT invertida 912, seguida por um filtro de síntese da LPC 913.

[000116] O sinal branqueado transmitido pelo módulo da LPC 901 no codificador da figura 9b é inserido no banco de filtro da MDCT 902. As linhas da MDCT, em decorrência da análise da MDCT, são transformadas codificadas com um algoritmo de codificação da transformada, que consiste em um modelo perceptivo que guia o tamanho da etapa de quantização desejado para diferentes partes do espectro MDCT. Os valores que determinam o tamanho da etapa de quantização são chamados fatores de escala, e há um valor de fator de escala necessário para cada partição, chamado de banda de fator de escala, do espectro MDCT. Em algoritmos de codificação da transformada da tecnologia anterior, os fatores de escala são transmitidos, por meio do fluxo de bits, ao decodificador.

[000117] De acordo com um aspecto da invenção, a curva de mascaramento perceptiva estimada dos parâmetros da LPC, da forma explicada em relação à figura 9, é usada durante a codificação dos fatores de escala usados na quantização. Uma outra possibilidade de estimar uma curva de mascaramento perceptiva é usar os coeficientes do filtro LPC não modificados para uma estimativa da distribuição de energia nas linhas da MDCT. Com esta estimativa de energia, um modelo psicoacústico, usado em esquemas de codificação transformada, pode ser aplicado tanto no codificador quanto no decodificador, para obter uma estimativa de uma curva de mascaramento.

[000118] Então, as duas representações de uma curva de mascaramento são combinadas, de forma que os fatores de escala a ser transmitidos pelo codificador por transformada representem a diferença entre os fatores de escala desejados e aqueles que podem ser derivados a partir do polinômio LPC transmitido ou o modelo psicoacústico com base em LPC. Este recurso retém a capacidade de ter um

quantizador com base em MDCT que tem a noção de fatores de escala comumente usados em codificadores de transformada, em uma estrutura da LPC, operando em um resíduo da LPC, e ainda tem a possibilidade de controlar ruído de quantização com base em banda de fator de escala de acordo com o modelo psicoacústico do codificador de transformada. A vantagem é que a transmissão da diferença dos fatores de escala custará menos bits, se comparada à transmissão dos valores de fator de escala absolutos, sem considerar os dados da LPC já presentes. Dependendo da taxa de bits, da duração do quadro ou de outros parâmetros, a quantidade de fator de escala residual a ser transmitido pode ser selecionada. Para ter controle completo de cada banda de fator de escala, um fator de escala delta pode ser transmitido com um esquema de codificação sem ruído apropriado. Em outros casos, o custo para transmitir fatores de escala pode ser adicionalmente reduzido por uma representação mais grosseira das diferenças do fator de escala. O caso especial com sobreprocessamento mais baixo é quando a diferença do fator de escala for ajustada em 0 para todas as bandas e nenhuma informação adicional for transmitida.

[000119] A figura 10 ilustra uma modalidade preferida da tradução de polinômios LPC em uma curva de ganho MDCT. Da forma esboçada na figura 2, a MDCT opera em um sinal branqueado, branqueado pelo filtro LPC 1001. A fim de reter o envelope espectral do sinal de entrada original, uma curva de ganho da MDCT é calculada pelo módulo da curva de ganho da MDCT 1070. A curva de ganho de equalização do domínio MDCT pode ser obtida pela estimação da resposta de magnitude do envelope espectral descrito pelo filtro LPC, para as frequências representadas pelas posições na transformada MDCT. Então, a curva de ganho pode ser aplicada nos dados MDCT, por exemplo, durante o cálculo o sinal de erro quadrático médio mínimo esboçado na figura 3 ou durante a estimação de uma curva de mascaramento per-

ceptiva para determinação do fator de escala como esboçada em relação à figura 9 exposta.

[000120] A figura 12 ilustra uma modalidade preferida da adaptação do cálculo do filtro de ponderação perceptiva com base no tamanho da transformada e/ou no tipo de quantizador. O polinômio LP  $A(z)$  é estimado pelo módulo da LPC 1201 na figura 16. Um módulo de modificação de parâmetro LPC 1271 recebe parâmetros da LPC, tal como o polinômio LPC  $A(z)$ , e gera um filtro de ponderação perceptiva  $A'(z)$ , pela modificação dos parâmetros da LPC. Por exemplo, a largura de banda do polinômio LPC  $A(z)$  é ampliada e/ou o polinômio é oscilado. Os parâmetros inseridos no módulo de chiado e oscilação de adaptação 1272 são os valores padrões de chiado e oscilação  $\rho$  e  $\lambda$ . Estas são regras predeterminadas modificadas dadas, baseadas no tamanho da transformada usada e/ou na estratégia de quantização  $Q$  usada. Os parâmetros de chiado e oscilação modificados  $\rho'$  e  $\lambda'$  são inseridos no módulo de modificação de parâmetro LPC 1271, traduzindo o envelope espectral do sinal de entrada, representado por  $A(z)$ , em uma curva de mascaramento perceptiva representada por  $A'(z)$ .

[000121] A seguir, será explicada a estratégia de quantização condicionada na duração do quadro, e a quantização com base em modelo condicionada em parâmetros variados de acordo com uma modalidade da invenção. Um aspecto da presente invenção é que ela utiliza diferentes estratégias de quantização para diferentes tamanhos de transformada ou tamanhos de quadro. Isto é ilustrado na figura 13, em que a duração do quadro é usada como um parâmetro de seleção para usar um quantizador baseado em modelo ou um quantizador não baseado em modelo. Percebe-se que este aspecto da quantização é independente de outros aspectos do codificador/decodificador divulgado e também pode ser aplicado em outros codecs. Um exemplo de um quantizador não baseado em modelo é o quantizador baseado na ta-

bela Huffman usado no padrão de codificação de áudio AAC. O quantizador baseado em modelo pode ser um Quantizador com Restrição de Entropia (ECQ) que emprega codificação aritmética. Entretanto, outros quantizadores também podem ser usados em modalidades da presente invenção.

[000122] De acordo com um aspecto independente da presente invenção, sugere-se comutar entre diferentes estratégias de quantização em função da duração do quadro, a fim de poder usar a estratégia de quantização ideal, dado uma duração do quadro em particular. Como um exemplo, a sequência de janelas pode indicar o uso de uma transformada longa para um segmento de música tonal muito estacionário do sinal. Para este tipo de sinal em particular, usando uma transformada longa, é altamente benéfico empregar uma estratégia de quantização que pode tirar vantagem do caráter "esparso" (isto é, tons discretos bem definidos) no espectro do sinal. Um método de quantização usado em AAC, em combinação com tabelas Huffman e agrupamento de linhas espectrais, também usados em AAC, é muito benéfico. Entretanto, e ao contrário, para segmentos de fala, a sequência de janelas pode, dado o ganho de codificação da LTP, indicar o uso de transformadas curtas. Para este tipo de sinal e tamanho da transformada, é benéfico empregar uma estratégia de quantização que não tenta encontrar ou introduzir espalhamento no espectro, mas, em vez disto, mantém uma energia de banda larga que, dada a LTP, reterá a característica tipo pulso do sinal de entrada original.

[000123] Uma visualização mais geral deste conceito é dada na figura 14, em que o sinal de entrada é transformado no domínio MDCT e subsequentemente quantizado por um quantizador controlado pelo tamanho da transformada ou duração do quadro usado para a transformada MDCT.

[000124] De acordo com um outro aspecto da invenção, o tamanho

da etapa do quantizador é adaptado em função dos dados LPC e/ou LTP. Isto permite uma determinação do tamanho da etapa dependente da dificuldade de um quadro e dos controles do número de bits que são alocados para codificação do quadro. Na figura 15, é dada uma ilustração sobre como a quantização com base em modelo pode ser controlada pelos dados LPC e LTP. Na parte de topo da figura 15, é dada uma visualização esquemática das linhas da MDCT. Abaixo do tamanho da etapa de quantização, é representado delta  $\Delta$  em função da frequência. Fica claro a partir deste exemplo em particular que o tamanho da etapa de quantização aumenta com a frequência, isto é, incorre-se em mais distorção da quantização para frequências mais altas. A curva delta é derivada dos parâmetros LPC e LTP por meio de um módulo de adaptação delta representado na figura 15a. A curva delta pode ser adicionalmente derivada do polinômio de predição  $A(z)$  por chiado e/ou oscilação, da forma explicada em relação à figura 13.

[000125] Uma função de ponderação perceptiva preferida derivada dos dados da LPC é dada na seguinte equação:

$$P(z) = \frac{1 - (1 - \tau)r_1 z^{-1}}{A(z/\rho)}$$

em que  $A(z)$  é o polinômio LPC,  $\tau$  é um parâmetro de oscilação,  $\rho$  controla o chiado e  $r_1$  é o primeiro coeficiente de reflexão calculado a partir do polinômio  $A(z)$ . Percebe-se que o polinômio  $A(z)$  pode ser recalculado em uma variedade de diferentes representações, a fim de extrair informação relevante do polinômio. Se houver interesse na inclinação espectral, a fim de aplicar uma "oscilação" para se opor à inclinação do espectro, o recálculo do polinômio em relação aos coeficientes de reflexão é preferido, já que o primeiro coeficiente de reflexão representa a inclinação do espectro.

[000126] Além do mais, os valores delta  $\Delta$  podem ser adaptados em função da variância do sinal de entrada  $\rho$ , o ganho de LTP  $g$ , e o pri-

meiro coeficiente de reflexão  $r_1$  derivado do polinômio de predição. Por exemplo, a adaptação pode ser baseada na seguinte equação:

$$\Delta' = \Delta(1 + r_1(1 - g^2))$$

[000127] A seguir, são esboçados aspectos de um quantizador baseado em modelo de acordo com uma modalidade da presente invenção. Na figura 16, um dos aspectos do quantizador baseado em modelo é visualizado. As linhas da MDCT são inseridas em um quantizador que emprega quantizadores escalares uniformes. Além do mais, deslocamentos aleatórios são inseridos no quantizador e usados como valores de deslocamento para os intervalos de quantização que movem os limites do intervalo. O quantizador proposto provê vantagens da quantização vetorial, ainda mantendo a buscabilidade de quantizadores escalares. O quantizador itera sobre um conjunto de diferentes valores de deslocamento, e calcula o erro de quantização para estes. O valor de deslocamento (ou valor do vetor de deslocamento) que minimiza a distorção da quantização para as linhas da MDCT em particular que estão sendo quantizadas é usado para quantização. Então, o valor de deslocamento é transmitido para o decodificador juntamente com as linhas da MDCT quantizadas. O uso de deslocamentos aleatórios introduz preenchimento de ruído no sinal decodificado dequantizado e, fazendo isto, evita furos espectrais no espectro quantizado. Isto é particularmente importante para baixas taxas de bit, em que muitas linhas da MDCT são de outra forma quantizadas em um valor zero, o que levará a furos audíveis no espectro do sinal reconstruído.

[000128] A figura 17 ilustra esquematicamente um Quantizador de Linhas da MDCT com base em Modelo (MBMLQ) de acordo com uma modalidade da invenção. O topo da figura 17 representa um codificador MBMLQ 1700. O codificador MBMLQ 1700 toma como entrada as linhas da MDCT em um quadro MDCT ou as linhas da MDCT da LTP

residual, se uma LTP estiver presente no sistema. O MBMLQ emprega modelos estatísticos das linhas da MDCT, e códigos-fontes são adaptados às propriedades do sinal com base em quadro a quadro da MDCT, produzindo eficiente compressão em um fluxo de bits.

[000129] Um ganho local das linhas da MDCT pode ser estimado como o valor RMS das linhas da MDCT, e as linhas da MDCT normalizadas no módulo de normalização de ganho 1720 antes da inserção no codificador MBMLQ 1700. O ganho local normaliza as linhas da MDCT e é um complemento à normalização de ganho da LP. Enquanto o ganho da LP se adapta às variações no nível do sinal em uma maior escala de tempo, o ganho local se adapta às variações em uma menor escala de tempo, produzindo maior qualidade de sons transitórios e inícios na fala. O ganho local é codificado por codificação em taxa fixa ou taxa variável e transmitido ao decodificador.

[000130] Um módulo de controle de taxa 1710 pode ser empregado para controlar o número de bits usados para codificar um quadro MDCT. Um índice de controle de taxa controla o número de bits usados. O índice de controle de taxa aponta para uma lista de tamanhos nominais de etapa do quantizador. A tabela pode ser classificada com tamanhos de etapa em ordem descendente (veja figura 17g).

[000131] O codificador MBMLQ é executado com um conjunto de diferentes índices de controle de taxa, e o índice de controle de taxa que produz uma contagem de bits menor que o número de bits concedidos, dado pelo controle do reservatório de bits, é usado para o quadro. O índice de controle de taxa varia lentamente, e isto pode ser explorado para reduzir complexidade da busca e para codificar eficientemente o índice. O conjunto de índices que é testado pode ser reduzido se o teste for iniciado ao redor do índice do quadro MDCT anterior. Igualmente, a eficiente codificação por entropia do índice é obtida se as probabilidades alcançarem o máximo ao redor do prévio valor do

índice. Por exemplo, para uma lista de 32 tamanhos de etapa, o índice de controle de taxa pode ser codificado usando 2 bits por quadro MDCT na média.

[000132] A figura 17 ilustra esquematicamente de forma adicional o decodificador MBMLQ 1750, em que o quadro MDCT é renormalizado em relação ao ganho se um ganho local for estimado no codificador 1700.

[000133] A figura 17a ilustra esquematicamente com mais detalhes o codificador de linhas da MDCT com base em modelo 1700, de acordo com uma modalidade. Ele compreende um módulo de pré-processamento do quantizador 1730 (veja figura 17c), um codificador com restrição de entropia com base em modelo 1740 (veja figura 17e), e um codificador aritmético 1720 que pode ser um codificador aritmético da tecnologia anterior. A tarefa do módulo de pré-processamento do quantizador 1730 é adaptar o codificador MBMLQ na estatística de sinal, em uma base quadro a quadro da MDCT. Ele toma como entrada outros parâmetros do codec, e deles deriva estatísticas úteis sobre o sinal, que podem ser usadas para modificar o comportamento do codificador com restrição de entropia com base em modelo 1740. O codificador com restrição de entropia com base em modelo 1740 é controlado, por exemplo, por um conjunto de parâmetros de controle: um tamanho da etapa do quantizador  $\Delta$  (delta, duração do intervalo), um conjunto de estimações de variância das linhas  $V$  da MDCT (um vetor; um valor estimado por linha da MDCT), uma curva de mascaramento perceptiva  $P_{\text{mod}}$ , uma matriz ou tabela de deslocamentos (aleatórios), e um modelo estatístico das linhas da MDCT que descrevem a forma da distribuição das linhas da MDCT e suas interdependências. Todos os parâmetros de controle supramencionados podem variar entre quadros da MDCT.

