



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

**(11) PI 0607646-7 B1**



**(22) Data do Depósito: 03/04/2006**

**(45) Data de Concessão: 25/05/2021**

---

**(54) Título:** MÉTODO E EQUIPAMENTO PARA ENCODIFICAÇÃO POR DIVISÃO DE BANDA DE SINAIS DE FALA

**(51) Int.Cl.:** G10L 19/02; G10L 19/24; G10L 21/0232.

**(52) CPC:** G10L 19/0208; G10L 19/24; G10L 21/0232.

**(30) Prioridade Unionista:** 01/04/2005 US 60/667,901; 22/04/2005 US 60/673,965.

**(73) Titular(es):** QUALCOMM INCORPORATED.

**(72) Inventor(es):** KOEN BERNARD VOS; ANANTHAPADMANABHAN A. KANDHADAI.

**(86) Pedido PCT:** PCT US2006012230 de 03/04/2006

**(87) Publicação PCT:** WO 2006/107836 de 12/10/2006

**(85) Data do Início da Fase Nacional:** 01/10/2007

**(57) Resumo:** MÉTODO E EQUIPAMENTO PARA ENCODIFICAÇÃO POR DIVISÃO DE BANDA DE SINAIS DE FALA. Um codificador de fala de banda larga de acordo com uma modalidade inclui um banco de filtros tendo um percurso de processamento de banda baixa e um percurso de processamento de banda alta. Os percursos de processamento têm respostas de frequência de sobreposição. Um primeiro codificador é configurado para codificar um sinal de fala produzido pelo percurso de processamento de banda baixa de acordo com uma primeira metodologia de codificação. Um segundo codificador é configurado para codificar um sinal de fala produzido pelo percurso de processamento de banda alta de acordo com uma segunda metodologia de codificação que é diferente da primeira metodologia de codificação.

**"MÉTODO E EQUIPAMENTO PARA ENCODIFICAÇÃO POR DIVISÃO DE  
BANDA DE SINAIS DE FALA"**

Campo da Invenção

[0001] A presente invenção refere-se ao processamento de sinais.

Descrição da Técnica Anterior

[0002] Comunicações de voz através da Rede de Telefonia Comutada Pública (PSTN) têm sido tradicionalmente limitadas em largura de banda à faixa de frequência de 300-3400 kHz. Novas redes para comunicações de voz, tal como telefonia celular e voz sobre IP (Protocolo Internet, VoIP), podem não ter os mesmos limites de largura de banda, e pode ser desejável transmitir e receber comunicações de voz que incluem uma faixa de frequência de banda larga através destas redes. Por exemplo, pode ser desejável suportar uma faixa de frequência de áudio que estenda abaixo até 50 Hz e/ou acima de 7 ou 8 kHz. Pode ser também desejável suportar outras aplicações, tal como conferência de áudio/vídeo ou áudio de alta qualidade, que pode ter conteúdo de fala de áudio em faixas fora dos limites tradicionais de PSTN.

[0003] A extensão da faixa suportada por um codificador de fala em frequências mais altas pode melhorar a inteligibilidade. Por exemplo, a informação que diferencia fricativos como 's' e 'f' está amplamente nas frequências altas. A extensão de banda alta pode também melhorar outras qualidades de fala, tal como presença. Por exemplo, mesmo uma vogal vozeada pode ter energia espectral bem acima do limite PSTN.

[0004] Uma abordagem à codificação de fala de banda larga envolve escalar uma técnica de codificação de fala de banda estreita (por exemplo, uma configurada para encodificar a faixa de 0-4 kHz) para cobrir o espectro de banda larga. Por exemplo, um sinal de fala pode ser amostrado

em uma taxa mais elevada para incluir componentes em frequências elevadas, e uma técnica de codificação de banda estreita pode ser reconfigurada para utilizar mais coeficientes de filtro para representar esse sinal de banda larga. Técnicas de codificação de banda estreita como CELP (Predição Linear Excitada por Livro-Código) são intensas por meio computacional, entretanto, e um codificador CELP de banda larga pode consumir ciclos de processamento em demasia para ser prático para muitas aplicações móveis e outras aplicações incorporadas. Encodificação do espectro total de um sinal de banda larga em uma qualidade desejada utilizando tal técnica também pode levar a um aumento inaceitavelmente grande na largura de banda. Além disso, a transcodificação de tal sinal encodificado seria necessária antes mesmo que sua parte de banda estreita pudesse ser transmitida para e/ou decodificada por um sistema que somente suporta codificação de banda estreita.

[0005] Outra abordagem à codificação de fala de banda larga envolve extrapolar o envelope espectral de banda alta a partir do envelope espectral de banda estreita encodificado. Embora tal abordagem possa ser implementada sem nenhum aumento na largura de banda e sem necessidade de transcodificação, o envelope espectral grosso ou estrutura formante da parte de banda alta de um sinal de fala genericamente não pode ser predito precisamente a partir do envelope espectral da parte de banda estreita.

[0006] Pode ser desejável implementar codificação de fala de banda larga de tal modo que pelo menos a parte de banda estreita do sinal encodificado possa ser enviada através de um canal de banda estreita (como um canal PSTN) sem transcodificar ou outra modificação significativa. A eficiência da extensão de codificação de banda larga também pode ser desejável, por exemplo, para evitar uma redução

significativa no número de usuários que podem ser servidos em aplicações como telefonia celular sem fio e difusão através de canais cabeados e sem fio.

#### Sumário da Invenção

[0007] Em uma modalidade, um equipamento inclui um primeiro codificador de fala configurado para codificar um sinal de fala de banda baixa; um segundo codificador de fala configurado para codificar um sinal de fala de banda alta; e um banco de filtros tendo (A) um percurso de processamento de banda baixa configurado para receber um sinal de fala de banda larga tendo conteúdo de frequência entre pelo menos 1000 e 6000 Hz e produzir o sinal de fala de banda baixa e (B) um percurso de processamento de banda alta configurado para receber o sinal de fala de banda larga e produzir o sinal de fala de banda alta. O sinal de fala de banda baixa se baseia em uma primeira parte do conteúdo de frequência do sinal de banda larga, a primeira parte incluindo a parte do sinal de banda larga entre 1000 e 2000 Hz. O sinal de fala de banda alta se baseia em uma segunda parte do conteúdo de frequência do sinal de banda larga, a segunda parte incluindo a parte do sinal de banda larga entre 5000 e 6000 Hz. Cada um dentre o sinal de fala de banda baixa e o sinal de fala de banda alta se baseia em uma terceira parte do conteúdo de frequência do sinal de banda larga, a terceira parte incluindo uma parte do sinal de banda larga entre 2000 e 5000 Hz que tem uma largura de pelo menos 250 Hz.

[0008] Em outra modalidade, um equipamento inclui um banco de filtros tendo (A) um percurso de processamento de banda baixa configurado para receber um sinal de fala de banda larga e produzir um sinal de fala de banda baixa com base em uma parte de frequência baixa do sinal de fala de banda larga e (B) um percurso de

processamento de banda alta configurado para receber o sinal de fala de banda larga e produzir um sinal de fala de banda alta com base em uma parte de frequência alta do sinal de fala de banda alta. Uma banda passante do percurso de processamento de banda baixa sobrepõe uma banda passante do percurso de processamento de banda alta. O equipamento também inclui um primeiro encodificador de fala, configurado para encodificar o sinal de fala de banda baixa em pelo menos um sinal de excitação de banda baixa encodificado e uma pluralidade de parâmetros de filtro de banda baixa; e um segundo encodificador de fala configurado para gerar um sinal de excitação de banda alta com base no sinal de excitação de banda baixa encodificado, e encodificar o sinal de banda alta, de acordo com o sinal de excitação de banda alta, em pelo menos uma pluralidade dos parâmetros de filtro de banda alta.

[0009] Em outra modalidade, um método de processamento de sinais inclui produzir um sinal de fala de banda baixa com base em um sinal de fala de banda larga tendo conteúdo de frequência entre pelo menos 1000 e 6000 Hz; encodificar o sinal de fala de banda baixa; produzir um sinal de fala de banda alta com base no sinal de fala de banda larga; e encodificar o sinal de fala de banda alta. Nesse método, a produção de um sinal de fala de banda baixa inclui produzir o sinal de fala de banda baixa com base em (A) uma primeira parte do conteúdo de frequência do sinal de banda larga, a primeira parte incluindo a parte do sinal de banda larga entre 1000 e 2000 Hz, e (B) uma terceira parte do conteúdo de frequência do sinal de banda larga, a terceira parte incluindo uma parte do sinal de banda larga entre 2000 e 5000 Hz que tem uma largura de pelo menos 250 Hz. Nesse método, a produção de um sinal de fala de banda alta inclui a produção do sinal de fala de banda alta com base em (C)

uma segunda parte do conteúdo de frequência do sinal de banda larga, a segunda parte incluindo a parte do sinal de banda larga entre 5000 e 6000 Hz, e (D) a terceira parte do conteúdo de frequência do sinal de banda larga.

Breve Descrição das Figuras

[0010] A FIGURA 1a mostra um diagrama em blocos de um codificador de fala de banda larga A100 de acordo com uma modalidade.

[0011] A FIGURA 1b mostra um diagrama em blocos de uma implementação A102 de codificador de fala de banda larga A100.

[0012] A FIGURA 2a mostra um diagrama em blocos de um decodificador de fala de banda larga B100 de acordo com uma modalidade.

[0013] A FIGURA 2b mostra um diagrama em blocos de uma implementação B102 de codificador de fala de banda larga B100.

[0014] A FIGURA 3a mostra um diagrama em blocos de uma implementação A112 do banco de filtros A110.

[0015] A FIGURA 3b mostra um diagrama em blocos de uma implementação B122 do banco de filtros B120.

[0016] A FIGURA 4a mostra cobertura de largura de banda das bandas baixa e alta para um exemplo do banco de filtros A110.

[0017] A FIGURA 4b mostra cobertura de largura de banda das bandas baixa e alta para outro exemplo de banco de filtros A110.

[0018] A FIGURA 4c mostra um diagrama em blocos de uma implementação A114 de banco de filtros A112.

[0019] A FIGURA 4d mostra um diagrama em blocos de uma implementação B124 do banco de filtros B122.

[0020] A FIGURA 5a mostra um exemplo de um gráfico de frequência vs. Amplitude log para um sinal de fala.

[0021] A FIGURA 5b mostra um diagrama em blocos de um sistema de codificação de predição linear básico.

[0022] A FIGURA 6 mostra um diagrama em blocos de uma implementação A122 de codificador de banda estreita A120.