[000134] A figura 17b ilustra esquematicamente um decodificador de

linhas da MDCT com base em modelo 1750 de acordo com uma modalidade da invenção. Ele toma como entrada bits de informação complementar do fluxo de bits e os decodifica em parâmetros que são inseridos no módulo de pré-processamento do quantizador 1760 (veja figura 17c). Preferivelmente, o módulo de pré-processamento do quantizador 1760 tem exatamente a mesma funcionalidade no codificador 1700 como no decodificador 1750. Os parâmetros que são inseridos no módulo de pré-processamento do quantizador 1760 são exatamente os mesmos no codificador e no decodificador. O módulo de pré-processamento do quantizador 1760 transmite um conjunto de parâmetros de controle (mesmo do codificador 1700) e estes são inseridos no módulo de computações de probabilidade 1770 (veja figura 17g; mesmo do codificador, veja figura 17e) e no módulo de dequantização 1780 (veja figura 17h; mesmo do codificador, veja figura 17e). As tabelas cdf do módulo de computações de probabilidade 1770, que representam as funções de densidade de probabilidade para todas as linhas da MDCT, dado o delta usado para quantização e a variância do sinal, são inseridas no decodificador aritmético (que pode ser qualquer codificador aritmético conhecido pelos versados na técnica) que, então, decodifica os bits das linhas da MDCT em índices das linhas da MDCT. Então, os índices das linhas da MDCT são dequantizados em linhas da MDCT pelo módulo de dequantização 1780.

[000135] A figura 17c ilustra esquematicamente aspectos do pré-processamento do quantizador de acordo com uma modalidade da invenção, que consistem em: i) computação do tamanho da etapa, ii) modificação da curva de mascaramento perceptiva, iii) estimação da variância das linhas da MDCT, iv) construção da tabela de deslocamento.

[000136] A computação do tamanho da etapa é explicada com mais detalhes na figura 17d. Ela compreende: i) uma tabela de busca, em

que o índice de controle de taxa aponta para uma tabela de tamanhos de etapa, produz um tamanho nominal da etapa  $\Delta_{\text{nom}}$  (delta\_nom), ii) adaptação de baixa energia e iii) adaptação passa-alta.

[000137] Normalmente, normalização de ganho resulta em que sons de alta energia e sons de baixa energia são codificados com o mesmo SNR segmental. Isto pode levar a um excessivo número de bits sendo usado nos sons de baixa energia. A adaptação de baixa energia proposta permite o ajuste fino de um acerto entre os sons de baixa energia e de alta energia. O tamanho da etapa pode aumentar quando a energia do sinal ficar baixa, da forma representada na figura 17d-ii), em que é mostrada uma curva exemplar para a relação entre a energia do sinal (ganho  $g$ ) e um fator de controle  $q_{Le}$ . O ganho do sinal  $g$  pode ser computado como o valor RMS do próprio sinal de entrada ou da LP residual. A curva de controle na figura 17d-ii) é apenas um exemplo, e outras funções de controle para aumentar o tamanho da etapa para sinais de baixa energia podem ser empregadas. No exemplo representado, a função de controle é determinada por seções lineares graduais que são definidas por limites  $T_1$  e  $T_2$  e o fator do tamanho da etapa  $L$ .

[000138] Sons passa-alta são perceptivamente menos importantes que sons passa-baixa. A função de adaptação passa-alta aumenta o tamanho da etapa quando o quadro MDCT for passa-alta, isto é, quando a energia do sinal no presente quadro MDCT for concentrada nas frequências superiores, resultando em menos bits gastos em tais quadros. Se LTP estiver presente e se o ganho de LTP  $g_{LTP}$  estiver próximo de 1, a LTP residual pode se tornar passa-alta; em um caso como este, é vantajoso não aumentar o tamanho da etapa. Este mecanismo é representado na figura 17d-iii), em que  $r$  é o 1º coeficiente de reflexão da LPC. A adaptação passa-alta proposta pode usar a seguinte equação:

$$q_{lp} = \begin{cases} 1+r(1-g^2) & \text{se } r > 0 \\ 1 & \text{se } r \leq 0 \end{cases}$$

[000139] A figura 17c-ii) ilustra esquematicamente a modificação da curva de mascaramento perceptiva que emprega um aumento de baixa frequência (LF) para remover artefatos de codificação "tipo ronco". O aumento LF pode ser fixo ou adaptativo, de forma que apenas uma parte abaixo do primeiro pico espectral seja aumentada. O aumento LF pode ser adaptado pelo uso dos dados de envelope LPC.

[000140] A figura 17c-iii) ilustra esquematicamente a estimação da variância das linhas da MDCT. Com um filtro de branqueamento LPC ativo, todas as linhas da MDCT têm variância unitária (de acordo com o envelope LPC). Depois da ponderação perceptiva no codificador com restrição de entropia com base em modelo 1740 (veja figura 17e), as linhas da MDCT têm variâncias que são o inverso da curva de mascaramento perceptiva quadrática, ou a curva de mascaramento modificada quadrática  $P_{mod}$ . Se uma LTP estiver presente, ela pode reduzir a variância das linhas da MDCT. Na figura 17c-iii), é representado um mecanismo que adapta as variâncias estimadas à LTP. A figura mostra uma função de modificação  $g_{LTP}$  em relação à frequência  $f$ . As variâncias modificadas podem ser determinadas por  $V_{LTPmod} = V \cdot q_{LTP}$ . O valor  $L_{LTP}$  pode ser em função do ganho de LTP, de forma que  $L_{LTP}$  fique mais próximo de 0 se o ganho de LTP estiver ao redor de 1 (indicando que a LTP encontrou uma boa correspondência), e  $L_{LTP}$  fique mais próximo de 1 se o ganho de LTP estiver ao redor de 0. A adaptação LTP proposta das variâncias  $V = \{v_1, v_2, \dots, v_j, \dots, v_N\}$  afeta apenas as linhas da MDCT abaixo de uma certa frequência ( $f_{LTPcutoff}$ ). Em resultado, variâncias da linha da MDCT abaixo da frequência de corte  $f_{LTPcutoff}$  são reduzidas, a redução sendo dependente do ganho de LTP.

[000141] A figura 17c-iv) ilustra esquematicamente a construção da tabela de deslocamento. A tabela de deslocamento nominal é uma ma-

triz preenchida com números pseudoaleatórios distribuídos entre -0,5 e 0,5. O número de colunas na matriz iguala o número de linhas da MDCT que são codificadas pelo MBMLQ. O número de colunas é ajustável e iguala o número de vetores de deslocamentos que são testados na otimização RD no codificador com restrição de entropia com base em modelo 1740 (veja figura 17e). A função da construção da tabela de deslocamento escalona a tabela de deslocamento nominal com o tamanho da etapa do quantizador, de forma que os deslocamentos sejam distribuídos entre  $-\Delta/2$  e  $+\Delta/2$ .

[000142] A figura 17g ilustra esquematicamente uma modalidade para uma tabela de deslocamento. O índice de deslocamento é um apontador para a tabela e seleciona um vetor de deslocamento escolhido  $O = \{o_1, o_2, \dots, O_n, \dots, O_N\}$ , em que  $N$  é o número de linhas da MDCT no quadro MDCT.

[000143] Da forma descrita a seguir, os deslocamentos proveem um meio para preenchimento de ruído. Melhores objetivo e qualidade perceptiva são obtidos se a dispersão dos deslocamentos for limitada para linhas da MDCT que têm baixa variância  $v_j$ , comparado com o tamanho da etapa do quantizador  $\Delta$ . Um exemplo de uma limitação como esta é descrito na figura 17c-iv), em que  $k_1$  e  $k_2$  são parâmetros de sintonia. A distribuição dos deslocamentos pode ser uniforme e distribuída entre  $-s$  e  $+s$ . Os limites  $s$  podem ser determinados de acordo com

$$s = \begin{cases} k_2 \sqrt{v_j} & \text{se } \sqrt{v_j} < k_1 \Delta \\ \frac{\Delta}{2} & \text{caso contrário} \end{cases}$$

[000144] Para linhas da MDCT de baixa variância (em que  $v_j$  é pequeno, se comparado com  $\Delta$ ), pode ser vantajoso tornar a distribuição de deslocamento não uniforme e dependente do sinal.

[000145] A figura 17e ilustra esquematicamente o codificador com restrição de entropia com base em modelo 1740 com mais detalhes.

As linhas da MDCT inseridas são perceptivamente ponderadas pela sua divisão com os valores da curva de mascaramento perceptiva, preferivelmente, derivados do polinômio LPC, resultando no vetor das linhas da MDCT ponderadas  $y = (y_1, \dots, y_N)$ . O objetivo da subsequente codificação é introduzir ruído de quantização branco nas linhas da MDCT no domínio perceptivo. No decodificador, é aplicado o inverso da ponderação perceptiva, que resulta em ruído de quantização que segue a curva de mascaramento perceptiva.

[000146] Primeiro, a iteração sobre os deslocamentos aleatórios é esboçada. As seguintes operações são realizadas para cada coluna  $j$  na matriz de deslocamento: Cada linha da MDCT é quantizada por um quantizador escalar uniforme deslocamento (USQ), em que cada quantizador é deslocado por seu próprio valor de deslocamento exclusivo tomado do vetor da coluna de deslocamento.

[000147] A probabilidade do mínimo intervalo de distorção de cada USQ é computada no módulo de computações de probabilidade 1770 (veja figura 17g). Os índices do USQ são codificados por entropia. O custo, em termos do número de bits exigidos para codificar os índices, é computado da forma mostrada na figura 17e, produzindo um comprimento de palavra código teórico  $R_j$ . O limite de sobrecarga do USQ

da linha da MDCT  $j$  pode ser computado como  $k_3 \cdot \sqrt{v_j}$ , em que  $k_3$  pode ser escolhido para ser qualquer número apropriado, por exemplo, 20. O limite de sobrecarga é o limiar para que o erro de quantização seja maior que metade do tamanho da etapa de quantização em magnitude.

[000148] Um valor de reconstrução escalar para cada linha da MDCT é computado pelo módulo de dequantização 1780 (veja figura 17h), produzindo o vetor MDCT quantizado  $\bar{y}$ . No módulo de otimização RD 1790, uma distorção  $D_j = d(y, \bar{y})$  é computada.  $d(y, \bar{y})$  pode ser o erro quadrático médio (MSE), ou uma outra medida de distorção percepti-

vamente mais relevante, por exemplo, com base em uma função de ponderação perceptiva. Em particular, uma medida de distorção que pondera MSE e a divergência em energia entre  $y$  e  $\bar{y}$  pode ser usada.

[000149] No módulo de otimização RD 1790, um custo  $C$  é computado, preferivelmente, com base na distorção  $D_j$  e/ou no comprimento de palavra código teórico  $R_j$  para cada coluna  $j$  na matriz de deslocamento. Um exemplo de uma função de custo é  $C = 10 * \log_{10}(D_j) + \lambda * R_j/N$ . O deslocamento que minimiza  $C$  é escolhido, e os correspondentes índices e probabilidades do USQ são transmitidos pelo codificador com restrição de entropia com base em modelo 1780.

[000150] Opcionalmente, a otimização RD pode ser adicionalmente melhorada, pela variação de outras propriedades do quantizador juntamente com o deslocamento. Por exemplo, em vez de usar a mesma estimativa de variância fixa  $V$  para cada vetor de deslocamento que é testado na otimização RD, o vetor de estimativa de variância  $V$  pode variar. Então, para o vetor da coluna de deslocamento  $m$ , será usada uma estimativa de variância  $k_m \cdot V$ , em que  $k_m$  pode abarcar, por exemplo, a faixa de 0,5 a 1,5, já que  $m$  varia de  $m = 1$  a  $m =$  (número de colunas na matriz de deslocamento). Isto torna a codificação por entropia e a computação MMSE menos sensível às variações na estatística do sinal de entrada que o modelo estatístico não pode capturar. Isto resulta em um custo  $C$ , no geral, mais baixo.

[000151] As linhas da MDCT dequantizadas podem ser adicionalmente refinadas pelo uso de um quantizador residual representado na figura 17e. O quantizador residual pode ser, por exemplo, um quantizador vetorial aleatório de taxa fixa.

[000152] A operação do Quantizador Escalar Uniforme (USQ) para quantização de linha  $n$  da MDCT é esquematicamente ilustrada na figura 17f, que mostra o valor da linha da MDCT  $n$  no mínimo intervalo de distorção com índice  $i_n$ . As marcações 'x' indicam o centro (ponto

médio) dos intervalos de quantização com tamanho da etapa  $\Delta$ . A origem do quantizador escalar é deslocada pelo deslocamento  $o_n$  de vetor de deslocamento  $O = \{o_1, o_2, \dots, o_n, \dots, o_N\}$ . Assim, os limites e pontos médios do intervalo são deslocados pelo deslocamento.

[000153] O uso de deslocamentos introduz preenchimento de ruído controlado por codificador no sinal quantizado, e fazendo isto, evita furos espectrais no espectro quantizado. Além do mais, deslocamentos aumentam a codificação eficiência pela provisão de um conjunto de alternativas de codificação que preenchem o espaço mais eficientemente que uma estrutura cúbica. Também, deslocamentos proveem variação nas tabelas de probabilidade que são computadas pelo módulo de computações de probabilidade 1770, que leva a codificação por entropia mais eficiente dos índices das linhas da MDCT (isto é, menos bits exigidos).

[000154] O uso de um tamanho da etapa variável  $\Delta$ (delta) permite precisão variável na quantização, de forma que mais precisão possa ser usada para sons perceptivamente importantes, e menos precisão possa ser usada para sons menos importantes.

[000155] A figura 17g ilustra esquematicamente as computações de probabilidade no módulo de computação de probabilidade 1770. As entradas deste módulo são o modelo estatístico aplicado para as linhas da MDCT, o tamanho da etapa do quantizador  $\Delta$ , o vetor de variância  $V$ , o índice de deslocamento e a tabela de deslocamento. A saída do módulo de computação de probabilidade 1770 são tabelas cdf. Para cada linha  $x_j$  da MDCT, o modelo estatístico (isto, é uma função de densidade de probabilidade, pdf) é avaliado. A área sob a função pdf para um intervalo  $i$  é a probabilidade  $p_{i,j}$  do intervalo. Esta probabilidade é usada para a codificação aritmética das linhas da MDCT.

[000156] A figura 17h ilustra esquematicamente o processo de dequantização realizado, por exemplo, no módulo de dequantização

1780. O centro da massa (valor MMSE)  $X_{MMSE}$  para o mínimo intervalo de distorção de cada linha da MDCT é computado juntamente com o ponto médio  $X_{MP}$  do intervalo. Considerando que um vetor  $N$ -dimensional das linhas da MDCT é quantizado, o valor MMSE escalar é subideal e, no geral, muito baixo. Isto resulta em uma perda de variância e desequilíbrio espectral na saída decodificada. Este problema pode ser mitigado pela decodificação de conservação da variância descrita na figura 17h, em que o valor de reconstrução é computado como uma soma ponderada do valor MMSE e do valor do ponto médio. Uma melhoria opcional adicional é adaptar a ponderação, de forma que o valor MMSE domine para fala e o ponto médio domine para sons que não fala. Isto produz fala, enquanto equilíbrio espectral e energia são conservados para sons que não fala.

[000157] Decodificação com conservação de variância de acordo com uma modalidade da invenção é alcançada pela determinação do ponto de reconstrução de acordo com a seguinte equação:

$$x_{dequant} = (1 - \chi)x_{MMSE} + x_{MP}$$

[000158] Decodificação com conservação de variância adaptativa pode ser baseada na seguinte regra para determinação do fator de interpolação:

$$\chi = \begin{cases} 0 & \text{se sons de fala} \\ 1 & \text{se sons de não fala} \end{cases}$$

[000159] Por exemplo, a ponderação adaptativa pode ser adicionalmente em função do ganho de predição LTP  $g_{LTP}$ :  $\chi = f(g_{LTP})$ . A ponderação adaptativa varia lentamente e pode ser eficientemente codificada por um código de entropia recursivo.

[000160] O modelo estatístico das linhas da MDCT que é usado nas computações de probabilidade (figura 17g) e na dequantização (figura 17h) deve refletir a estatística do sinal real. Em uma versão, o modelo estatístico considera que as linhas da MDCT são independentes e dis-

tribuídas por modelo Laplaciano. Uma outra versão modela as linhas da MDCT como modelos Gaussianos independentes. Uma versão modela as linhas da MDCT como modelos de mistura Gaussiana, incluindo interdependências entre as linhas da MDCT nos quadros MDCT e entre eles. Uma outra versão adapta o modelo estatístico à estatística de sinal em linha. Os modelos estatísticos adaptativos podem ser adaptados para frente e/ou para trás.

[000161] Um outro aspecto da invenção que refere-se aos pontos de reconstrução modificados do quantizador é esquematicamente ilustrado na figura 19, em que é representado um quantizador invertido usado no decodificador de uma modalidade. O módulo também tem, além das entradas normais de um quantizador invertido, isto é, as linhas quantizadas e a informação sobre o tamanho da etapa de quantização (tipo de quantização), informação sobre o ponto de reconstrução do quantizador. O quantizador invertido desta modalidade pode usar múltiplos tipos de pontos de reconstrução durante a determinação de um valor reconstruído  $\bar{y}_n$  do índice de quantização correspondente  $i_n$ . Como exposto, valores de reconstrução  $\bar{y}$  são adicionalmente usados, por exemplo, no codificador das linhas da MDCT (veja figura 17) para determinar a quantização residual para entrada no quantizador residual. Além do mais, reconstrução da quantização é realizada no quantizador invertido 304, para reconstrução de um quadro MDCT codificado para uso no armazenamento temporário LTP (veja figura 3) e, naturalmente, no decodificador.

[000162] Por exemplo, o quantizador invertido pode escolher o ponto médio de um intervalo de quantização como o ponto de reconstrução, ou o ponto de reconstrução MMSE. Em uma modalidade da presente invenção, o ponto de reconstrução do quantizador é escolhido para ser o valor médio entre o centre e os pontos de reconstrução MMSE. No geral, o ponto de reconstrução pode ser interpolado entre o ponto mé-

dio e o ponto de reconstrução MMSE, por exemplo, dependendo de propriedades do sinal, tal como periodicidade do sinal. Informação de periodicidade do sinal pode ser derivada do módulo LTP, por exemplo. Este recurso permite que o sistema controle a distorção e a conservação de energia. O ponto de reconstrução central garantirá a conservação de energia, enquanto que o ponto de reconstrução MMSE garantirá mínima distorção. Então, dado o sinal, o sistema pode adaptar o ponto de reconstrução para onde o melhor acerto for provido.

[000163] A presente invenção incorpora adicionalmente um novo formato de codificação da sequência de janelas. De acordo com uma modalidade da invenção, as janelas usadas para a transformada MDCT são de tamanhos diádicos, e pode variar apenas um fator dois de tamanho, de janela para janela. Por exemplo, tamanhos de transformada diádicos são 64, 128, ..., 2048 amostras correspondentes a 4, 8, ..., 128 ms em taxa de amostragem de 16 kHz. No geral, são propostas janelas de tamanho variável que podem tomar uma pluralidade de tamanhos de janela entre um tamanho de janela mínimo e um tamanho máximo. Em uma sequência, tamanhos de janela consecutivos podem variar apenas em um fator de dois, de forma que suaves sequências de tamanhos de janela se desenvolvam sem mudanças abruptas. As sequências de janelas definidas por uma modalidade, isto é, limitadas a tamanhos diádicos e permitidas variar apenas em fator dois em tamanho, de janela para janela, tem diversas vantagens. Primeiramente, nenhuma janela de início ou fim específica é necessária, isto é, janelas com bordas agudas. Isto mantém uma boa resolução de tempo/frequência. Em segundo lugar, a sequência de janelas se torna muito eficiente para codificar, isto é, para sinalizar para um decodificador qual sequência de janelas em particular é usada. Finalmente, a sequência de janelas sempre se adequará habilmente em uma estrutura de hiperquadro.

[000164] A estrutura de hiperquadro é usada durante a operação do codificador em um sistema do mundo real, em que certos parâmetros de configuração do decodificador precisam ser transmitidos, a fim de poder iniciar o decodificador. Estes dados são comumente armazenados em um campo de cabeçalho no fluxo de bits que descreve o sinal de áudio codificado. A fim de minimizar taxa de bits, o cabeçalho não é transmitido para cada quadro dos dados codificados, particularmente, em um sistema proposto pela presente invenção, em que os tamanhos do quadro MDCT podem variar de muito pequenos a muito grandes. Portanto, a presente invenção propõe agrupar uma certa quantidade de quadros MDCT em um hiperquadro, em que os dados do cabeçalho são transmitidos no começo do hiperquadro. Tipicamente, o hiperquadro é definido com uma duração de tempo específica. Portanto, é necessário tomar cuidado, de forma que as variações dos tamanhos do quadro MDCT se adequem em uma duração de hiperquadro constante pré-definida. A sequência de janelas inventiva exposta garante que a sequência de janelas selecionada sempre se adeque em uma estrutura de hiperquadro.

[000165] De acordo com uma modalidade da presente invenção, o atraso LTP e o ganho de LTP são codificados de uma maneira de taxa variável. Isto é vantajoso uma vez que, em função da efetividade da LTP para sinais periódicos estacionários, o atraso LTP tende a ser o mesmo durante segmentos um tanto longos. Portanto, isto pode ser explorado por meio de codificação aritmética, resultando em um atraso LTP de taxa variável e codificação de ganho de LTP.

[000166] Similarmente, uma modalidade da presente invenção também tira vantagem de um reservatório de bits e de codificação com taxa variável para a codificação dos parâmetros da LP. Além do mais, codificação da LP recursiva é preceituada pela presente invenção.

[000167] Um outro aspecto da presente invenção é o tratamento de

um reservatório de bits para tamanhos de quadro variáveis no codificador. Na figura 18, é esboçada uma unidade de controle do reservatório de bits 1800 de acordo com a presente invenção. Além de uma medida de dificuldade provida como entrada, a unidade de controle do reservatório de bits também recebe informação sobre a duração do quadro atual. Um exemplo de uma medida de dificuldade para uso na unidade de controle do reservatório de bits é entropia perceptiva, ou o logaritmo do espectro de potência. Controle do reservatório de bits é importante em um sistema em que as durações do quadro podem variar durante um conjunto de diferentes durações do quadro. A unidade de controle do reservatório de bits sugerida 1800 considera o tamanho de quadro durante o cálculo do número de bits concedidos para que o quadro seja codificado, como será esboçado a seguir.

[000168] O reservatório de bits é aqui definido com uma certa quantidade fixa de bits em um armazenamento temporário que precisa ser maior que o número médio de bits que um quadro é permitido usar para uma dada taxa de bits. Se ele for do mesmo tamanho, nenhuma variação no número de bits para um quadro será possível. O controle do reservatório de bits sempre considera o nível do reservatório de bits antes de eliminar bits que serão concedidos ao algoritmo de codificação como número permitido de bits para o quadro real. Assim, um reservatório de bits cheio significa que o número de bits disponíveis no reservatório de bits iguala o tamanho do reservatório de bits. Depois da codificação do quadro, o número de bits usados será subtraído do armazenamento temporário e o reservatório de bits é atualizado pela adição do número de bits que representa a taxa de bits constante. Portanto, o reservatório de bits fica vazio, se o número de bits no reservatório de bits antes da codificação de um quadro for igual ao número médio de bits por quadro.

[000169] Na figura 18a, é representado o conceito básico do controle

do reservatório de bits. O codificador provê meio para calcular quão difícil é codificar o quadro real, comparado com o quadro prévio. Para uma dificuldade média de 1,0, o número de bits concedidos depende do número de bits disponíveis no reservatório de bits. De acordo com uma dada linha de controle, mais bits que correspondente a uma taxa média de bits serão eliminados do reservatório de bits, se o reservatório de bits estiver quase cheio. No caso de um reservatório de bits vazio, menos bits, comparado com a média de bits, serão usados para a codificação do quadro. Este comportamento produz um nível médio do reservatório de bits para uma maior sequência de quadros com dificuldade média. Para quadros com uma maior dificuldade, a linha de controle pode ser deslocada para cima, com o efeito que permite-se que quadros difíceis de codificar usem mais bits no mesmo nível do reservatório de bits. Desta maneira, para quadros fáceis de codificar, o número de bits permitido para um quadro será menor apenas deslocando para baixo a linha de controle da figura 18a do caso de dificuldade média para o caso de dificuldade fácil. Modificações diferentes do simples deslocamento da linha de controle também são possíveis. Por exemplo, da forma mostrada na figura 18a, a inclinação da curva de controle pode mudar, dependendo da dificuldade do quadro.

[000170] Durante o cálculo do número de bits concedidos, os limites na extremidade inferior do reservatório de bits precisam ser obedecidos a fim de não eliminar mais bits do armazenamento temporário do que o permitido. Um esquema de controle do reservatório de bits que inclui o cálculo dos bits concedidos por uma linha de controle, mostrado na figura 18a, é apenas um exemplo das possíveis relações do nível do reservatório de bits e da medida de dificuldade pelos bits concedidos. Também, outros algoritmos de controle terão em comum os severos limites na extremidade inferior do nível do reservatório de bits que impede que um reservatório de bits viole a restrição do reservató-

rio de bits vazio, também os limites na extremidade superior, em que o codificador será forçado a gravar bits preenchidos, se um número muito baixo de bits for consumido pelo codificador.

[000171] Para que um mecanismo de controle como este possa tratar um conjunto de tamanhos de quadro variáveis, este simples algoritmo de controle precisa ser adaptado. A medida de dificuldade a ser usada precisa ser normalizada, de forma que os valores de dificuldade de quadros de diferentes durações sejam comparáveis. Para cada duração do quadro, haverá uma diferente faixa permitida para os bits concedidos e, em virtude de o número médio de bits por quadro ser diferente para um tamanho de quadro variável, conseqüentemente, cada duração do quadro tem sua própria equação de controle com suas próprias limitações. Um exemplo é mostrado na figura 18b. Uma importante modificação do caso do quadro de duração fixa é o limite mais baixo permitido do algoritmo de controle. Em vez do número médio de bits para o tamanho do quadro real, que corresponde ao caso da taxa fixa de bits, agora, o número médio de bits para o maior tamanho de quadro permitido é o valor mais baixo permitido para o nível do reservatório de bits antes de eliminar os bits para o quadro real. Esta é uma das principais diferenças em relação ao controle do reservatório de bits para tamanhos de quadro fixos. Esta restrição garante que um quadro seguinte com o maior tamanho possível do quadro possa utilizar pelo menos o número médio de bits para este tamanho do quadro.

[000172] A medida de dificuldade pode ser baseada, por exemplo, em um cálculo de entropia perceptiva (PE) que é derivado dos limites de mascaramento de um modelo psicoacústico, como é feito em AAC, ou como uma alternativa à contagem de bits de uma quantização com tamanho da etapa fixa, como é feito na parte ECQ de um codificador de acordo com uma modalidade da presente invenção. Estes valores podem ser normalizados em relação aos tamanhos de quadro variá-

veis, o que pode ser realizado por uma simples divisão pelo tamanho do quadro, e o resultado será uma PE respectivamente a uma contagem de bits por amostra. Uma outra etapa de normalização pode ocorrer em relação à dificuldade média. Com este propósito, uma média móvel em relação aos quadros passados pode ser usada, resultando em um valor de dificuldade maior que 1,0 para quadros difíceis ou menor que 1,0 para quadros fáceis. No caso de um codificador de dois passos ou de uma maior antecipação, valores de dificuldade de quadros futuros também podem ser levados em consideração para esta normalização da medida de dificuldade.