[0023] A FIGURA 7 mostra um diagrama em blocos de uma implementação B112 de decodificador de banda estreita B110.

[0024] A FIGURA 8a mostra um exemplo de um gráfico de frequências vs. Amplitude log para um sinal residual para fala vozeada.

[0025] A FIGURA 8b mostra um exemplo de um gráfico de tempo vs. Amplitude log para um sinal residual para fala vozeada.

[0026] A FIGURA 9 mostra um diagrama em blocos de um sistema de codificação de predição linear básico que também executa predição a longo prazo.

[0027] A FIGURA 10 mostra um diagrama em blocos de uma implementação A202 de codificador de banda alta A200.

[0028] A FIGURA 11 mostra um diagrama em blocos de uma implementação A302 de gerador de excitação de banda alta A300.

[0029] A FIGURA 12 mostra um diagrama em blocos de uma implementação A402 do extensor de espectro A400.

[0030] A FIGURA 12a mostra gráficos de espectros de sinais em vários pontos em um exemplo de uma operação de extensão espectral.

[0031] A FIGURA 12b mostra gráficos de espectros de sinais em vários pontos em outro exemplo de uma operação de extensão espectral.

[0032] A FIGURA 13 mostra um diagrama em blocos de uma implementação A304 de gerador de excitação de banda alta A302.

[0033] A FIGURA 14 mostra um diagrama em blocos de uma implementação A306 de gerador de excitação de banda alta A302.

[0034] A FIGURA 15 mostra um fluxograma para uma tarefa de cálculo de envelope T100.

[0035] A FIGURA 16 mostra um diagrama em blocos de uma implementação 492 do combinador 490.

[0036] A FIGURA 17 ilustra uma abordagem para calcular uma medição de periodicidade de sinal de banda alta S30.

[0037] A FIGURA 18 mostra um diagrama em blocos de uma implementação A312 de gerador de excitação de banda alta A302.

[0038] A FIGURA 19 mostra um diagrama em blocos de uma implementação A314 do gerador de excitação de banda alta A302.

[0039] A FIGURA 20 mostra um diagrama em blocos de uma implementação A316 do gerador de excitação de banda alta A302.

[0040] A FIGURA 21 mostra um fluxograma para uma tarefa de cálculo de ganho T200.

[0041] A FIGURA 22 mostra um fluxograma para uma implementação T210 de tarefa de cálculo de ganho T200.

[0042] A FIGURA 23a mostra um diagrama de uma função de janelamento.

[0043] A FIGURA 23b mostra uma aplicação de uma função de janelamento como mostrado na FIGURA 23a para subquadros de um sinal de fala.

[0044] A FIGURA 24 mostra um diagrama em blocos para uma implementação B202 de decodificador de banda alta B200.

[0045] A FIGURA 25 mostra um diagrama em blocos de uma implementação AD10 de codificador de fala de banda larga A100.

[0046] A FIGURA 26a mostra um diagrama esquemático de uma implementação D122 de linha de retardo D120.

[0047] A FIGURA 26b mostra um diagrama esquemático de uma implementação D124 de linha de retardo D120.

[0048] A FIGURA 27 mostra um diagrama esquemático de uma implementação D130 de linha de retardo D120.

[0049] A FIGURA 28 mostra um diagrama em blocos de uma implementação AD12 de codificador de fala de banda larga AD10.

[0050] A FIGURA 29 mostra um fluxograma de um método de processamento de sinal MD100 de acordo com uma modalidade.

[0051] A FIGURA 30 mostra um fluxograma para um método M100 de acordo com uma modalidade.

[0052] A FIGURA 31a mostra um fluxograma para um método M200 de acordo com uma modalidade.

[0053] A FIGURA 31b mostra um fluxograma para uma implementação M210 do método M200.

[0054] A FIGURA 32 mostra um fluxograma para um método M300 de acordo com uma modalidade.

[0055] As FIGURAS 33-36b mostram respostas de impulso e frequência para operações de filtragem mostradas na FIGURA 4c.

[0056] As FIGURAS 37a-39b mostram respostas de impulso e frequência para operações de filtragem mostradas na FIGURA 4d.

[0057] Nas FIGURAs e descrição associada, os mesmos números de referência se referem a elementos ou sinais iguais ou análogos.

#### Descrição Detalhada da Invenção

[0058] Modalidades como descrito aqui incluem sistemas, métodos e equipamentos que podem ser configurados para fornecer uma extensão para um codificador de fala de banda estreita para suportar transmissão e/ou armazenamento de sinais de fala de banda larga em um aumento de largura de banda de aproximadamente 800 a 1000 bps (bits por segundo). Vantagens em potencial de tais implementações incluem codificação incorporada para suportar compatibilidade com sistemas de banda estreita, alocação e realocação relativamente fáceis de bits entre os canais de codificação de banda estreita e banda alta, para evitar uma operação de síntese de banda larga intensa de forma computacional, e manter uma taxa de amostragem baixa para sinais a serem processados por rotinas de forma computacional de codificação de forma de onda intensas.

[0059] A menos que expressamente limitado por seu contexto, o termo "calcular" é utilizado aqui para indicar quaisquer de seus significados comuns, como computar, gerar e selecionar de uma lista de valores. Onde o termo "compreendendo" é utilizado na presente descrição e reivindicações, não exclui outros elementos ou operações. O termo "A é baseado em B" é utilizado para indicar quaisquer de seus significados comuns, incluindo os casos (i) "A é

igual a B" e (ii) "A é baseado em pelo menos B". O termo "Protocolo Internet" inclui versão 4, como descrito em IETF (Força Tarefa da Engenharia da Internet) RFV (Solicitação por Comentários) 791, e versões subseqüentes como versão 6.

[0060] A FIGURA 1a mostra um diagrama em blocos de um encodificador de fala de banda larga A100 de acordo com uma modalidade. O banco de filtros A110 é configurado para filtrar um sinal de fala de banda larga S10 para produzir um sinal de banda estreita S20 e um sinal de banda alta S30. O encodificador de banda estreita A120 é configurado para encodificar o sinal de banda estreita S20 para produzir parâmetros de filtro de banda estreita (NB) S40 e um sinal residual de banda estreita S50. Como descrito em detalhes adicionais aqui, o encodificador de banda estreita A120 é tipicamente configurado para produzir parâmetros de filtro de banda estreita S40 e sinal de excitação de banda estreita encodificado S50 como índices de livro-código ou em outra forma quantizada. Encodificador de banda alta A200 é configurado para encodificar um sinal de banda alta S30 de acordo com informações no sinal de excitação de banda estreita encodificado S50 para produzir parâmetros de codificação de banda alta S60. Como descrito em detalhes adicionais aqui, o encodificador de banda alta A200 é tipicamente configurado para produzir parâmetros de codificação de banda alta S60 como índices de livro-código ou em outra forma quantizada. Um exemplo específico de encodificador de fala de banda larga A100 é configurado para encodificar o sinal de fala de banda larga S10 em uma taxa de aproximadamente 8,55 kbps (kilobits por segundo), com aproximadamente 7,55 kbps sendo utilizado para parâmetros de filtro de banda estreita S40 e sinal de excitação de banda estreita encodificado S50, e aproximadamente 1 kbps sendo utilizado para parâmetros de codificação de banda alta S60.

[0061] Pode ser desejado combinar os sinais de banda alta e banda estreita encodificados em um único fluxo de bits. Por exemplo, pode ser desejado multiplexar os sinais encodificados juntos para transmissão (por exemplo, através de um canal de transmissão cabeado, óptico ou sem fio), ou para armazenamento, como um sinal de fala de banda larga encodificado. A FIGURA 1b mostra um diagrama em blocos de uma implementação A102 de encodificador de fala de banda larga A100 que inclui um multiplexador A130 configurado para combinar parâmetros de filtro de banda estreita S40, sinal de excitação de banda estreita encodificado S50, e parâmetros de filtro de banda alta S60 em um sinal multiplexado S70.

[0062] Um equipamento incluindo encodificador A102 também pode incluir circuitos configurados para transmitir sinal multiplexado S70 em um canal de transmissão como um canal cabeado, óptico ou sem fio. Tal equipamento pode ser também configurado para executar uma ou mais operações de encodificação de canal no sinal, como encodificação de correção de erro (por exemplo, encodificação convolucional de taxa compatível) e/ou encodificação de detecção de erro (por exemplo, encodificação de redundância cíclica), e/ou uma ou mais camadas de encodificação de protocolo de rede (por exemplo, Ethernet, TCP/IP, cdma2000).

[0063] Pode ser desejável que o multiplexador A130 seja configurado para incorporar o sinal de banda estreita encodificado (incluindo parâmetros de filtro de banda estreita S40 e sinal de excitação de banda estreita encodificado S50) como um subfluxo separável de sinal multiplexado S70, de tal modo que o sinal de banda estreita encodificado possa ser recuperado e decodificado independentemente de outra parte de sinal multiplexado S70 como um sinal de banda alta e/ou banda baixa. Por exemplo,

o sinal multiplexado S70 pode ser disposto de tal modo que o sinal de banda estreita encodificado possa ser recuperado por extração dos parâmetros de filtro de banda alta S60. Uma vantagem potencial de tal característica é evitar a necessidade de transcodificar o sinal de banda larga encodificado antes de passar o mesmo para um sistema que suporta decodificação do sinal de banda estreita, porém, não suporta decodificação da parte de banda alta.

[0064] A FIGURA 2a é um diagrama em blocos de um decodificador de fala de banda larga B100, de acordo com uma modalidade. O decodificador de banda estreita B110 é configurado para decodificar parâmetros de filtro de banda estreita S40 e sinal de excitação de banda estreita encodificado S50 para produzir um sinal de banda estreita S90. O decodificador de banda alta B200 é configurado para decodificar parâmetros de codificação de banda alta S60, de acordo com um sinal de excitação de banda estreita S80, com base em sinal de excitação de banda estreita encodificado S50, para produzir um sinal de banda alta S100. Neste exemplo, o decodificador de banda estreita B110 é configurado para fornecer sinal de excitação de banda estreita S80 para o decodificador de banda alta B200. O banco de filtros B120 é configurado para combinar sinal de banda estreita S90 e sinal de banda alta S100 para produzir um sinal de fala de banda larga S110.