[000173] Um outro aspecto da invenção refere-se a especificações do tratamento do reservatório de bits para ECQ. O gerenciamento do reservatório de bits para ECQ funciona sob a consideração de que ECQ produz uma qualidade aproximadamente constante durante o uso de um tamanho da etapa do quantizador constante para codificação. Tamanho da etapa do quantizador constante produz uma taxa variável, e o objetivo do reservatório de bits é manter a variação no tamanho da etapa do quantizador entre diferentes quadros tão pequena quanto possível, ainda não violando as restrições do armazenamento temporário do reservatório de bits. Além da taxa produzida pelo ECQ, informação adicional (por exemplo, ganho e atraso da LTP) é transmitida com base no quadro da MDCT. No geral, a informação adicional também é codificada por entropia e, assim, consome taxa diferente de quadro para quadro.

[000174] Em uma modalidade da invenção, um controle do reservatório de bits proposto tenta minimizar a variação do tamanho da etapa ECQ pela introdução de três variáveis (veja figura 18c):

- $R_{ECQ\_AVG}$ : Taxa média do ECQ por amostra previamente usada;
- $\Delta_{ECQ\_AVG}$ : Tamanho médio da etapa do quantizador previ-

amente usada.

[000175] Ambas estas variáveis são dinamicamente atualizadas para refletir a estatística de codificação mais recente.

-  $R_{ECQ\_AVG\_DES}$ : A taxa ECQ correspondente à taxa de bits total média.

[000176] Este valor diferirá de  $R_{ECQ\_AVG}$  no caso de o nível do reservatório de bits mudar durante o quadro de tempo da janela do cálculo de média, por exemplo, uma taxa de bits maior ou menor que a taxa média de bits especificada foi usada durante este quadro de tempo. Ele também é atualizado à medida que a taxa da informação complementar muda, de forma que a taxa total iguale a taxa de bits especificada.

[000177] O controle do reservatório de bits usa estes três valores para determinar uma suposição inicial no delta a ser usado para o quadro atual. Ele faz isto encontrando  $\Delta_{ECQ\_AVG\_DES}$  na curva  $R_{ECQ}-\Delta$  mostrada na figura 18c, que corresponde a  $R_{ECQ\_AVG\_DES}$ . Em um segundo estágio, possivelmente, este valor é modificado se a taxa não estiver de acordo com as restrições do reservatório de bits. A curva  $R_{ECQ}-\Delta$  exemplar na figura 18c é baseada na seguinte equação:

$$R_{ECQ} = \frac{1}{2} \log_2 \frac{\alpha}{\Delta^2}$$

[000178] Certamente, outros relacionamentos matemáticos entre  $R_{ECQ}$  e  $\Delta$  também podem ser usados.

[000179] No caso estacionário,  $R_{ECQ\_AVG}$  ficará próximo de  $R_{ECQ\_AVG\_DES}$ , e a variação em  $\Delta$  será muito pequena. No caso não estacionário, a operação de cálculo de média garantirá uma suave variação de  $\Delta$ .

[000180] Embora o exposto tenha sido divulgado em relação a modalidades em particular da presente invenção, entende-se que o conceito inventivo não é limitado às modalidades descritas. Por outro lado,

a divulgação apresentada neste pedido habilitará versados na técnica a entender e realizar a invenção. Versados na técnica entendem que várias modificações podem ser feitas sem fugir do espírito e do escopo da invenção, da forma exposta exclusivamente pelas concretizações.

## REIVINDICAÇÕES

1. Sistema de codificação de áudio, **caracterizado pelo fato de que** compreende:

uma unidade de predição linear (LP) (201) para filtrar um sinal de áudio com base em um filtro LP, a unidade LP operando em um primeiro tamanho de quadro do sinal de áudio;

uma unidade de transformação de comprimento adaptativa (202), para transformar um quadro do sinal de áudio em um domínio transformado, a transformada sendo uma Transformada de Cosseno Discreta Modificada (MDCT) que opera em um segundo tamanho de quadro variável;

uma unidade de quantização (203) para quantizar um sinal do domínio MDCT;

uma unidade de geração da curva de ganho (1470) para gerar curvas de ganho no domínio MDCT com base em respostas de magnitude do filtro LP; e

uma unidade de mapeamento (1500) para mapear parâmetros LP para quadros correspondentes do sinal de domínio MDCT; curvas de ganho no domínio MDCT para correspondentes quadros do sinal do domínio MDCT; em que a unidade de mapeamento (1500) interpola polinômios de LP gerados em uma taxa correspondente ao primeiro comprimento de quadro, de modo a corresponder aos quadros do sinal do domínio MDCT gerado em uma taxa correspondente ao segundo comprimento de quadro para uso como pesos perceptivos na unidade de quantização (203).

2. Sistema de codificação de áudio, de acordo com a reivindicação 1, **caracterizado pelo fato de que** compreende:

uma unidade de controle de sequência de janelas, para determinar, para um bloco do sinal de áudio, as segundas durações do quadro para sobrepor janelas MDCT.

3. Sistema de codificação de áudio, de acordo com a reivindicação 1 ou 2, **caracterizado pelo fato de que** compreende uma unidade de modelagem perceptiva que modifica uma característica do filtro LP pelo chiado e/ou oscilação de um polinômio LPC gerado pela unidade de predição linear para um quadro LPC.

4. Sistema de codificação de áudio, de acordo com qualquer uma das reivindicações 1 a 3, **caracterizado pelo fato de que** compreende:

uma unidade de divisão de frequência para dividir o sinal de áudio em um componente de banda baixa e um componente de banda alta; e

um codificador de banda alta para codificar o componente de banda alta,

em que o componente de banda baixa é inserido na unidade de predição linear e na unidade de transformação.

5. Sistema de codificação de áudio, de acordo com a reivindicação 4, **caracterizado pelo fato de que** a unidade de divisão de frequência compreende um banco de filtro especular na quadratura e uma unidade de síntese de filtro especular na quadratura, configurada para fazer amostragem descendente do sinal de áudio.

6. Sistema de codificação de áudio, de acordo com a reivindicação 4 ou 5, **caracterizado pelo fato de que** o limite entre a banda baixa e a banda alta é variável e a unidade de divisão de frequência determina a frequência de cruzamento com base em propriedades do sinal de áudio e/ou em exigências de largura de banda do codificador.

7. Sistema de codificação de áudio, de acordo com qualquer uma das reivindicações 4 a 6, **caracterizado pelo fato de que** o codificador de banda alta é um codificador de replicação de banda espectral.

8. Sistema de codificação de áudio, de acordo com qualquer uma das reivindicações 1 a 7, **caracterizado pelo fato de que** compreende:

uma unidade de estimação do fator de escala (1360) para estimar fatores de escala para controlar o ruído de quantização da unidade de quantização (203).

9. Sistema de codificação de áudio, de acordo com a reivindicação 8, **caracterizado pelo fato de que** os fatores de escala são determinados com base nas curvas de ganho no domínio MDCT mapeadas.

10. Sistema de codificação de áudio, de acordo com qualquer uma das reivindicações 1 a 9, **caracterizado pelo fato de que** compreende uma unidade estéreo paramétrica para calcular uma representação estéreo paramétrica dos canais esquerdo e direito inseridos.

11. Decodificador de áudio, **caracterizado pelo fato de que** compreende:

uma unidade de dequantização (211) para reconstruir linhas de MDCT quantizadas recebidas em um fluxo de bits de entrada e gerar um sinal do domínio transformado;

uma unidade de transformação MDCT invertida de comprimento adaptativo (212) para transformar inversamente um sinal do domínio transformado em um sinal no domínio do tempo, a unidade de transformação MDCT invertida (212) operando em um comprimento de quadro variável;

uma unidade de geração da curva de ganho (1470) para gerar curvas de ganho no domínio MDCT com base em respostas de magnitude dos filtros de predição linear, em que parâmetros para os filtros de predição linear são recebidos no fluxo de bits e correspondem a um primeiro comprimento de quadro; e

uma unidade de mapeamento (1500) para mapear parâme-

tros de LP para correspondentes quadros do sinal do domínio MDCT; em que a unidade de mapeamento (1500) interpola polinômios de LP gerados em uma taxa correspondente ao primeiro comprimento de quadro, de modo a corresponder aos quadros do sinal do domínio de transformada gerado em uma taxa correspondente ao segundo comprimento de quadro para uso como pesos perceptivos na unidade de dequantização (211).

12. Método de codificação de áudio, **caracterizado pelo fato de que** compreende as etapas de:

realizar uma análise de predição linear (LP) em um sinal de áudio, a análise LP operando em um primeiro comprimento de quadro e gerando parâmetros de filtro da LP;

transformar um quadro do sinal de áudio em um domínio da Transformada de Cosseno Discreta Modificada (MDCT), a MDCT operando em um segundo comprimento de quadro variável;

quantizar um sinal do domínio MDCT;

gerar curvas de ganho no domínio MDCT com base em respostas de magnitude dos filtros LP gerados; e

mapear parâmetros de LP para correspondentes quadros do sinal do domínio MDCT; em que o mapeamento compreende interpolar polinômios de LP gerados em uma taxa correspondente ao primeiro comprimento de quadro, de modo a corresponder aos quadros do sinal do domínio MDCT gerado em uma taxa correspondente ao segundo comprimento de quadro para uso como pesos perceptivos para quantizar o sinal do domínio MDCT.

13. Método de decodificação de áudio, **caracterizado pelo fato de que** compreende as etapas de:

reconstruir linhas MDCT quantizadas recebidas em um fluxo de bits de entrada e gerar um sinal do domínio transformado;

fazer transformada MDCT invertida de um sinal do domínio

transformado em um sinal no domínio do tempo, a MDCT invertida operando em um comprimento de quadro variável;

gerar curvas de ganho no domínio MDCT com base em respostas de magnitude de filtros de predição linear, em que parâmetros para os filtros de predição linear são recebidos no fluxo de bits e correspondem a um primeiro comprimento de quadro; e

mapear parâmetros de LP para correspondentes quadros do sinal do domínio MDCT; em que o mapeamento compreende interpolar polinômios de LP gerados em uma taxa correspondente ao primeiro comprimento de quadro, de modo a corresponder aos quadros do sinal do domínio MDCT gerado em uma taxa correspondente ao segundo comprimento de quadro para uso como pesos perceptivos para dequantizar o sinal do domínio MDCT.

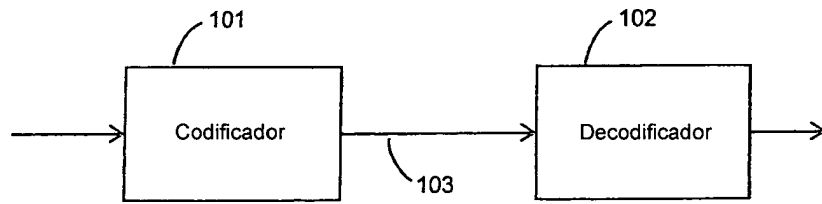


Fig. 1

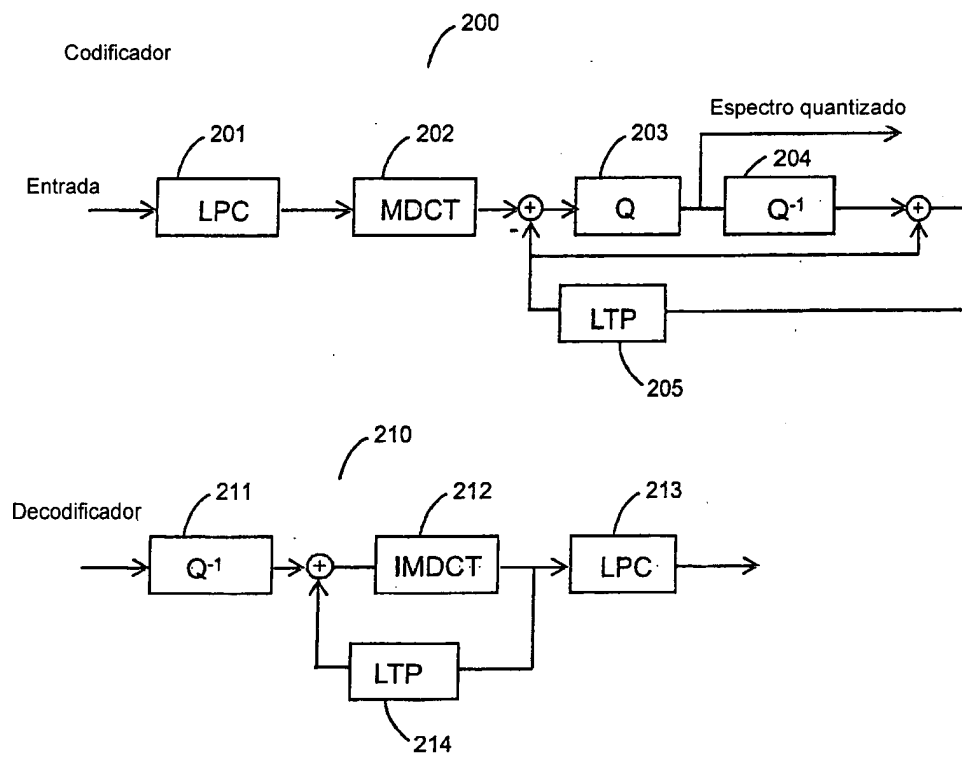


Fig. 2

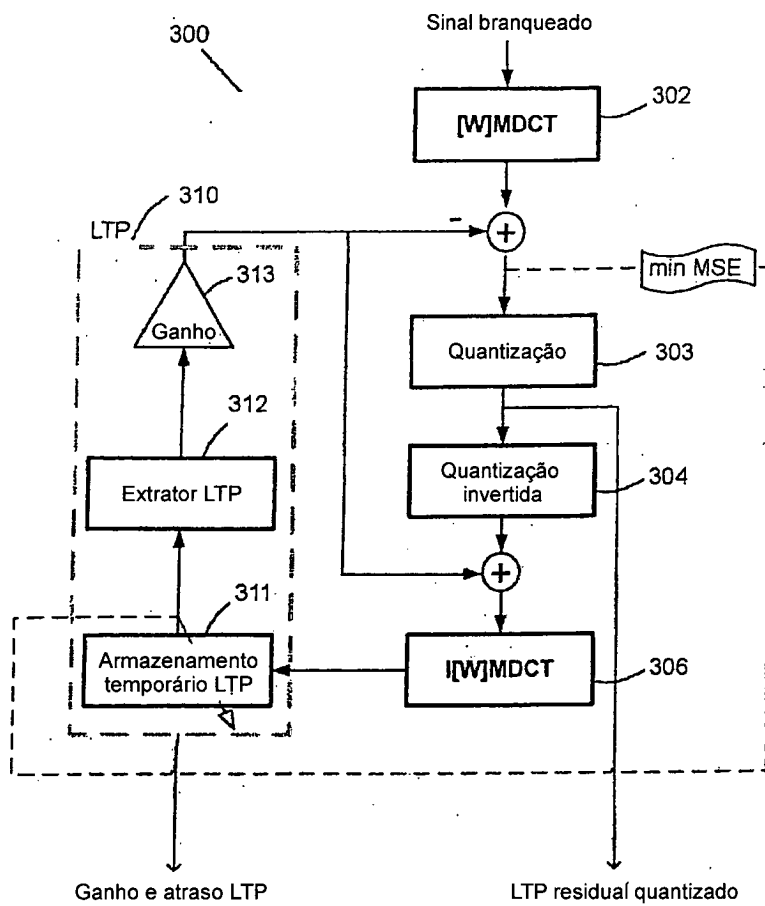


Fig. 3

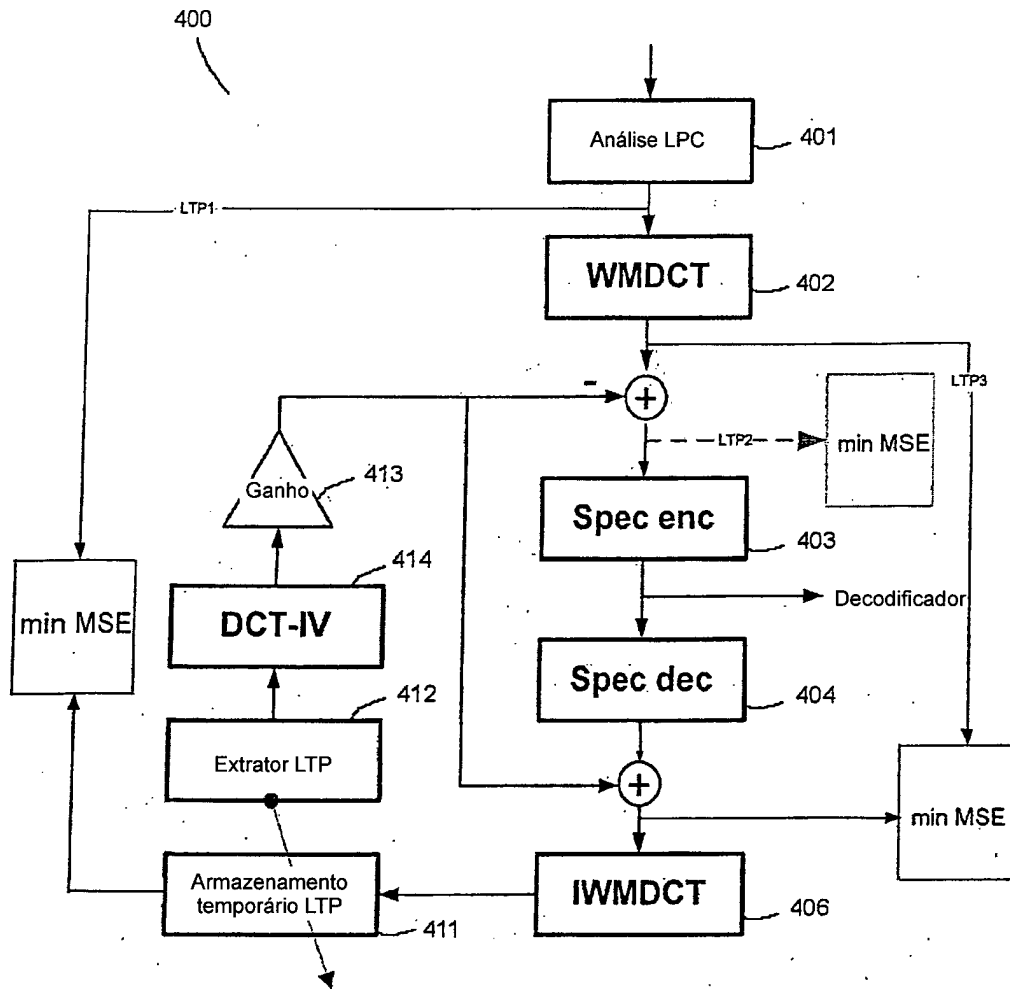


Fig. 4

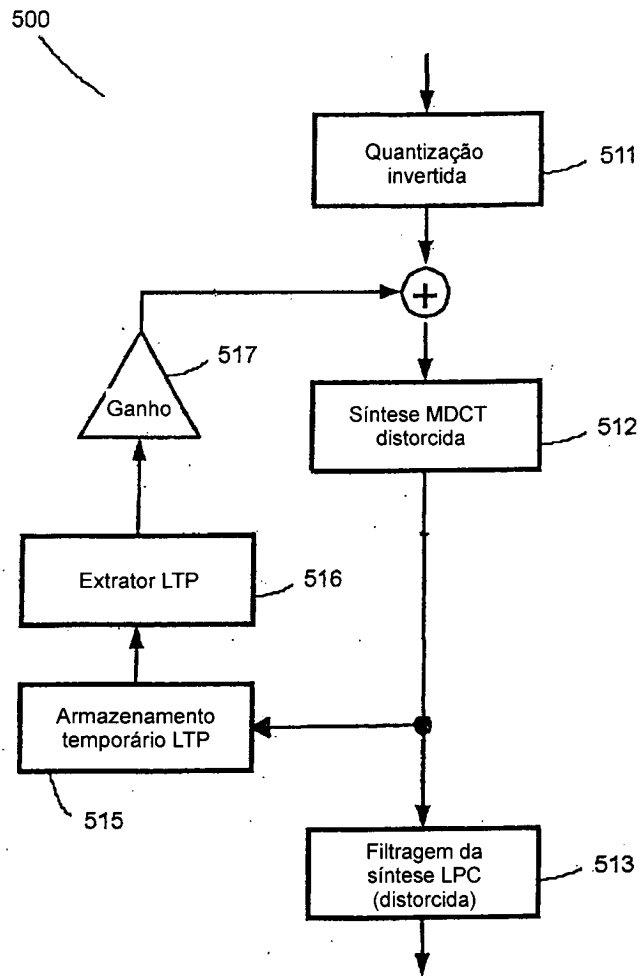


Fig. 5

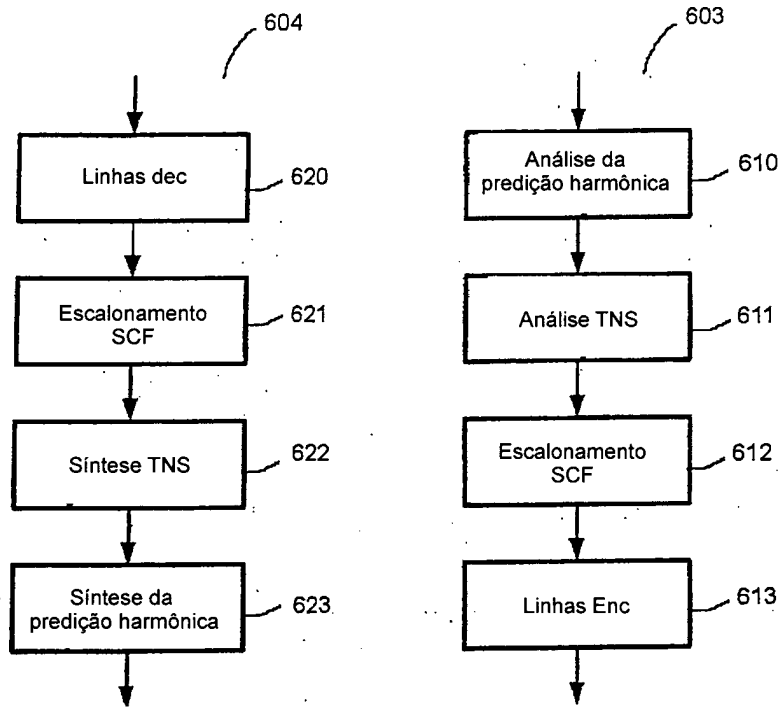


Fig. 6

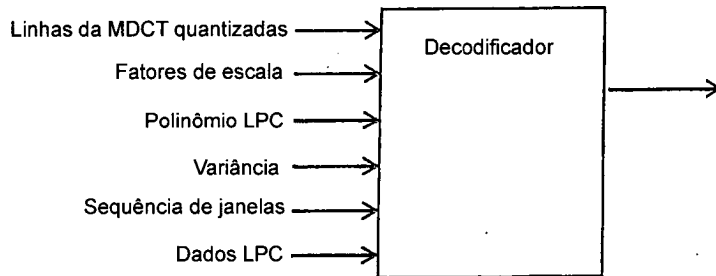
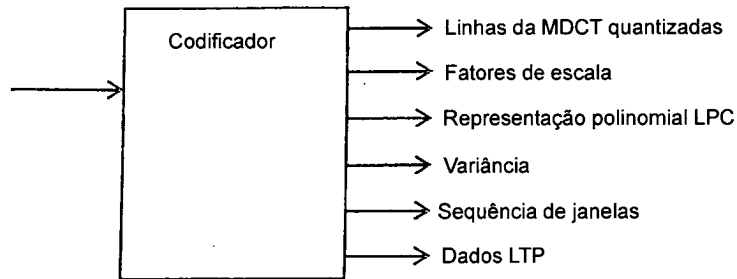