[0065] A FIGURA 2b é um diagrama em blocos de uma implementação B102 de decodificador de fala de banda larga B100 que inclui um demultiplexador B130 configurado para produzir sinais encodificados S40, S50 e S60 a partir do sinal multiplexado S70. Um equipamento incluindo decodificador B102 pode incluir circuitos configurados para receber sinal multiplexado S70 a partir de um canal de transmissão como um canal cabeado, óptico ou sem fio. Tal

equipamento pode ser também configurado para executar uma ou mais operações de decodificação de canal no sinal, como decodificação de correção de erro (por exemplo, decodificação convolucional compatível em taxa) e/ou decodificação de detecção de erro (por exemplo, decodificação de redundância cíclica), e/ou uma ou mais camadas de decodificação de protocolo de rede (por exemplo, Ethernet, TCP/IP, cdma2000).

[0066] O banco de filtros A110 é configurado para filtrar um sinal de entrada de acordo com um esquema por divisão de banda para produzir uma sub-banda de baixa frequência e uma sub-banda de alta frequência. Dependendo dos critérios de projeto para a aplicação específica, as sub-bandas de saída podem ter larguras de banda iguais ou desiguais e podem ser sobrepostas ou não sobrepostas. Uma configuração de banco de filtros A110 que produza mais do que duas sub-bandas, também é possível. Por exemplo, tal banco de filtros pode ser configurado para produzir um ou mais sinais de banda baixa que incluem componentes em uma faixa de frequências abaixo daquela de sinal de banda estreita S20 (como a faixa de 50-300 Hz). Também é possível que tal banco de filtros seja configurado para produzir um ou mais sinais de banda alta adicionais que incluem componentes em uma faixa de frequência acima daquela do sinal de banda alta S30 (como uma faixa de 14-20, 16-20 ou 16-32 kHz). Em tal caso, o codificador de fala de banda larga A100 pode ser implementado para codificar esse sinal ou sinais separadamente, e o multiplexador A130 pode ser configurado para incluir o sinal ou sinais codificados adicionais no sinal multiplexado S70 (por exemplo, como uma parte separável).

[0067] A FIGURA 3a mostra um diagrama em blocos de uma implementação A112 do banco de filtros A110 que é

configurada para produzir dois sinais de sub-banda tendo taxas de amostragem reduzidas. O banco de filtros A110 é disposto para receber um sinal de fala de banda larga S10 tendo uma parte de alta frequência (ou banda alta) e uma parte de baixa frequência (ou banda baixa). O banco de filtros A112 inclui um percurso de processamento de banda baixa configurado para receber sinal de fala de banda larga S10 e para produzir sinal de fala de banda estreita S20, e um percurso de processamento de banda alta configurado para receber sinal de fala de banda larga S10 e produzir sinal de fala de banda alta S30. O filtro passa-baixa 110 filtra o sinal de fala de banda larga S10 para passar uma sub-banda de baixa frequência selecionada, e o filtro passa-alta 130 filtra o sinal de fala de banda larga S10 para passar uma sub-banda de alta frequência selecionada. Como os dois sinais de sub-banda têm mais larguras de banda estreitas do que o sinal de fala de banda larga S10, suas taxas de amostragem podem ser reduzidas até certo ponto sem perda de informações. O subamostrador 120 reduz a taxa de amostragem do sinal passa-baixa de acordo com um fator de decimação desejado (por exemplo, removendo amostras do sinal e/ou substituindo amostras com valores médios), e o subamostrador 140 reduz de modo semelhante a taxa de amostragem do sinal passa-alta de acordo com outro fator de decimação desejado.

[0068] A FIGURA 3b mostra um diagrama em blocos de uma implementação correspondente B122 do banco de filtros B120. O sobreamostrador 150 aumenta a taxa de amostragem do sinal de banda estreita S90 (por exemplo, por inserção de zero e/ou por duplicação de amostras), e o filtro passa-baixa 160 filtra o sinal superamostrado para passar somente uma parte de banda baixa (por exemplo, evitar "aliasing"). De modo semelhante, o sobreamostrador 170 aumenta a taxa de amostragem do sinal de banda alta S100 e o filtro passa-alta

180 filtra o sinal superamostrado para passar somente uma parte de banda alta. Os dois sinais de banda passante são então somados para formar o sinal de fala de banda larga S110. Em algumas implementações do decodificador B100, o banco de filtros B120 é configurado para produzir uma soma ponderada dos dois sinais de banda passante de acordo com uma ou mais ponderações recebidas e/ou calculadas por decodificador de banda alta B200. Uma configuração de banco de filtros B120 que combina mais do que dois sinais de banda passante, também é considerada.

[0069] Cada um dos filtros 110, 130, 160, 180 pode ser implementado como um filtro de resposta de impulso finita (FIR) ou como um filtro de resposta de impulso infinita (IIR). As respostas de frequência de filtros encodificadores 110 e 130 podem ter regiões de transição simétricas ou diferentemente formadas entre a banda de corte e a banda passante. De modo semelhante, as respostas de frequência dos filtros decodificadores 160 e 180 podem ter regiões de transição simétricas ou diferentemente moldadas entre a banda de corte e a banda passante. Pode ser desejável, porém, não é estritamente necessário que o filtro passa-baixa 110 tenha a mesma resposta que o filtro passa-baixa 160, e para o filtro passa-alta 130 ter a mesma resposta que o filtro passa-alta 180. Em um exemplo, os dois pares de filtros 110, 130 e 160, 180 são bancos de filtros em espelho de quadratura (QMF), com o par de filtros 110, 130 tendo os mesmos coeficientes que o par de filtros 160, 180.

[0070] Em um exemplo típico, o filtro passa-baixa 110 tem uma banda passante que inclui a faixa limitada de PSTN de 300-3400 Hz (por exemplo, a banda de 0 a 4 kHz). As FIGURAS 4a e 4b mostram larguras de banda relativas de sinal de fala de banda larga S10, sinal de banda estreita

S20, e sinal de banda alta S30 em dois exemplos de implementação diferentes. Nesses dois exemplos específicos, o sinal de fala de banda larga S10 tem uma taxa de amostragem de 16 kHz (representando componentes de frequência compreendidos na faixa de 0 a 8 kHz), e o sinal de banda estreita S20 tem uma taxa de amostragem de 8 kHz (representando componentes de frequência compreendidos na faixa de 0 a 4 kHz).

[0071] No exemplo da FIGURA 4a, não há sobreposição significativa entre as duas sub-bandas. Um sinal de banda alta S30 como mostrado nesse exemplo pode ser obtido utilizando um filtro passa-alta 130 com uma banda passante de 4-8 kHz. Em tal caso, pode ser desejável reduzir a taxa de amostragem para 8 kHz por subamostragem do sinal filtrado por um fator de dois. Tal operação que pode se esperar que reduza significativamente a complexidade computacional de operações de processamento adicionais no sinal, moverá a energia de banda passante para baixo da faixa de 0 a 4 kHz sem perda de informação.

[0072] No exemplo alternativo da FIGURA 4b, as sub-bandas superior e inferior têm uma sobreposição apreciável, de tal modo que a região de 3,5 a 4 kHz é descrita pelos dois sinais de sub-banda. Um sinal de banda alta S30 como nesse exemplo pode ser obtido utilizando um filtro passa-alta 130 com uma banda passante de 3,5 - 7 kHz. Em tal caso, pode ser desejável reduzir a taxa de amostragem para 7 kHz por subamostragem do sinal filtrado por um fator de 16/7. Tal operação, que pode se esperar que reduza significativamente a complexidade computacional de operações de processamento adicionais no sinal, moverá a energia de banda passante para baixo da faixa de 0 a 3,5 kHz sem perda de informação.

[0073] Em um aparelho de telefone típico para comunicação telefônica, um ou mais dos transdutores (isto é, o microfone e o auricular ou alto-falante) não tem uma resposta apreciável sobre a faixa de frequência de 7 - 8 kHz. No exemplo da FIGURA 4b, a parte de sinal de fala de banda larga S10 entre 7 e 8 kHz não está incluída no sinal encodificado. Outros exemplos específicos de filtro passa-alta 130 têm bandas passantes de 3,5 - 7,5 kHz e 3,5 - 8 kHz.

[0074] Em algumas implementações, o fornecimento de uma sobreposição entre sub-bandas como no exemplo da FIGURA 4b permite o uso de um filtro passa-baixa e/ou passa-alta tendo um roll-off suave sobre a região sobreposta. Tais filtros são tipicamente de projeto mais fácil, menos complexos de forma computacional, e/ou introduzem menos retardo do que os filtros com respostas mais nítidas ou "brick-wall". Os filtros tendo regiões de transição nítida tendem a ter lóbulos secundários mais elevados (que podem causar aliasing) do que filtros de ordem similar que têm roll-offs suaves. Os filtros tendo regiões de transição nítida também podem ter respostas de impulso longas que podem causar artefatos de toque. Para implementações de banco de filtros tendo um ou mais filtros IIR, a permissão de um roll-off suave sobre a região sobreposta pode permitir o uso de um filtro ou filtros cujos pólos são mais distantes do círculo de unidade, que pode ser importante para assegurar uma implementação de ponto fixo estável.

[0075] A sobreposição de sub-bandas permite uma mistura suave de banda baixa e banda alta que pode levar a um número menor de artefatos audíveis, aliasing reduzido e/ou uma transição menos perceptível de uma banda para a outra. Além disso, a eficiência de codificação de

codificador de banda estreita A120 (por exemplo, um codificador de forma de onda) pode cair com frequência crescente. Por exemplo, a qualidade de codificação do codificador de banda estreita pode ser reduzida em taxas de bit baixas, especialmente na presença de ruído de segundo plano. Em tais casos, a provisão de uma sobreposição das sub-bandas pode aumentar a qualidade de componentes de frequência reproduzida na região sobreposta.

[0076] Além disso, a sobreposição de sub-bandas permite uma mistura suave de banda baixa e banda alta que pode levar a um número menor de artefatos audíveis, aliasing reduzido e/ou transição menos perceptível de uma banda para a outra. Tal característica pode ser especialmente desejável para uma implementação na qual o codificador de banda estreita A120 e codificador de banda alta A200 operam de acordo com metodologias de codificação diferentes. Por exemplo, técnicas de codificação diferentes podem produzir sinais que soam bem diferentes. Um codificador que encodifica um envelope espectral na forma de índices de livro-código pode produzir um sinal tendo um som diferente de um codificador que encodifica o espectro de amplitude. Um codificador de domínio de tempo (por exemplo, uma modulação de código-pulso ou codificador PCM) pode produzir um sinal tendo um som diferente de um codificador de domínio de frequência. Um codificador que encodifica um sinal com uma representação do envelope espectral e o sinal residual correspondente pode produzir um sinal tendo um som diferente de um codificador que encodifica um sinal com somente uma representação do envelope espectral. Um codificador que encodifica um sinal como uma representação de sua forma de onda pode produzir uma saída tendo um som diferente daquele de um codificador senoidal. Em tais casos, o uso de filtros tendo regiões de transição nítida para definir sub-bandas

não de sobreposição pode levar a uma transição abrupta e perceptivelmente notável entre as sub-bandas no sinal de banda larga sintetizado.

[0077] Embora bancos de filtros QMF tendo respostas de frequência de sobreposição complementares sejam freqüentemente utilizados em técnicas de sub-banda, tais filtros são inadequados para pelo menos algumas das implementações de codificação de banda larga descritas aqui. Um banco de filtros QMF no codificador é configurado para criar um grau significativo de aliasing que é cancelado no banco de filtros QMF correspondente no decodificador. Tal arranjo pode não ser apropriado para uma aplicação na qual o sinal incorre em uma quantidade significativa de aliasing entre os bancos de filtros, visto que a distorção pode reduzir a eficácia da propriedade de cancelamento de alias. Por exemplo, aplicações descritas aqui incluem implementações de codificação configuradas para operar em taxas de bit muito baixas. Como consequência da taxa de bit muito baixa, o sinal decodificado provavelmente deve aparecer significativamente distorcido em comparação com o sinal original, de tal modo que o uso de bancos de filtros QMF pode levar a aliasing não cancelado. As aplicações que utilizam bancos de filtros QMF têm tipicamente taxas de bits mais elevadas (por exemplo, acima de 12 kbps para AMR, e 64 kbps para G.722).

[0078] Adicionalmente, um codificador pode ser configurado para produzir um sinal sintetizado que é similar de forma perceptível ao sinal original, porém que difere na realidade significativamente do sinal original. Por exemplo, um codificador que deriva a excitação de banda alta a partir do residual de banda estreita como descrito aqui pode produzir tal sinal, visto que o residual de banda alta real pode ser totalmente ausente do sinal decodificado. O uso de

bancos de filtros QMF em tais aplicações pode levar a um grau significativo de distorção causado por aliasing não cancelado.

[0079] A quantidade de distorção causada por aliasing QMF pode ser reduzida se a sub-banda afetada for estreita, visto que o efeito do aliasing é limitado a uma largura de banda igual à largura da sub-banda. Para exemplos como descrito aqui nos quais cada sub-banda inclui aproximadamente metade da largura de banda de banda larga, entretanto, a distorção causada por aliasing não cancelado poderia afetar uma parte significativa do sinal. A qualidade do sinal também pode ser afetada pela localização da banda de frequência sobre a qual o aliasing não cancelado ocorre. Por exemplo, a distorção criada próximo ao centro de um sinal de fala de banda larga (por exemplo, entre 3 e 4 kHz) pode ser muito mais questionável do que a distorção que ocorre próximo a uma borda do sinal (por exemplo, acima de 6 kHz).

[0080] Embora as respostas dos filtros de um banco de filtros QMF sejam estritamente relacionadas entre si, os percursos de banda baixa e banda alta de bancos de filtros A110 e B120 podem ser configurados para ter espectros que são totalmente não relacionados fora a sobreposição das duas sub-bandas. Define-se a sobreposição das duas sub-bandas como a distância a partir do ponto no qual a resposta de frequência do filtro de banda alta cai para -20 dB até o ponto no qual a resposta de frequência do filtro de banda baixa cai para -20 dB. Em vários exemplos de banco de filtros A110 e/ou B120, essa sobreposição varia em torno de 200 Hz até em torno de 1 kHz. A faixa de aproximadamente 400 a aproximadamente 600 Hz pode representar uma compensação desejável entre eficiência de codificação e suavidade perceptível. Em um exemplo específico como mencionado acima, a sobreposição é em torno de 500 Hz.

[0081] Pode ser desejável implementar o banco de filtros A112 e/ou B122 para executar operações como ilustrado nas FIGURAS 4a e 4b em vários estágios. Por exemplo, a FIGURA 4c mostra um diagrama em blocos de uma implementação A114 do banco de filtros A112 que executa um equivalente funcional de operações de subamostragem e filtragem passa-alta utilizando uma série de interpolação, reamostragem, decimação e outras operações. Tal implementação pode ser mais fácil de projetar e/ou pode permitir reutilização de blocos funcionais de lógica e/ou código. Por exemplo, o mesmo bloco funcional pode ser utilizado para executar as operações de decimação para 14 kHz e decimação para 7 kHz como mostrado na FIGURA 4c. A operação de reversão espectral pode ser implementada multiplicando o sinal com a função  $e^{jn\pi}$  ou a seqüência  $(-1)^n$ , cujos valores se alternam entre +1 e -1. A operação de moldagem espectral pode ser implementada como um filtro passa-baixa configurado para moldar o sinal para obter uma resposta de filtro geral desejada.

[0082] As FIGURAS 33, 34a, 34b e 35a mostram respostas de impulso e frequência para exemplos de implementação, respectivamente, do filtro passa-baixa, interpolação para 34 kHz, reamostragem para 28 kHz, e a decimação para 14 kHz como mostrado na FIGURA 4c. A FIGURA 35b mostra respostas de impulso e frequência combinadas para aquelas implementações da interpolação para 34 kHz, a reamostragem para 28 kHz, e a decimação para 14 kHz. As FIGURAS 36a e 36b mostram respostas de impulso e frequência para exemplos de implementação, respectivamente, da decimação para 7 kHz e a operação de formatação espectral como mostrado na FIGURA 4c.

[0083] Observa-se que como consequência da operação de reversão espectral, o espectro de sinal de banda

alta S30 é revertido. Operações subseqüentes no encodificador e decodificador correspondente podem ser configuradas de acordo. Por exemplo, o gerador de excitação de banda alta A300 como descrito aqui, pode ser configurado para produzir um sinal de excitação de banda alta S120 que tem também uma forma revertida espectralmente.

[0084] A FIGURA 4d mostra um diagrama em blocos de uma implementação B124 do banco de filtros B122 que executa um equivalente funcional de operações de superamostragem e filtragem passa-alta utilizando uma série de operações de interpolação, reamostragem e outras. O banco de filtros B124 inclui uma operação de reversão espectral na banda alta que reverte uma operação similar como executada, por exemplo, em um banco de filtros do encodificador como banco de filtros A114. Nesse exemplo específico, o banco de filtros B124 também inclui filtros de corte (notch filter) na banda baixa e banda alta que atenuam um componente do sinal em 7100 Hz, embora tais filtros sejam opcionais e não necessitem ser incluídos.

[0085] As FIGURAS 37a e 37b mostram respostas de impulso e de frequência para exemplos de implementação, respectivamente, do filtro passa-baixa e filtro de corte de banda baixa como mostrado na FIGURA 4d. As FIGURAS 38a, 38b, 39a e 39b mostram respostas de impulso e frequência para exemplos de implementação, respectivamente, da interpolação para 14 kHz, interpolação para 28 kHz, a reamostragem para 16 kHz, e filtro de corte de banda alta como mostrado na FIGURA 4d.

[0086] O codificador de banca estreita A120 é implementado de acordo com um modelo de filtro de fonte que codifica o sinal de fala de entrada como (A) um conjunto de parâmetros que descrevem um filtro e (B) um sinal de excitação que aciona o filtro descrito para produzir uma

reprodução sintetizada do sinal de fala de entrada. A FIGURA 5a mostra um exemplo de um envelope espectral de um sinal de fala. Os picos que caracterizam esse envelope espectral representam ressonâncias do aparelho vocal e são denominados formantes. A maioria dos codificadores de fala codifica pelo menos essa estrutura espectral grosseira como um conjunto de parâmetros como coeficientes de filtro.

[0087] A FIGURA 5b mostra um exemplo de um arranjo de filtro-fonte básico como aplicado para codificação do envelope espectral de sinal de banda estreita S20. Um módulo de análise calcula um conjunto de parâmetros que caracterizam um filtro correspondendo ao som de fala durante um período de tempo (tipicamente 20 ms). Um filtro branqueador (também denominado um filtro de erro de predição ou análise) configurado de acordo com estes parâmetros de filtro remove o envelope espectral para aplainar espectralmente o sinal. O sinal branco resultante (também denominado um residual) tem menos energia e, desse modo, menos variância e é mais fácil de codificar do que o sinal de fala original. Os erros que resultam de codificação do sinal residual também podem ser espalhados mais uniformemente sobre o espectro. Os parâmetros de filtro e residual são tipicamente quantificados para transmissão eficiente através do canal. No decodificador, um filtro de síntese configurado de acordo com os parâmetros de filtro é excitado por um sinal baseado no residual para produzir uma versão sintetizada do som de fala original. O filtro de síntese é tipicamente configurado para ter uma função de transferência que é o inverso da função de transferência do filtro branqueador.

[0088] A FIGURA 6 mostra um diagrama em blocos de uma implementação básica A122 do codificador de banda estreita A120. Nesse exemplo, um módulo de análise de

codificação de predição linear (LPC) 210 encodifica o envelope espectral de sinal de banda estreita S20 como um conjunto de coeficientes de predição linear (LP) (por exemplo, coeficientes de um filtro todo de pólos  $1/A(z)$ ). O módulo de análise processa tipicamente o sinal de entrada como uma série de quadros de não sobreposição, com um novo conjunto de coeficientes sendo calculado para cada quadro. O período de quadro é genericamente um período através do qual se pode esperar que o sinal seja localmente estacionário; um exemplo comum é 20 milisegundos (equivalente a 160 amostras em uma taxa de amostragem de 8 kHz). Em um exemplo, o módulo de análise LPC 210 é configurado para calcular um conjunto de dez coeficientes de filtro LP para caracterizar a estrutura formante de cada quadro de 20 milisegundos. Também é possível implementar o módulo de análise para processar o sinal de entrada como uma série de quadros de sobreposição.

[0089] O módulo de análise pode ser configurado para analisar as amostras de cada quadro diretamente, ou as amostras podem ser ponderadas primeiramente de acordo com uma função de janelamento (por exemplo, uma janela Hamming). A análise também pode ser executada sobre uma janela que é maior do que o quadro, como uma janela de 30 ms. Essa janela pode ser simétrica (por exemplo, 5-20-5, de tal modo que inclui os 5 milisegundos imediatamente antes e após o quadro de 20 milisegundos) ou assimétrica (por exemplo, 10-20 de tal modo que inclui os últimos 10 milisegundos do quadro precedente). Um módulo de análise LPC é tipicamente configurado para calcular os coeficientes de filtro LP utilizando uma recursão de Levinson-Durbin ou algoritmo de Leroux-Gueguen. Em outra implementação, o módulo de análise pode ser configurado para calcular um conjunto de

coeficientes cepstrais para cada quadro em vez de um conjunto de coeficientes de filtro LP.