Fig. 7

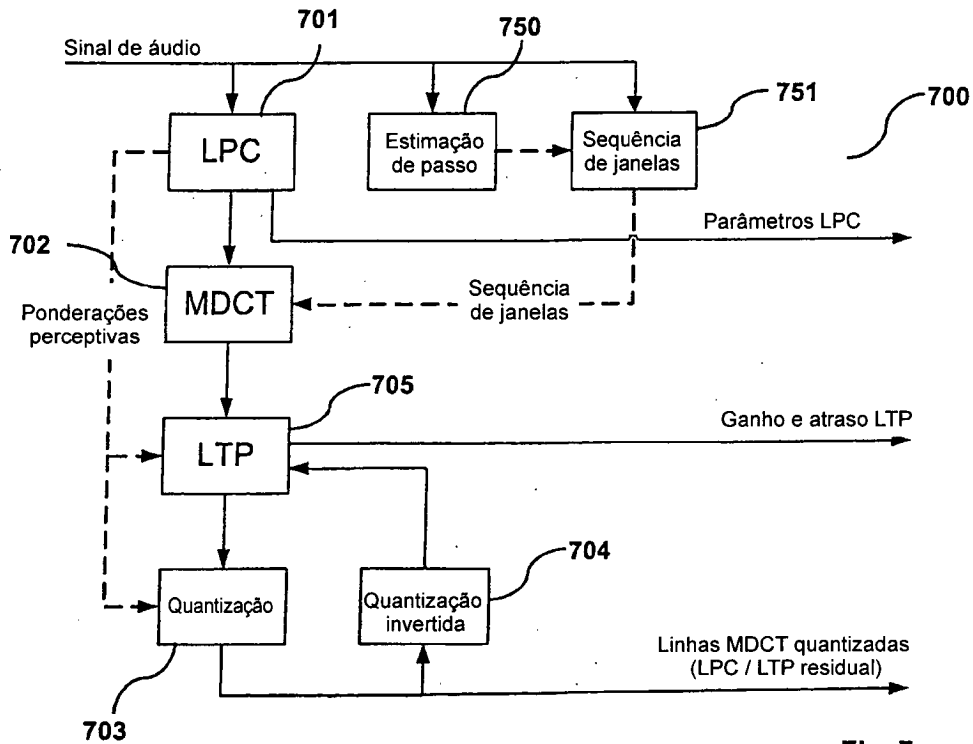


Fig. 7a

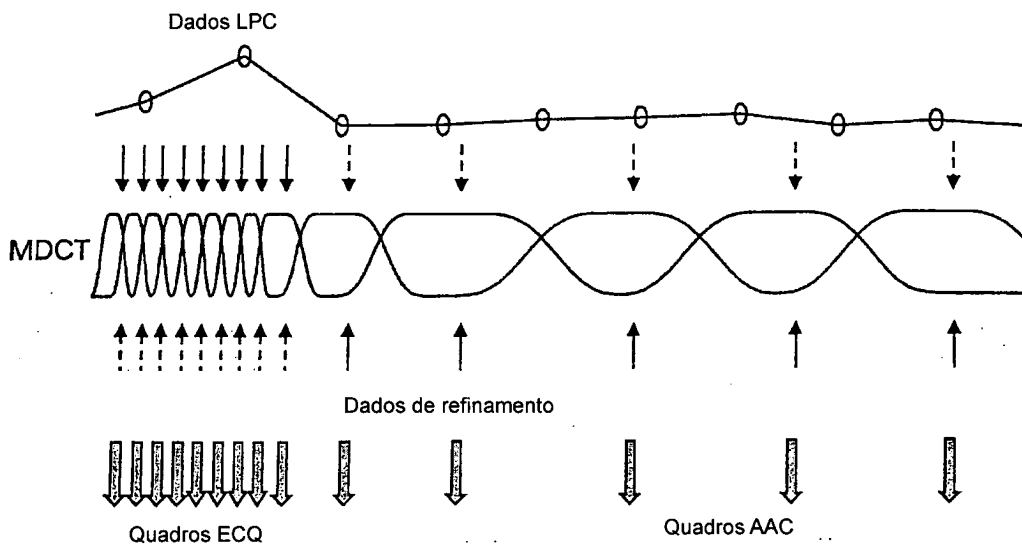


Fig. 8

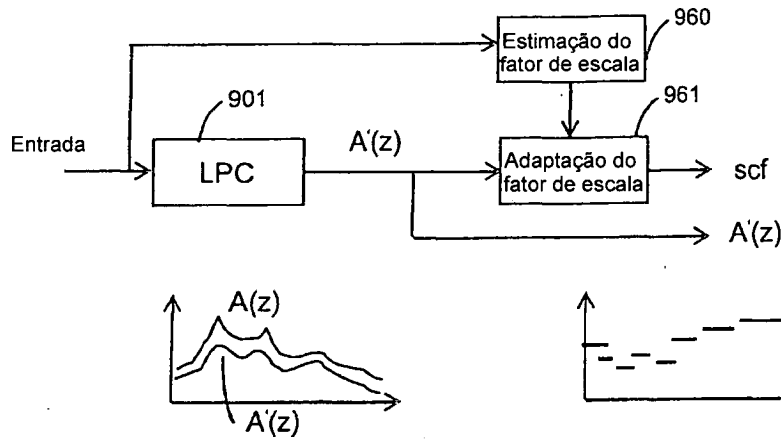


Fig. 9

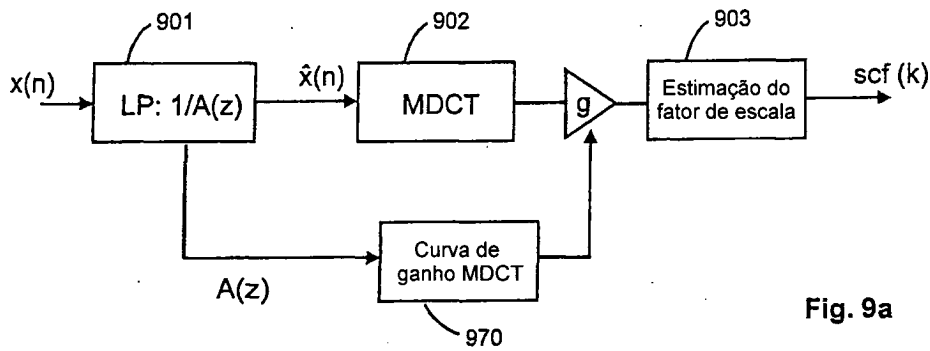


Fig. 9a

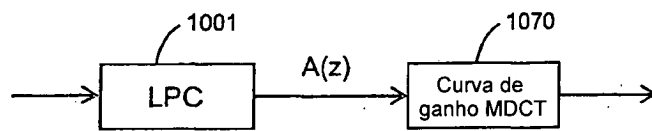


Fig. 10

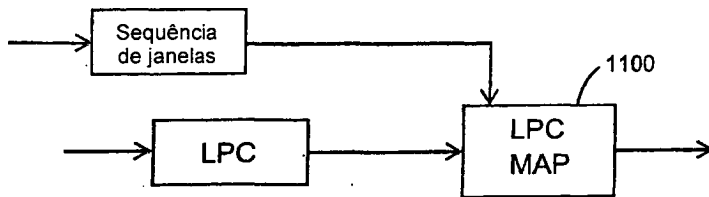


Fig. 11

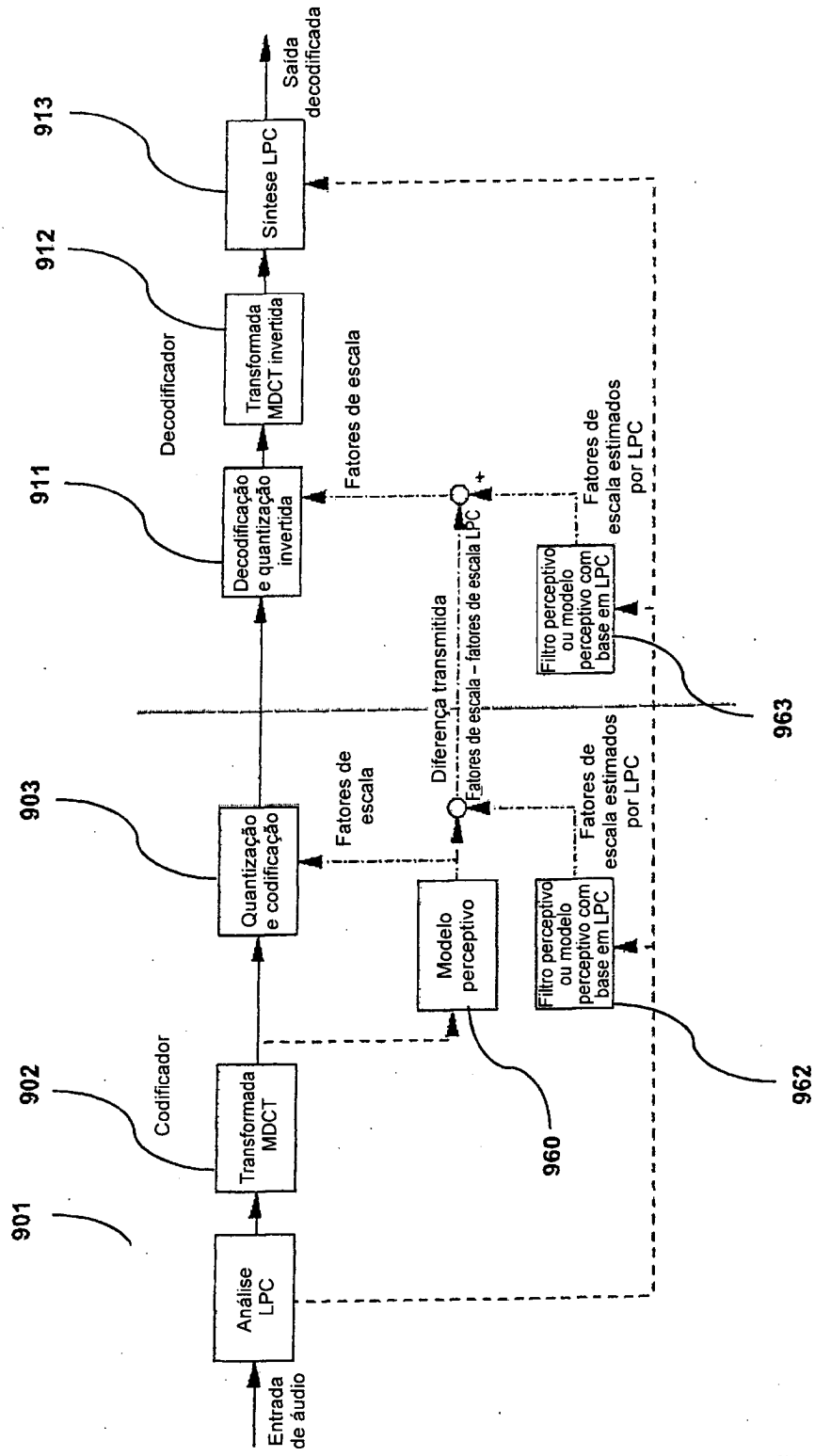


Fig. 9b

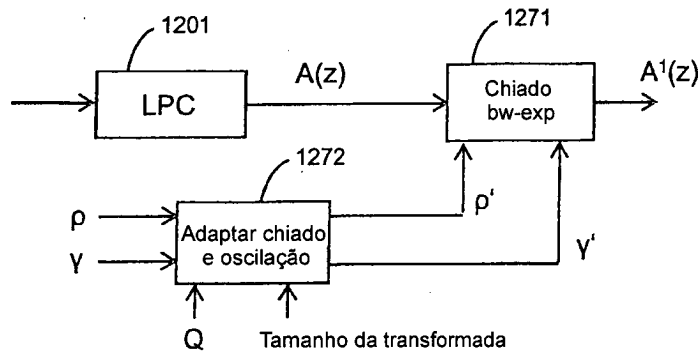


Fig. 12

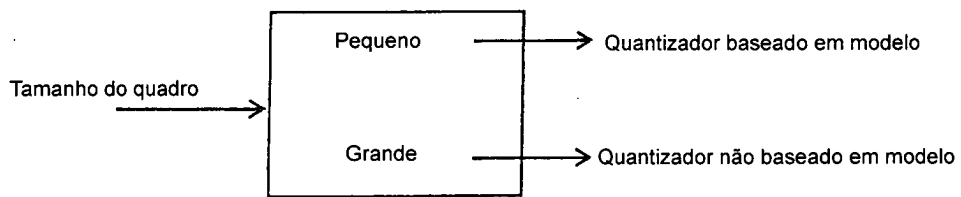


Fig. 13