[0090] A taxa de saída do codificador A120 pode ser reduzida significativamente, com relativamente pouco efeito sobre a qualidade de reprodução, pela quantificação dos parâmetros de filtro. Coeficientes de filtro de predição linear são difíceis de quantificar eficientemente e são normalmente mapeados em outra representação, como pares de linha espectral (LSPs) ou frequências de linha espectral (LSFs), para quantificação e/ou encodificação de entropia. No exemplo da FIGURA 6, transformação de coeficiente de filtro LP em LSF 220 transforma o conjunto de coeficientes de filtro LP em um conjunto correspondente de LSFs. Outras representações de um-para-um de coeficientes de filtro LP incluem coeficientes PARCOR; valores de razão de área log; pares de imitância espectral (ISPs); e frequências de imitância espectral (ISFs) que são utilizados no codec GSM (Sistema Global para Comunicação Móvel) AMR-WB (Banda Larga Adaptativo de Múltiplas-Taxas). Tipicamente uma transformação entre um conjunto de coeficientes de filtro LP e um conjunto correspondente de LSFs é reversível, porém as modalidades também incluem implementações do codificador A120 no qual a transformação não é reversível sem erro.

[0091] O quantificador 230 é configurado para quantificar o conjunto de LSFs de banda estreita (ou outra representação de coeficiente) e o codificador de banda estreita A122 é configurado para transmitir o resultado dessa quantificação como os parâmetros de filtro de banda estreita S40. Tal quantificador inclui tipicamente um quantificador de vetor que encodifica o vetor de entrada como um índice para uma entrada de vetor correspondente em uma tabela ou livro-código.

[0092] Como visto na FIGURA 6, o encodificador de banda estreita A122 também gera um sinal residual pela passagem do sinal de banda estreita S20 através de um filtro branqueador 260 (também denominado um filtro de erro de predição ou análise) que é configurado de acordo com o conjunto de coeficientes de filtro. Nesse exemplo específico, o filtro branqueador 260 é implementado como um filtro FIR, embora implementações IIR também possam ser utilizadas. Esse sinal residual conterá tipicamente informações importantes de forma perceptível do quadro de fala, como estrutura de longo prazo em relação a pitch, que não é representado em parâmetros de filtro de banda estreita S40. O quantificador 270 é configurado para calcular uma representação quantizada desse sinal residual para saída como sinal de excitação de banda estreita codificado S50. Tal quantificador inclui tipicamente um quantificador de vetor que encodifica o vetor de entrada como um índice para uma entrada de vetor correspondente em uma tabela ou livro-código. Alternativamente, tal quantificador pode ser configurado para enviar um ou mais parâmetros a partir do qual o vetor pode ser gerado dinamicamente no decodificador, em vez de recuperado a partir de armazenamento, como em um método de livro-código disperso. Tal método é utilizado em esquemas de codificação como CELP algébrico (predição linear de excitação de livro-código) e codecs como 3GPP2 (Projeto 2 de Parceria da Terceira Geração) EVRC (Compressão e Descompressão a Taxa Variável Aperfeiçoada).

[0093] É desejável que o encodificador de banda estreita A120 gere o sinal de excitação de banda estreita encodificado de acordo com os mesmos valores de parâmetro de filtro que serão disponíveis para o decodificador de banda estreita correspondente. Desse modo, o sinal de excitação de banda estreita encodificado resultante já pode responder até

certo ponto por não idealidades nestes valores de parâmetro, como erro de quantificação. Por conseguinte, é desejável configurar o filtro branqueador utilizando os mesmos valores de coeficiente que estarão disponíveis no decodificador. No exemplo básico de encodificador A122 como mostrado na FIGURA 6, o quantificador inverso 240 desquantiza os parâmetros de codificação de banda estreita S40, a transformação de coeficiente de LSF-para filtro LP 250 mapeia os valores resultantes de volta para um conjunto correspondente de coeficientes de filtro LP, e esse conjunto de coeficientes é utilizado para configurar o filtro branqueador 260 para gerar o sinal residual que é quantificado pelo quantizador 270.

[0094] Algumas implementações de encodificador de banda estreita A120 são configuradas para calcular sinal de excitação de banda estreita encodificado S50 pela identificação de um entre um conjunto de vetores de livro-código que melhor associa com o sinal residual. Observa-se, entretanto, que o encodificador de banda estreita A120 também pode ser implementado para calcular uma representação quantizada do sinal residual sem gerar efetivamente o sinal residual. Por exemplo, o encodificador de banda estreita A120 pode ser configurado para utilizar um número de vetores de livro-código para gerar sinais sintetizados correspondentes (por exemplo, de acordo com um conjunto atual de parâmetros de filtro), e selecionar o vetor de livro-código associado ao sinal gerado que melhor associa com o sinal de banda estreita original S20 em um domínio ponderado de forma perceptível.

[0095] A FIGURA 7 mostra um diagrama em blocos de uma implementação B112 de decodificador de banda estreita B110. O quantizador inverso 310 desquantiza os parâmetros de filtro de banda estreita S40 (nesse caso, para um conjunto

de LSFs), e a transformação de coeficiente de LSF para filtro LP 320 transforma os LSFs em um conjunto de coeficientes de filtro (por exemplo, como descrito acima com referência ao quantificador inverso 240 e transformação 250 de codificador de banda estreita A122). O quantizador inverso 340 desquantiza o sinal residual de banda estreita S40 para produzir um sinal de excitação de banda estreita S80. Com base nos coeficientes de filtro e sinal de excitação de banda estreita S80, o filtro de síntese de banda estreita 330 sintetiza o sinal de banda estreita S90. Em outras palavras, o filtro de síntese de banda estreita 330 é configurado para moldar espectralmente o sinal de excitação de banda estreita S80 de acordo com os coeficientes de filtro desquantizados para produzir o sinal de banda estreita S90. O decodificador de banda estreita B112 também provê sinal de excitação de banda estreita S80 para um codificador de banda alta A200, que utiliza o mesmo para derivar o sinal de excitação de banda alta S120 como descrito aqui. Em algumas implementações como descrito abaixo, o decodificador de banda estreita B110 pode ser configurado para fornecer informações adicionais para o decodificador de banda alta B200 que se refere ao sinal de banda estreita, como inclinação espectral, ganho de pitch e atraso, e modo de fala.

[0096] O sistema do codificador de banda estreita A122 e decodificador de banda estreita B112 é um exemplo básico de um codec de fala de análise por síntese. A codificação de predição linear de excitação de livro-código (CELP) é uma família popular de codificação de análise-por-síntese, e implementações de tais codificadores podem executar codificação de forma de onda do residual, incluindo tais operações como seleção de entradas a partir de livros-código fixo e adaptável, operações de minimização de erro, e/ou operações de ponderação perceptível. Outras

implementações de codificação de análise por síntese incluem codificação de predição linear de excitação misturada (MELP), CELP algébrico (ACELP), CELP de relaxamento (RCELP), excitação de pulso regular (RPE), CELP de multi-pulsos (MPE), e predição linear excitada por soma de vetor (VSELP). Métodos de codificação relacionados incluem codificação de excitação de multi-bandas (BEM) e interpolação de forma de onda de protótipo (PWI). Os exemplos de codecs de fala de análise por síntese padronizados incluem o codec de taxa total ETSI (European Telecommunications Standards Institute)-GSM (GSM 06.10), que utiliza predição linear excitada residual (RELP); o codec de taxa total intensificado GSM (ETSI-GSM 06.60); o codificador ITU (International Telecommunication Union) padrão 11.8 kb/s G.729 Anexo E; os codecs IS (Padrão interino)-641 para IS-136 (um esquema de acesso múltiplo por divisão de tempo); os codecs de taxa múltipla adaptável GSM (GSM-AMR); e o codec 4GV<sup>TM</sup> (Vocoder<sup>TM</sup> de Quarta geração) (QUALCOMM Incorporated, San Diego, CA). O encodificador de banda estreita A120 e decodificador correspondente B110 podem ser implementados de acordo com qualquer uma destas tecnologias, ou qualquer outra tecnologia de codificação de fala (quer conhecida ou a ser desenvolvida) que representa um sinal de fala como (A) um conjunto de parâmetros que descrevem um filtro e (B) um sinal de excitação utilizado para acionar o filtro descrito para reproduzir o sinal de fala.

[0097] Mesmo após o filtro branqueador ter removido o envelope espectral grosseiro do sinal de banda estreita S20, uma quantidade considerável de estrutura harmônica fina pode permanecer, especialmente para fala vozeada. A FIGURA 8a mostra um gráfico espectral de um exemplo de um sinal residual, como pode ser produzido por um filtro branqueador, para um sinal vozeado como uma vogal. A

estrutura periódica visível nesse exemplo é relacionada a pitch, e diferentes sons vozeados falados pela mesma pessoa podem ter diferentes estruturas formantes, porém estruturas de pitch similares. A FIGURA 8b mostra um gráfico de domínio de tempo de um exemplo de tal sinal residual que mostra uma seqüência de pulsos de pitch no tempo.

[0098] A eficiência de codificação e/ou qualidade de fala pode ser aumentada utilizando um ou mais valores de parâmetro para encodificar características da estrutura de pitch. Uma característica importante da estrutura de pitch é a frequência da primeira harmônica (também denominada a frequência fundamental), que é tipicamente na faixa de 60 a 400 Hz. Essa característica é tipicamente codificada como o inverso da frequência fundamental, também denominado atraso de pitch (atraso de pitch). O atraso de pitch indica o número de amostras em um período de pitch e pode ser codificado como um ou mais índices de livro-código. Sinais de fala de homens tendem a ter atraso de pitch maiores do que sinais de fala de mulheres.

[0099] Outra característica de sinal referente à estrutura de pitch é a periodicidade, que indica a intensidade da estrutura harmônica ou, em outras palavras, o grau no qual o sinal é harmônico ou não harmônico. Dois indicadores típicos de periodicidade são cruzamentos zero e funções de autocorrelação normalizadas (NACFs). A periodicidade também pode ser indicada pelo ganho de pitch, que é comumente codificado como um ganho de livro-código (por exemplo, um ganho de livro-código adaptável quantizado).