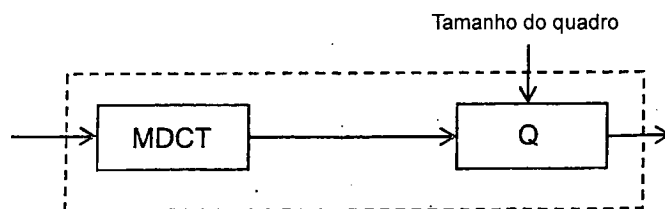


Fig. 14

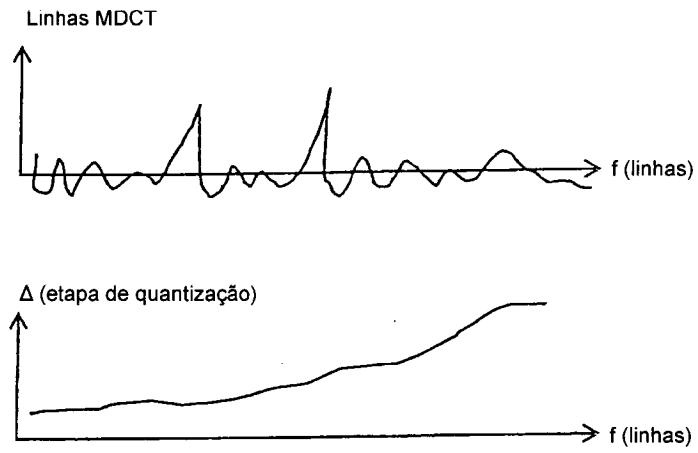


Fig. 15

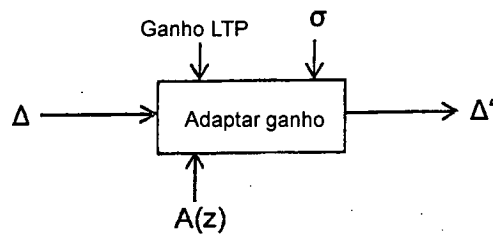


Fig. 15a

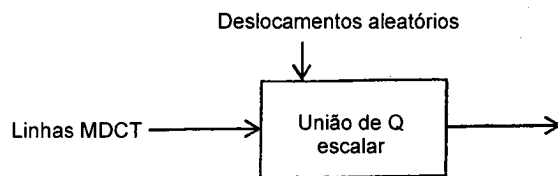


Fig. 16

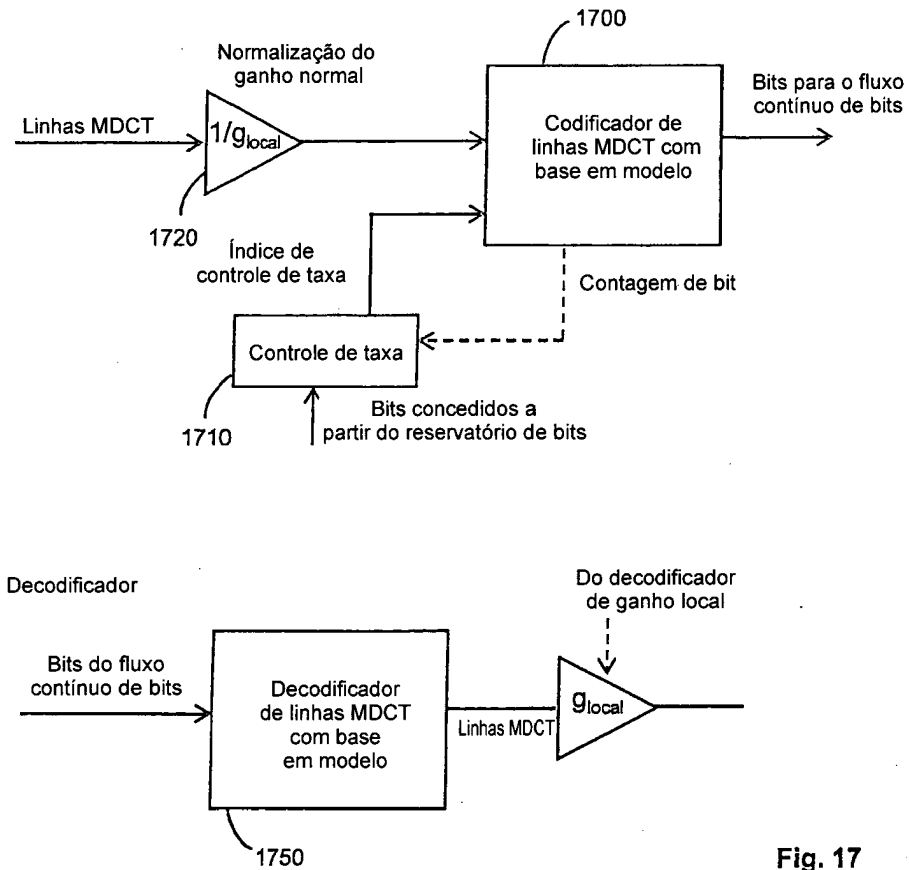


Fig. 17

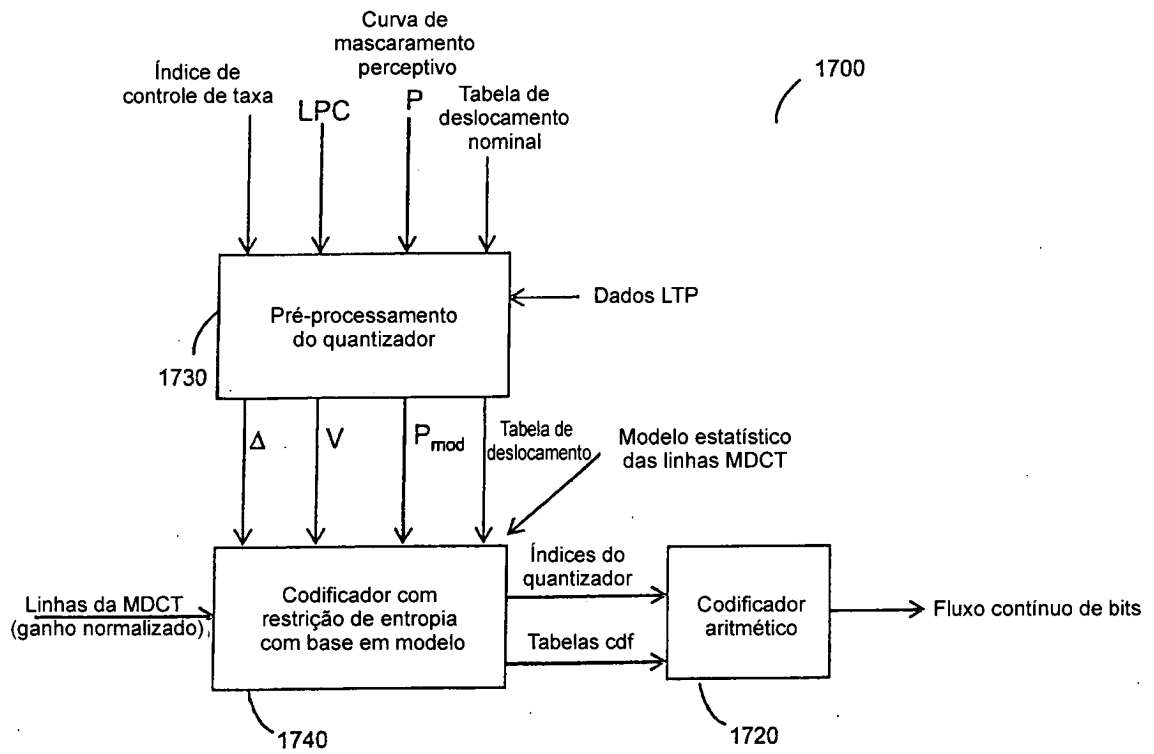


Fig. 17a

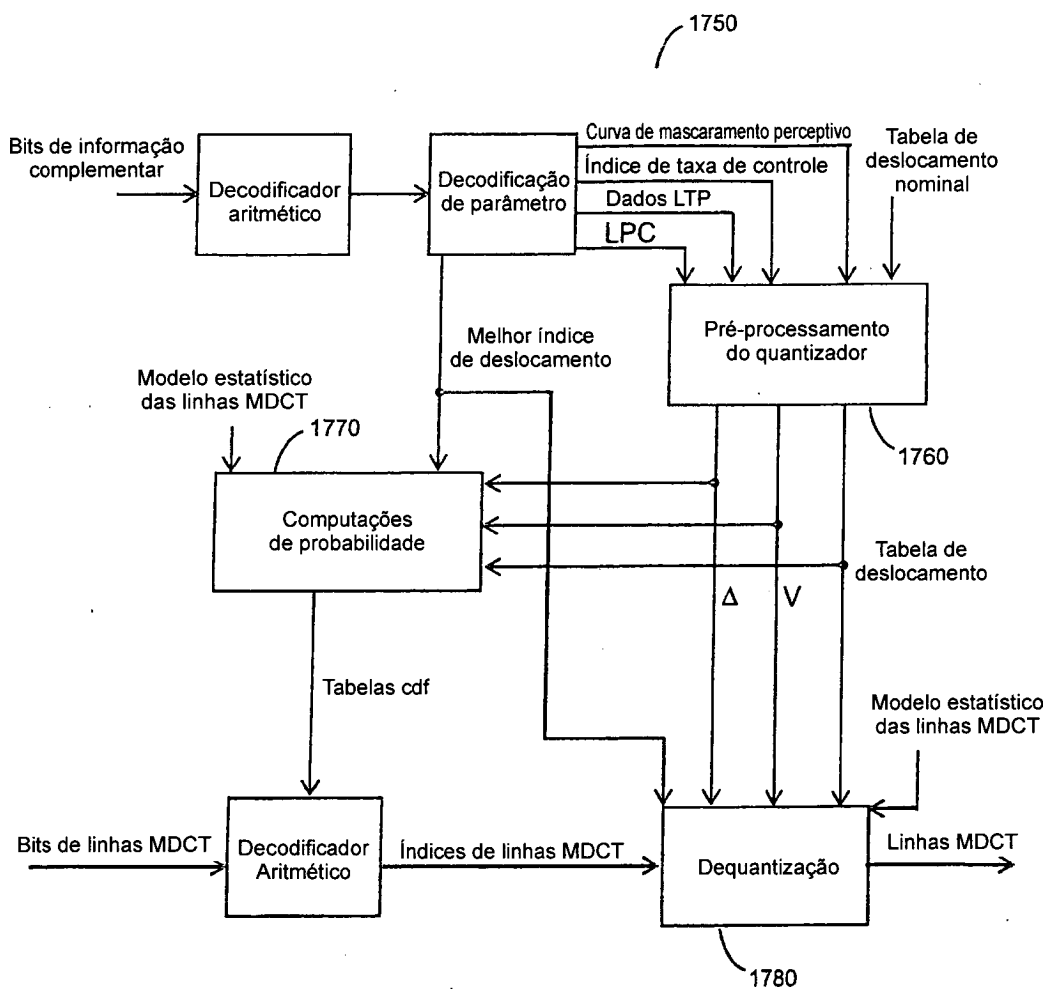


Fig. 17b

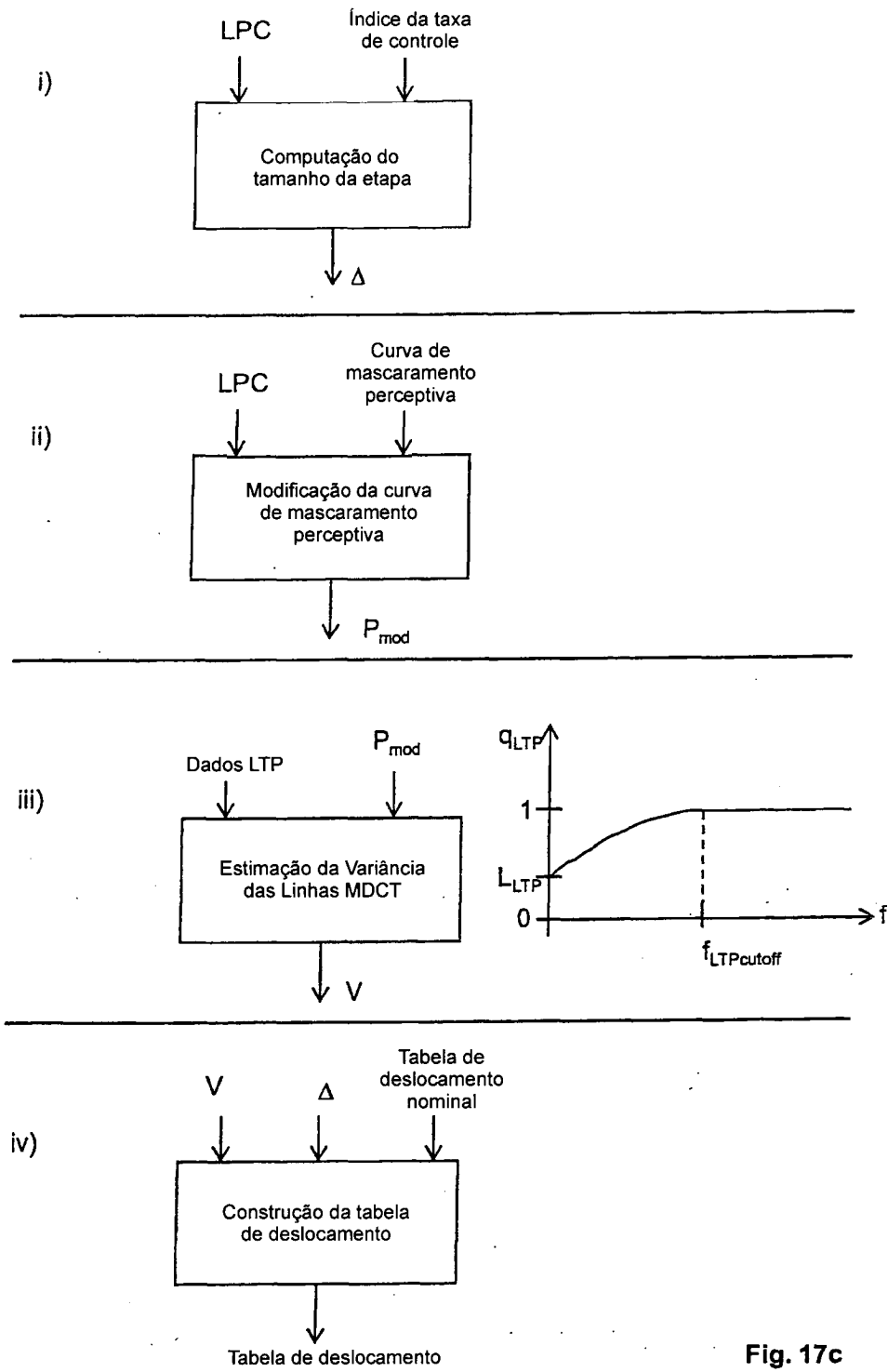


Fig. 17c

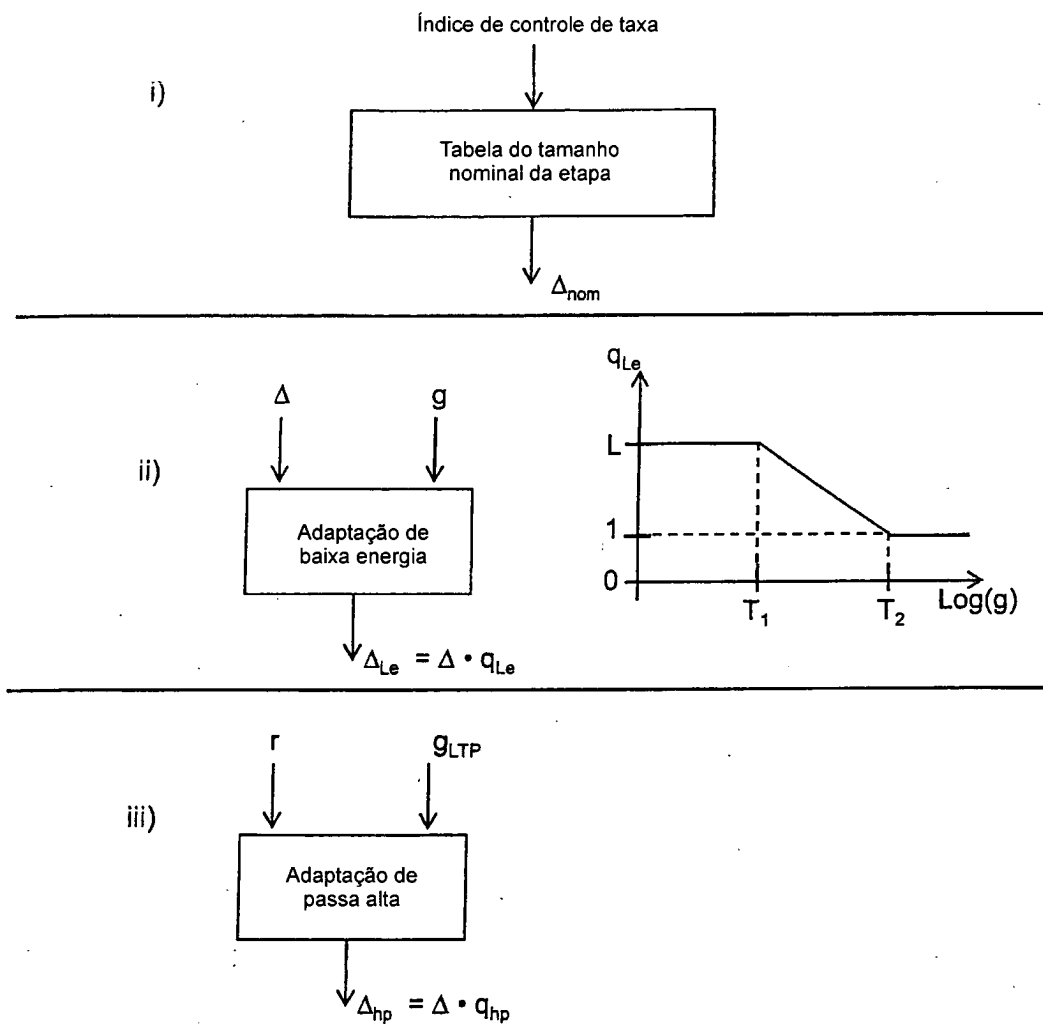


Fig. 17d