[0100] O encodificador de banda estreita A120 pode incluir um ou mais módulos configurados para encodificar a estrutura harmônica de sinal de banda estreita S20 de longo

prazo. Como mostrado na FIGURA 9, um paradigma CELP típico que pode ser utilizado inclui um módulo de análise LPC de malha aberta (open-loop), que codifica as características de curto prazo ou envelope espectral grosseiro, seguido por um estágio de análise de predição de longo prazo de malha fechada (closed-loop), que codifica a estrutura harmônica ou pitch fino. As características de curto prazo são encodificadas como coeficientes de filtro, e as características de longo prazo são encodificadas como valores para parâmetros como um atraso de pitch e ganho de pitch. Por exemplo, o encodificador de banda estreita A120 pode ser configurado para transmitir sinal de excitação de banda estreita encodificado S50 em uma forma que inclui um ou mais índices de livro-código (por exemplo, um índice de livro-código fixo e um índice de livro-código adaptável) e valores de ganho correspondentes. O cálculo dessa representação quantizada do sinal residual de banda estreita (por exemplo, por quantizador 270) pode incluir a seleção de tais índices e cálculo destes valores. A codificação da estrutura de pitch pode incluir também interpolação de uma forma de onda de protótipo de pitch, cuja operação pode incluir calcular uma diferença entre pulsos de pitch sucessivos. A modelagem da estrutura de longo prazo pode ser desabilitada para quadros correspondendo à fala não-vozeada, que é tipicamente como ruído e não estruturada.

[0101] Uma implementação de decodificador de banda estreita B110 de acordo com um paradigma como mostrado na FIGURA 9 pode ser configurada para transmitir sinal de excitação de banda estreita S80 para decodificador de banda alta B200 após recuperação da estrutura de longo prazo (estrutura harmônica ou pitch). Por exemplo, tal decodificador pode ser configurado para transmitir sinal de excitação de banda estreita S80 como uma versão desquantizada

de sinal de excitação de banda estreita encodificado S50. Evidentemente, também é possível implementar o decodificador de banda estreita B110 de tal modo que o decodificador de banda alta B200 execute desquantização de sinal de excitação de banda estreita encodificado S50 para obter sinal de excitação de banda estreita S80.

[0102] Em uma implementação de encodificador de fala de banda larga A100 de acordo com um paradigma como mostrado na FIGURA 9, o encodificador de banda alta A200 pode ser configurado para receber o sinal de excitação de banda estreita como produzido pela análise de curto prazo ou filtro branqueador. Em outras palavras, o encodificador de banda estreita A120 pode ser configurado para transmitir o sinal de excitação de banda estreita para o encodificador de banda alta A200 antes de encodificar a estrutura de longo prazo. É desejável, entretanto, que o encodificador de banda alta A200 receba a partir do canal de banda estreita a mesma informação de codificação que será recebida pelo decodificador de banda alta B200, de tal modo que os parâmetros de codificação produzidos pelo encodificador de banda alta A200 já possam responder até certo ponto por não idealidades nesta informação. Desse modo, pode ser preferível que o encodificador de banda alta A200 reconstrua o sinal de excitação de banda estreita S80 a partir do mesmo sinal de excitação de banda estreita encodificado parametrizado e/ou quantizado S50 para ser transmitido pelo encodificador de fala de banda larga A100. Uma vantagem potencial dessa abordagem é o cálculo mais preciso dos fatores de ganho de banda alta S60b descritos abaixo.

[0103] Além dos parâmetros que caracterizam a estrutura de curto prazo e/ou longo prazo de sinal de banda estreita S20, o encodificador de banda estreita A120 pode produzir valores de parâmetro que se referem a outras

características de sinal de banda estreita S20. Esses valores, que podem ser adequadamente quantizados para saída pelo codificador de fala de banda larga A100, podem ser incluídos entre os parâmetros de filtro de banda estreita S40 ou transmitidos separadamente. O codificador de banda alta A200 pode ser também configurado para calcular parâmetros de codificação de banda alta S60 de acordo com um ou mais desses parâmetros adicionais (por exemplo, após desquantificação). No decodificador de fala de banda larga B100, o decodificador de banda alta B200 pode ser configurado para receber os valores de parâmetro através do decodificador de banda estreita B110 (por exemplo, após desquantificação). Alternativamente, o decodificador de banda alta B200 pode ser configurado para receber (e possivelmente desquantificar) os valores de parâmetro diretamente.

[0104] Em um exemplo de parâmetros de codificação de banda estreita adicionais, o codificador de banda estreita A120 produz valores para inclinação espectral e parâmetros de modo de fala para cada quadro. Inclinação espectral se refere ao formato do envelope espectral sobre a banda passante e é tipicamente representado pelo primeiro coeficiente de reflexão quantizado. Para a maioria dos sons de voz, a energia espectral diminui com frequência crescente, de tal modo que o primeiro coeficiente de reflexão é negativo e pode se aproximar de -1. A maioria dos sons não-vozeados tem um espectro que também é plano, de tal modo que o primeiro coeficiente de reflexão esteja próximo a zero, ou tenha mais energia em frequências elevadas, de tal modo que o primeiro coeficiente de reflexão é positivo e pode se aproximar de +1.

[0105] O modo de fala (também denominado modo de voz) indica se o quadro atual representa fala vozeada ou não vozeada. Esse parâmetro pode ter um valor binário baseado

em uma ou mais medições de periodicidade (por exemplo, cruzamentos zero, NACFs, ganho de pitch) e/ou atividade de voz para o quadro, como uma relação entre tal medição e um valor limite. Em outras implementações, o parâmetro de modo de fala tem um ou mais outros estados para indicar modos como silêncio ou ruído de fundo, ou uma transição entre silêncio e fala vozeada.

[0106] O encodificador de banda alta A200 é configurado para codificar sinal de banda alta S30 de acordo com um modelo de filtro-fonte, com a excitação para esse filtro sendo baseada no sinal de excitação de banda estreita encodificado. A FIGURA 10 mostra um diagrama em blocos de uma implementação A202 do encodificador de banda alta A200 que é configurado para produzir um fluxo de parâmetros de codificação de banda alta S60 incluindo parâmetros de filtro de banda alta S60a e fatores de ganho de banda alta S60b. O gerador de excitação de banda alta A300 deriva um sinal de excitação de banda alta S120 a partir do sinal de excitação de banda estreita encodificado S50. O módulo de análise A210 produz um conjunto de valores de parâmetro que caracterizam o envelope espectral de sinal de banda alta S30. Nesse exemplo específico, o módulo de análise A210 é configurado para executar análise LPC a fim de produzir um conjunto de coeficientes de filtro LP para cada quadro de sinal de banda alta S30. A transformação de coeficiente de filtro de predição linear em LSF 410 transforma o conjunto de coeficientes de filtro LP em um conjunto correspondente de LSFs. Como observado acima com referência ao módulo de análise 210 e transformação 220, o módulo de análise A210 e/ou transformação 410 pode ser configurado para utilizar outros conjuntos de coeficiente (por exemplo, coeficientes cepstrais) e/ou representações de coeficiente (por exemplo, ISPs).

[0107] O quantizador 420 é configurado para quantizar o conjunto de LSFs de banda alta (ou outra representação de coeficiente, como ISPs) e o encodificador de banda alta A202 é configurado para transmitir o resultado dessa quantização como os parâmetros de filtro de banda alta S60a. Tal quantizador inclui tipicamente um quantizador de vetor que codifica o vetor de entrada como um índice para uma entrada de vetor correspondente em uma tabela ou livro-código.

[0108] O encodificador de banda alta A202 também inclui um filtro de síntese A220 configurado para produzir um sinal de banda alta sintetizado S130 de acordo com o sinal de excitação de banda alta S120 e o envelope espectral codificado (por exemplo, o conjunto de coeficientes de filtro LP) produzido por módulo de análise A210. O filtro de síntese A220 é tipicamente implementado como um filtro IIR, embora implementações FIR também possam ser utilizadas. Em um exemplo específico, o filtro de síntese A220 é implementado como um filtro auto-regressivo linear de sexta ordem.

[0109] O calculador de fator de ganho de banda alta A230 calcula uma ou mais diferenças entre os níveis do sinal de banda alta original S30 e sinal de banda alta sintetizado S130 para especificar um envelope de ganho para o quadro. O quantizador 430, que pode ser implementado como um quantizador de vetor que encodifica o vetor de entrada como um índice para uma entrada de vetor correspondente em uma tabela ou livro-código, quantiza o valor ou valores especificando o envelope de ganho, e o encodificador de banda alta A202 é configurado para transmitir o resultado dessa quantização como fatores de ganho de banda alta S60b.

[0110] Em uma implementação como mostrado na FIGURA 10, o filtro de síntese A220 é disposto para receber

os coeficientes de filtro a partir do módulo de análise A210. Uma implementação alternativa do codificador de banda alta A202 inclui um quantizador inverso e transformação inversa configurada para decodificar os coeficientes de filtro a partir de parâmetros de filtro de banda alta S60a, e, nesse caso, o filtro de síntese A220 é disposto para receber os coeficientes de filtro decodificados. Tal arranjo alternativo pode suportar um cálculo mais preciso do envelope de ganho pelo calculador de ganho de banda alta A230.

[0111] Em um exemplo específico, o módulo de análise A210 e o calculador de ganho de banda alta A230 transmitem um conjunto de seis LSFs e um conjunto de cinco valores de ganho por quadro, respectivamente, de tal modo que uma extensão de banda larga do sinal de banda estreita S20 pode ser obtida com apenas onze valores adicionais por quadro. O ouvido tende a ser menos sensível a erros de frequência em frequências altas, de tal modo que a codificação de banda alta em uma ordem de LPC baixa pode produzir um sinal tendo uma qualidade perceptível comparável com a codificação de banda estreita em ordem de LPC mais elevada. Uma implementação típica do codificador de banda alta A200 pode ser configurada para transmitir 8 a 12 bits por quadro para reconstrução de alta qualidade do envelope espectral e outros 8 a 12 bits por quadro para reconstrução de alta qualidade do envelope temporal. Em outro exemplo específico, o módulo de análise A210 transmite um conjunto de oito LSFs por quadro.

[0112] Algumas implementações do codificador de banda alta A200 são configuradas para produzir um sinal de excitação de banda alta S120 pela geração de um sinal de ruído aleatório tendo componentes de frequência de banda alta e modulação por amplitude do sinal de ruído de acordo com o envelope de domínio de tempo do sinal de banda estreita

S20, sinal de excitação de banda estreita S80, ou sinal de banda alta S30. Embora tal método com base em ruído possa produzir resultados adequados para sons não vozeados, entretanto, pode não ser desejável para sons vozeados, cujos residuais são normalmente harmônicos e conseqüentemente têm alguma estrutura periódica.

[0113] O gerador de excitação de banda alta A300 é configurado para gerar sinal de excitação de banda alta S120 estendendo o espectro de sinal de excitação de banda estreita S80 para dentro da faixa de frequência de banda alta. A FIGURA 11 mostra um diagrama em blocos de uma implementação A302 de gerador de excitação de banda alta A300. O quantizador inverso 450 é configurado para desquantizar o sinal de excitação de banda estreita encodificado S50 a fim de produzir o sinal de excitação de banda estreita S80. O extensor de espectro A400 é configurado para produzir um sinal harmonicamente estendido S160 com base no sinal de excitação de banda estreita S80. O combinador 470 é configurado para combinar um sinal de ruído aleatório gerado pelo gerador de ruído 480 e um envelope de domínio de tempo calculado pelo calculador de envelope 460 para produzir um sinal de ruído modulado S170. O combinador 490 é configurado para misturar o sinal harmonicamente estendido S60 e o sinal de ruído modulado S170 para produzir um sinal de excitação de banda alta S120.

[0114] Em um exemplo, o extensor de espectro A400 é configurado para executar uma operação de dobramento espectral (também denominado espelhamento) sobre o sinal de excitação de banda estreita S80 para produzir um sinal harmonicamente estendido S160. Dobramento espectral pode ser executado por um sinal de excitação de enchimento zero S80 e então aplicando um filtro passa-alta para reter alias. Em outro exemplo, um extensor de espectro A400 é configurado

para produzir um sinal harmonicamente estendido S160 por tradução espectralmente do sinal de excitação de banda estreita S80 para banda alta (por exemplo, através de superamostragem seguido por multiplicação com um sinal cosseno de frequência constante).

[0115] Métodos de tradução e dobramento espectral podem produzir sinais estendidos espectralmente cuja estrutura harmônica é descontínua com a estrutura harmônica original do sinal de excitação de banda estreita S80 em fase e/ou frequência. Por exemplo, tais métodos podem produzir sinais tendo picos que não são geralmente localizados em múltiplos da frequência fundamental, que podem causar pequenos artefatos sonoros no sinal de fala reconstruído. Esses métodos tendem também a produzir harmônica de alta frequência que têm características tonais não naturalmente fortes. Além disso, como um sinal PSTN pode ser amostrado em 8 kHz porém limitado por banda a não mais do que 3400 Hz, o espectro superior do sinal de excitação de banda estreita S80 pode conter pouca ou nenhuma energia, de tal modo que um sinal estendido gerado de acordo com uma operação de tradução espectral ou dobramento espectral pode ter um vazio espectral acima de 3400 Hz.

[0116] Outros métodos para gerar um sinal harmonicamente estendido S160 incluem identificar uma ou mais frequências fundamentais do sinal de excitação de banda estreita S80 e gerar tons harmônicos de acordo com a informação. Por exemplo, a estrutura harmônica de um sinal de excitação pode ser caracterizada pela frequência fundamental juntamente com informações de fase e amplitude. Outra implementação do gerador de excitação de banda alta A300 gera um sinal harmonicamente estendido S160 com base na amplitude e frequência fundamentais (como indicado, por exemplo, pelo atraso de pitch e ganho de pitch). A menos que

o sinal estendido harmonicamente seja coerente em fase com o sinal de excitação de banda estreita S80, entretanto, a qualidade da fala decodificada resultante pode não ser aceitável.

[0117] Uma função não-linear pode ser utilizada para criar um sinal de excitação de banda alta que é coerente em fase com a excitação de banda estreita e conserva a estrutura harmônica sem descontinuidade de fase. Uma função não-linear também pode fornecer um nível de ruído aumentado entre harmônica de alta frequência, que tende a soar mais natural do que a harmônica de alta frequência tonal produzida por métodos como dobramento espectral e tradução espectral. Funções não-lineares sem memória, típicas que podem ser aplicadas por várias implementações de extensor de espectro A400 incluem a função de valor absoluto (também denominada retificação de onda completa), retificação de meia-onda, quadrática, cúbica, e de recorte. Outras implementações de extensor de espectro A400 podem ser configuradas para aplicar uma função não-linear tendo memória.

[0118] A FIGURA 12 é um diagrama em blocos de uma implementação A402 de extensor de espectro A400 que é configurado para aplicar uma função não-linear a fim de estender o espectro de sinal de excitação de banda estreita S80. O sobreamostrador 510 é configurado para sobreamostrar o sinal de excitação de banda estreita S80. Pode ser desejável sobreamostrar o sinal suficientemente para minimizar aliasing após aplicação da função não-linear. Em um exemplo específico, o sobreamostrador 510 sobreamostra o sinal por um fator de oito. O sobreamostrador 510 pode ser configurado para executar a operação de superamostragem por enchimento zero do sinal de entrada e filtragem passa-baixa do resultado. O calculador de função não-linear 520 é configurado para aplicar uma função não-linear ao sinal

superamostrado. Uma vantagem potencial da função de valor absoluto em relação a outras funções não-lineares para extensão espectral, como a quadrática, é que a normalização de energia não é necessária. Em algumas implementações, a função de valor absoluto pode ser aplicada eficientemente por extração ou apagamento do bit de sinal de cada amostra. O calculador de função não-linear 520 pode ser também configurado para executar uma distorção de amplitude do sinal superamostrado ou estendido espectralmente.

[0119] O subamostrador 530 é configurado para subamostrar o resultado estendido espectralmente da aplicação da função não-linear. Pode ser desejável que o subamostrador 530 execute uma operação de filtragem de passa-faixa para selecionar uma banda de frequência desejada do sinal estendido espectralmente antes de reduzir a taxa de amostragem (por exemplo, para reduzir ou evitar aliasing ou corrupção por uma imagem indesejável). Pode ser também desejável que o subamostrador 530 reduza a taxa de amostragem em mais de um estágio.

[0120] Na FIGURA 12a é um diagrama que mostra os espectros de sinais em vários pontos em um exemplo de uma operação de extensão espectral, onde a escala de frequência é igual através de vários gráficos. O gráfico (a) mostra o espectro de um exemplo de sinal de excitação de banda estreita S80. O gráfico (b) mostra o espectro após o sinal S80 ter sido superamostrado por um fator de oito. O gráfico (c) mostra um exemplo do espectro estendido após aplicação de uma função não-linear. O gráfico (d) mostra o espectro após filtragem passa-baixa. Nesse exemplo, a banda passante estende até o limite de frequência superior do sinal de banda alta S30 (por exemplo, 7kHz ou 8 kHz).

[0121] O gráfico (e) mostra o espectro após um primeiro estágio de subamostragem, no qual a taxa de

amostragem é reduzida por um fator de quatro para obter um sinal de banda larga. O gráfico (f) mostra o espectro após uma operação de filtragem passa-alta para selecionar a parte de banda alta do sinal estendido, e o gráfico (g) mostra o espectro após um segundo estágio de subamostragem, no qual a taxa de amostragem é reduzida por um fator de dois. Em um exemplo específico, o subamostrador 530 executa a filtragem passa-alta e o segundo estágio de subamostragem pela passagem do sinal de banda larga através do filtro passa-alta 130 e o subamostrador 140 do banco de filtros A112 (ou outras estruturas ou rotinas tendo a mesma resposta) para produzir um sinal estendido espectralmente tendo a faixa de frequência e a taxa de amostragem do sinal de banda alta S30.

[0122] Como pode ser visto no gráfico (g), a subamostragem do sinal passa-alta mostrado no gráfico (f) causa uma reversão de seu espectro. Nesse exemplo, o subamostrador 530 é também configurado para executar uma operação de inversão espectral no sinal. O gráfico (h) mostra um resultado da aplicação da operação de inversão espectral, que pode ser executada pela multiplicação do sinal com a função  $e^{jn\pi}$  ou a seqüência  $(-1)^n$ , cujos valores alternam entre +1 e -1. Tal operação é equivalente ao deslocamento do espectro digital do sinal no domínio de frequência por uma distância de  $\pi$ . Observa-se que o mesmo resultado pode também ser obtido pela aplicação de operações de inversão espectral e subamostragem em uma ordem diferente. As operações de superamostragem e/ou subamostragem também podem ser configuradas para incluir reamostragem a fim de obter um sinal estendido espectralmente tendo a taxa de amostragem de sinal de banda alta S30 (por exemplo, 7 kHz).

[0123] Como observado acima, bancos de filtros A110 e B120 podem ser implementados de tal modo que um ou ambos os sinais de banda estreita e banda alta S20, S30 tem

uma forma revertida espectralmente na saída do banco de filtros A110, é codificado e decodificado na forma revertida espectralmente, e é novamente revertido espectralmente no banco de filtros B120 antes de ser transmitido em sinal de fala de banda larga S110. Em tal caso, evidentemente, uma operação de inversão espectral como mostrado na FIGURA 12a não seria necessária, visto que seria desejável para o sinal de excitação de banda alta S120 ter também uma forma revertida espectralmente.

[0124] As várias tarefas de superamostragem e subamostragem de uma operação de extensão espectral como executado por extensor de espectro A402 podem ser configuradas e dispostas em muitos modos diferentes. Por exemplo, a FIGURA 12b é um diagrama que mostra os espectros de sinais em vários pontos em outro exemplo de uma operação de extensão espectral, onde a escala de frequência é igual através dos vários gráficos. O gráfico (a) mostra o espectro de um exemplo de sinal de excitação de banda estreita S80. O gráfico (b) mostra o espectro após o sinal S80 ter sido superamostrado por um fator de dois. O gráfico (c) mostra um exemplo do espectro estendido após aplicação de uma função não-linear. Neste caso, aliasing que pode ocorrer nas frequências mais elevadas é aceito.

[0125] O gráfico (d) mostra o espectro após uma operação de reversão espectral. O gráfico (e) mostra o espectro após um único estágio de subamostragem, no qual a taxa de amostragem é reduzida por um fator de dois para obter o sinal estendido espectralmente, desejado. Nesse exemplo, o sinal está na forma revertida espectralmente e pode ser utilizado em uma implementação de codificador de banda alta A200 que processou o sinal de banda alta S30 em tal forma.