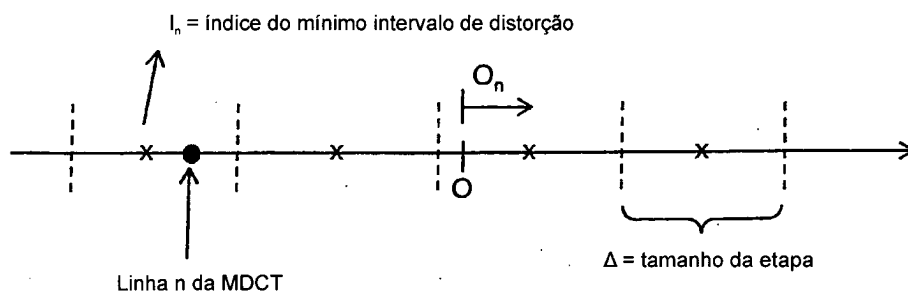


Fig. 17f

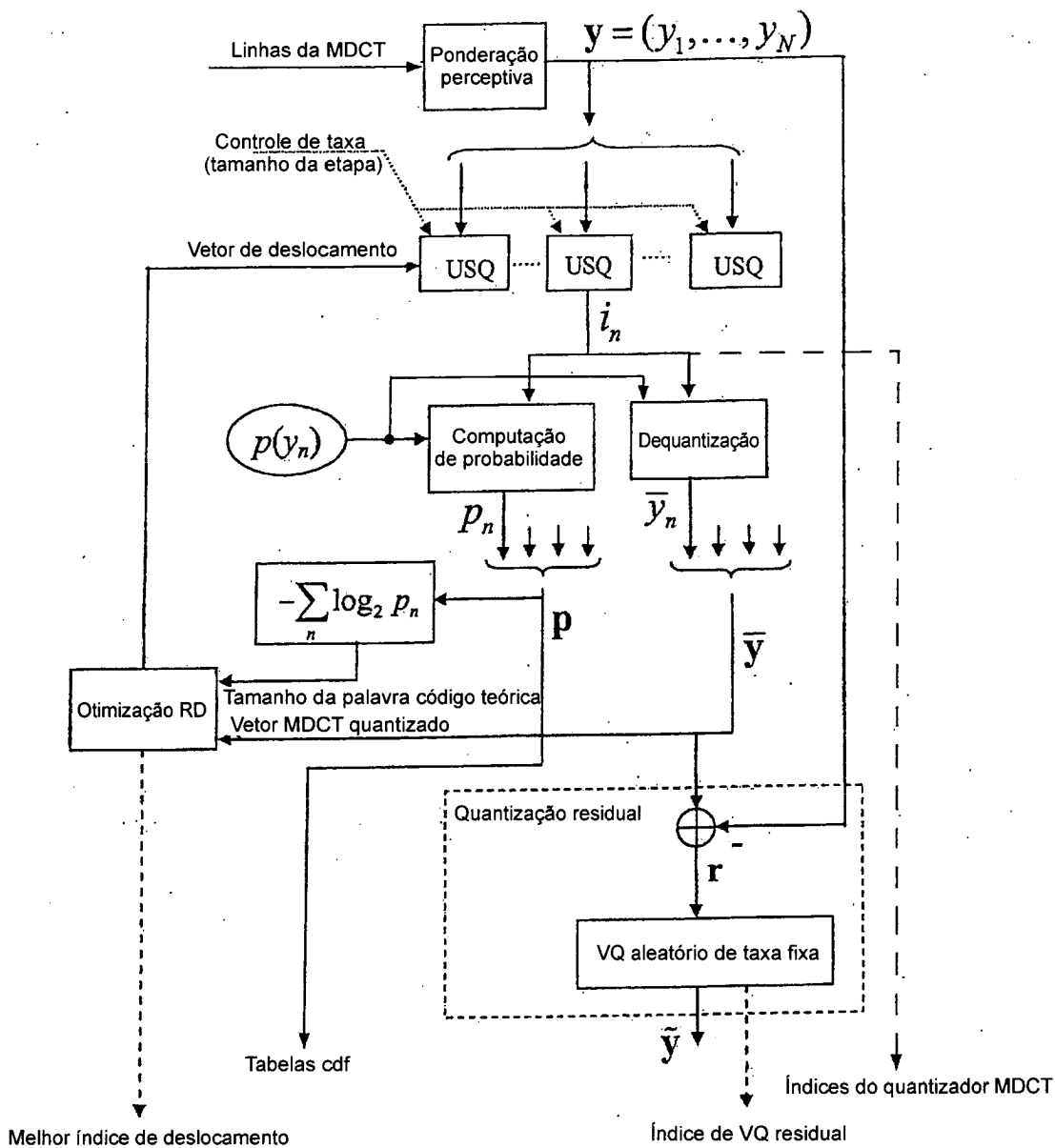


Fig. 17e

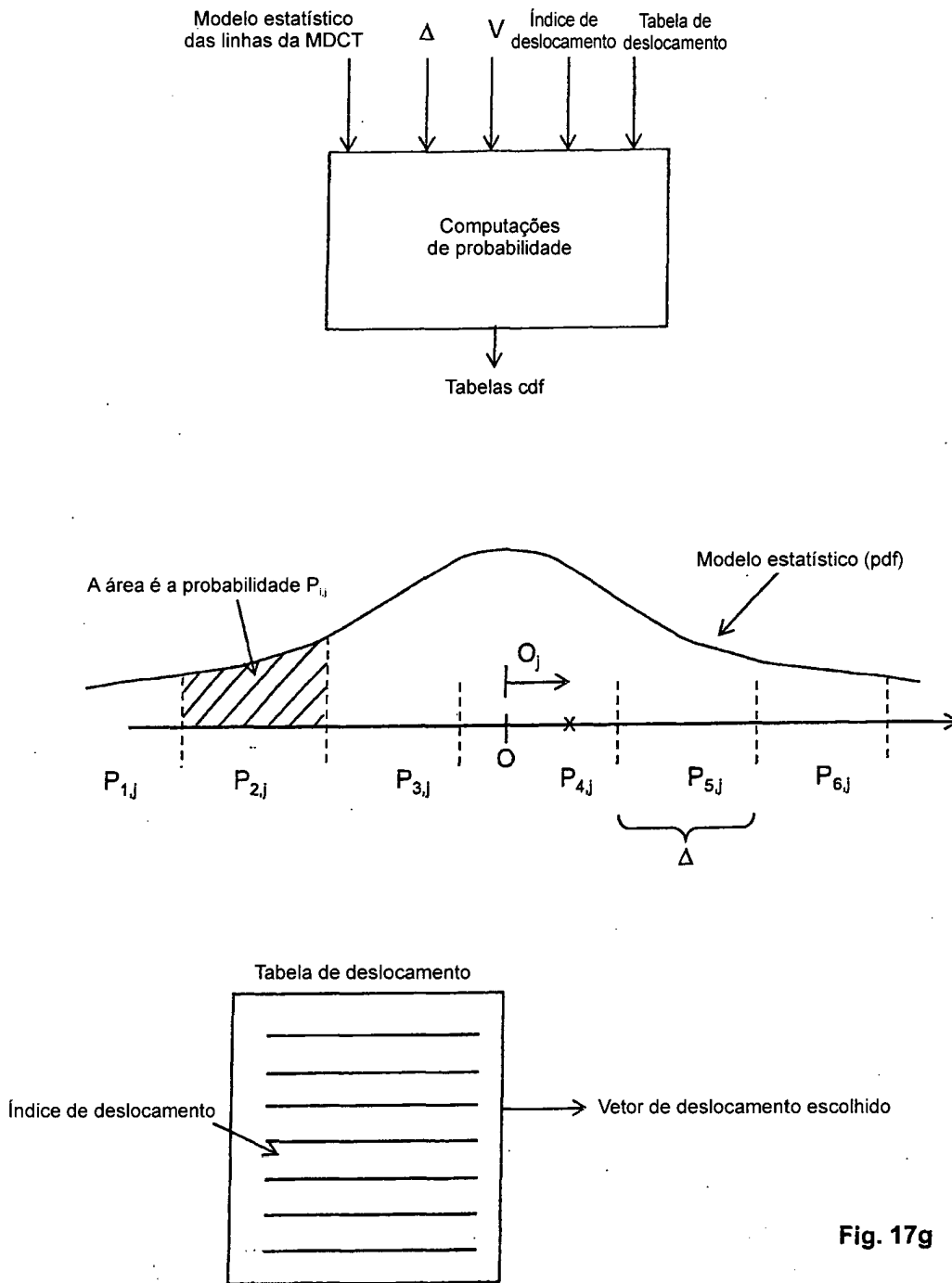


Fig. 17g

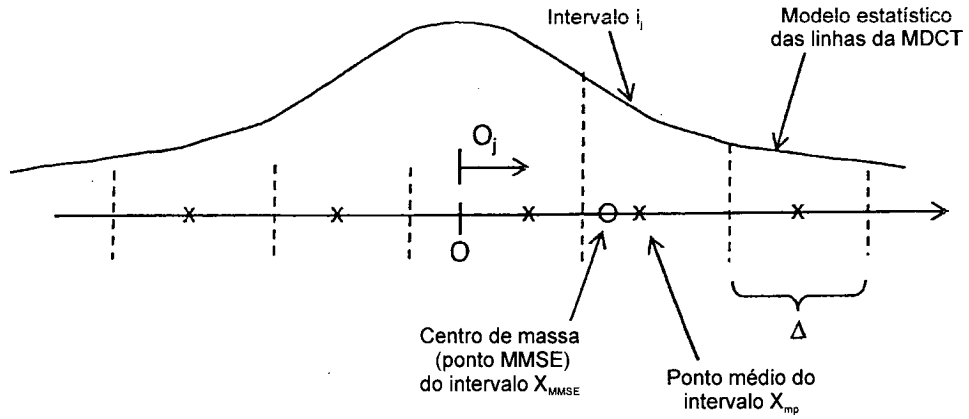


Fig. 17h

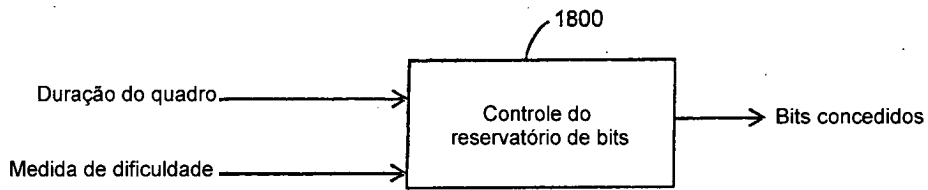


Fig. 18

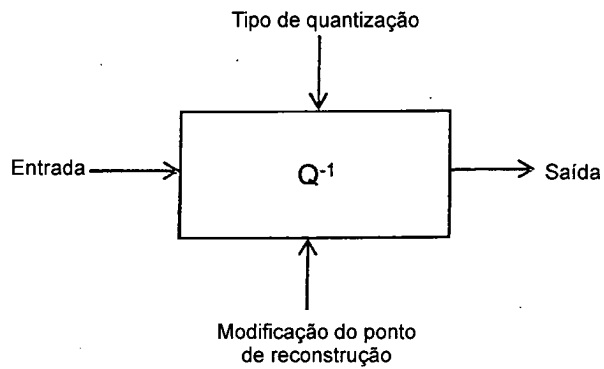


Fig. 19

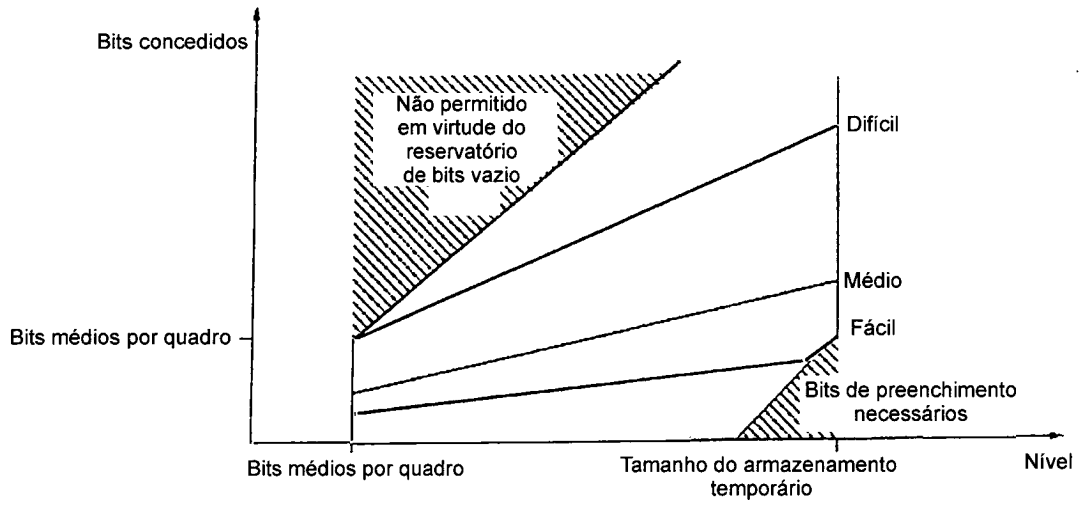


Fig. 18a

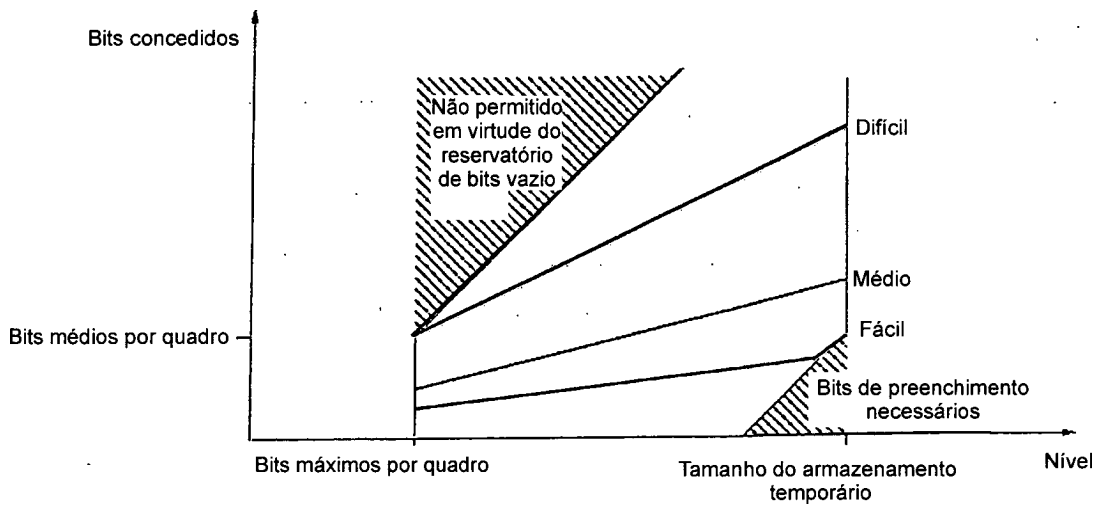


Fig. 18b

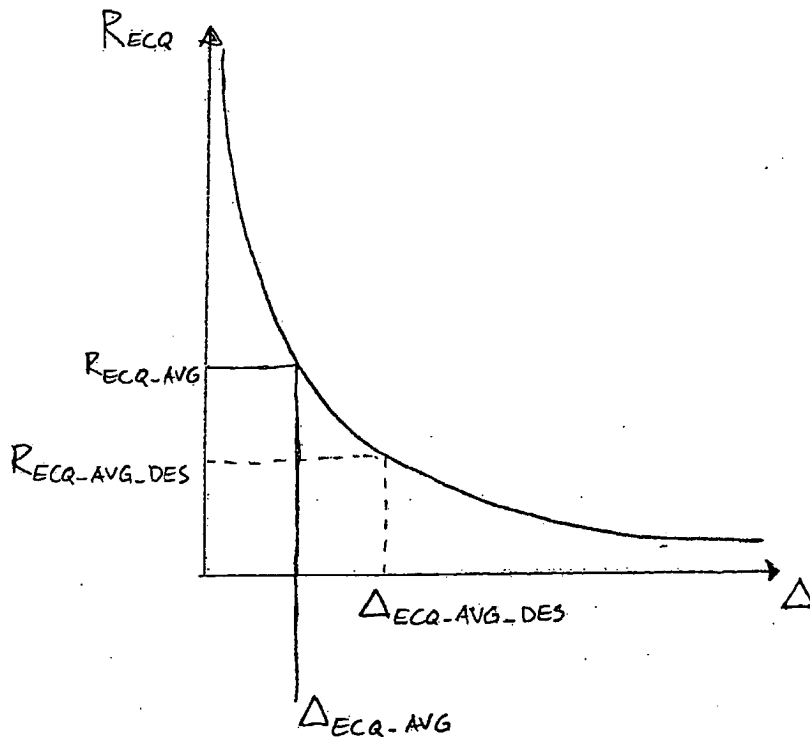


Fig. 18c