[0126] O sinal estendido espectralmente produzido pelo calculador de função não-linear 520 provavelmente tem uma queda acentuada em amplitude à medida que a frequência aumenta. O extensor espectral A402 inclui um aplainador espectral 540 configurado para executar uma operação de branqueamento no sinal subamostrado. O aplainador espectral 540 pode ser configurado para executar uma operação de branqueamento fixa ou executar uma operação de branqueamento adaptável. Em um exemplo específico de branqueamento adaptável, o aplainador espectral 540 inclui um módulo de análise LPC configurado para calcular um conjunto de quatro coeficientes de filtro a partir do sinal subamostrado e um filtro de análise de quarta ordem configurado para branquear o sinal de acordo com esses coeficientes. Outras implementações de extensor de espectro A400 incluem configurações nas quais o aplainador espectral 540 opera no sinal estendido espectralmente antes do subamostrador 530.

[0127] O gerador de excitação de banda alta A300 pode ser implementado para transmitir o sinal estendido harmonicamente S160 como o sinal de excitação de banda alta S120. Em alguns casos, entretanto, o uso somente de um sinal harmonicamente estendido como a excitação de banda alta pode resultar em artefatos audíveis. A estrutura harmônica de fala é geralmente menos acentuada na banda alta do que na banda baixa, e o uso de estrutura harmônica em demasia no sinal de excitação de banda alta pode resultar em um som de zumbido. Esse artefato pode ser especialmente perceptível em sinais de fala de mulheres.

[0128] As modalidades incluem implementações de gerador de excitação de banda alta A300 que são configuradas para misturar o sinal estendido harmonicamente S160 com um sinal de ruído. Como mostrado na FIGURA 11, o gerador de

excitação de banda alta A302 inclui um gerador de ruído 480 que é configurado para produzir um sinal de ruído aleatório. Em um exemplo, o gerador de ruído 480 é configurado para produzir um sinal de ruído pseudoaleatório branco de variância unitária, embora em outras implementações o sinal de ruído não necessite ser branco e possa ter uma densidade de potência que varia com a frequência. Pode ser desejável que o gerador de ruído 480 seja configurado para transmitir o sinal de ruído como função determinista de tal modo que seu estado possa ser duplicado no decodificador. Por exemplo, o gerador de ruído 480 pode ser configurado para transmitir o sinal de ruído como uma função determinista de informações codificadas mais cedo dentro do mesmo quadro, como os parâmetros de filtro de banda estreita S40 e/ou sinal de excitação de banda estreita encodificado S50.

[0129] Antes de ser misturado com o sinal harmonicamente estendido S160, o sinal de ruído aleatório produzido pelo gerador de ruído 480 pode ser modulado em amplitude para ter um envelope de domínio de tempo que se aproxima da distribuição de energia no tempo do sinal de banda estreita S20, sinal de banda alta S30, sinal de excitação de banda estreita S80, ou sinal harmonicamente estendido S160. Como mostrado na FIGURA 11, o gerador de excitação de banda alta A302 inclui um combinador 470 configurado para modular em amplitude o sinal de ruído produzido pelo gerador de ruído 480 de acordo com um envelope de domínio de tempo calculado pelo calculador de envelope 460. Por exemplo, o combinador 470 pode ser implementado como um multiplicador disposto para escalar a saída do gerador de ruído 480 de acordo com o envelope de domínio de tempo calculado pelo calculador de envelope 460 para produzir um sinal de ruído modulado S170.

[0130] Em uma implementação A304 do gerador de excitação de banda alta A302, como mostrado no diagrama em blocos da FIGURA 13, o calculador de envelope 460 é disposto para calcular o envelope do sinal harmonicamente estendido S160. Em uma implementação A306 do gerador de excitação de banda alta A302, como mostrado no diagrama de blocos da FIGURA 14, o calculador de envelope 460 é disposto para calcular o envelope do sinal de excitação de banda estreita S80. Implementações adicionais do gerador de excitação de banda alta A302 podem ser, de outro modo, configurados para acrescentar ruído ao sinal harmonicamente estendido S160 de acordo com locais dos pulsos de pitch de banda estreita no tempo.

[0131] O calculador de envelope 460 pode ser configurado para executar um cálculo de envelope como uma tarefa que inclui uma série de subtarefas. A FIGURA 15 mostra um fluxograma de um exemplo T100 dessa tarefa. A subtarefa T110 calcula o quadrado de cada amostra do quadro do sinal cujo envelope deve ser modelado (por exemplo, o sinal de excitação de banda estreita S80 ou o sinal harmonicamente estendido S160) para produzir uma seqüência de valores quadráticos. A subtarefa T120 executa uma operação de suavização na seqüência de valores quadráticos. Em um exemplo, a subtarefa T120 aplica um filtro passa-baixa IIR de primeira ordem na seqüência de acordo com a expressão:

$$Y(n) = ax(n) + (1-a)y(n-1), \quad (1)$$

onde  $x$  é a entrada de filtro,  $y$  é a saída de filtro,  $n$  é um índice de domínio de tempo, e  $a$  é um coeficiente de suavização tendo um valor entre 0,5 e 1. O valor do coeficiente de suavização  $a$  pode ser fixo ou, em uma implementação alternativa, pode ser adaptável de acordo com uma indicação de ruído no sinal de entrada, de tal modo que

a esteja mais próximo a 1 na ausência de ruído e mais próximo a 0,5 na presença de ruído. A subtarefa T130 aplica uma função de raiz quadrada a cada amostra da seqüência suavizada para produzir o envelope de domínio de tempo.

[0132] Tal implementação de calculador de envelope 460 pode ser configurada para executar as diversas subtarefas da tarefa T100 em modo serial e/ou paralelo. Em implementações adicionais da tarefa T100, a subtarefa T110 pode ser precedida por uma operação de passa-faixa configurada para selecionar uma parte de frequência desejada do sinal cujo envelope deve ser modelado, como a faixa de 3-4 kHz.

[0133] O combinador 490 é configurado para misturar o sinal harmonicamente estendido S160 e o sinal de ruído modulado S170 para produzir o sinal de excitação de banda alta S120. As implementações do combinador 490 podem ser configuradas, por exemplo, para calcular sinal de excitação de banda alta S120 como uma soma de sinal harmonicamente estendido S160 e sinal de ruído modulado S170. Tal implementação do combinador 490 pode ser configurada para calcular o sinal de excitação de banda alta S120 como uma soma ponderada pela aplicação de um fator de ponderação em sinal harmonicamente estendido S160 e/ou em sinal de ruído modulado S170 antes da soma. Cada fator de ponderação pode ser calculado de acordo com um ou mais critérios e pode ser um valor fixo ou, alternativamente, um valor adaptável que é calculado com base em quadro por quadro ou subquadro por subquadro.

[0134] A FIGURA 16 mostra um diagrama em blocos de uma implementação 492 de combinador 490 que é configurado para calcular sinal de excitação de banda alta S120 como uma soma ponderada de sinal estendido harmonicamente S160 e sinal de ruído modulado S170. O combinador 492 é configurado para

ponderar o sinal harmonicamente estendido S160 de acordo com o fator de ponderação harmônico S180, para ponderar o sinal de ruído modulado S170 de acordo com o fator de ponderação de ruído S190, e transmitir o sinal de excitação de banda alta S120 como uma soma dos sinais ponderados. Nesse exemplo, o combinador 492 inclui um calculador de fator de ponderação 550 que é configurado para calcular o fator de ponderação harmônica S180 e o fator de ponderação de ruído S190.

[0135] O calculador de fator de ponderação 550 pode ser configurado para calcular fatores de ponderação S180 e S190 de acordo com uma razão desejada de conteúdo de harmônica para conteúdo de ruído em sinal de excitação de banda alta S120. Por exemplo, pode ser desejável para o combinador 492 produzir sinal de excitação de banda alta S120 para ter uma razão de energia harmônica para energia de ruído similar àquela do sinal de banda alta S30. Em algumas implementações do calculador de fator de ponderação 550, fatores de ponderação S180, S190 são calculados de acordo com um ou mais parâmetros referentes a uma periodicidade de sinal de banda estreita S20 ou do sinal residual de banda estreita, como ganho de pitch e/ou modo de fala. Tal implementação de calculador de fator de ponderação 550 pode ser configurada para atribuir um valor ao fator de ponderação harmônica S180 que é proporcional ao ganho de pitch, por exemplo, e/ou para atribuir um valor ao fator de ponderação de ruído S190 para sinais de fala não-vozeados do que para sinais de fala vozeados.

[0136] Em outras implementações, o calculador de fator de ponderação 550 é configurado para calcular valores para fator de ponderação de harmônica S180 e/ou fator de ponderação de ruído S190 de acordo com uma medição de periodicidade de sinal de banda alta S30. Em tal exemplo, o calculador de fator de ponderação 550 calcula o fator de

ponderação de harmônica S180 como o valor máximo do coeficiente de autocorrelação de sinal de banda alta S30 para o quadro ou subquadro atual, onde a autocorrelação é executada sobre uma faixa de busca que inclui um retardo de um atraso de pitch e não inclui um retardo de amostras zero. A FIGURA 17 mostra um exemplo de uma faixa de busca de comprimento de  $n$  amostras que é centrado em torno de um retardo de um atraso de pitch e tem uma largura não maior do que um atraso de pitch.

[0137] A FIGURA 17 também mostra um exemplo de outra abordagem na qual o calculador de fator de ponderação 550 calcula uma medição de periodicidade de sinal de banda alta S30 em vários estágios. Em um primeiro estágio, o quadro atual é dividido em um número de subquadros, e o retardo para o qual o coeficiente de autocorrelação é máximo é identificado separadamente para cada subquadro. Como mencionado acima, a autocorrelação é executada sobre uma faixa de busca que inclui um retardo de um atraso de pitch e não inclui um retardo de amostras zero.

[0138] Em um segundo estágio, um quadro retardado é construído pela aplicação do retardo identificado correspondente em cada subquadro, concatenando os subquadros resultantes para construir um quadro retardado de forma ótima, e calculando o fator de ponderação de harmônica S180 como o coeficiente de correlação entre o quadro original e o quadro retardado de forma ótima. Em uma alternativa adicional, o calculador de fator de ponderação 550 calcula o fator de ponderação de harmônica S180 como uma média dos coeficientes de autocorrelação máximos obtidos no primeiro estágio para cada subquadro. As implementações de calculador de fator de ponderação 550 também podem ser configuradas para escalar o coeficiente de correlação e/ou

para combinar o mesmo com outro valor, para calcular o valor para o fator de ponderação de harmônica S180.

[0139] Pode ser desejável que o calculador de fator de ponderação 550 calcule uma medição de periodicidade de sinal de banda alta S30 somente em casos onde uma presença de periodicidade no quadro é, de outro modo, indicada. Por exemplo, o calculador de fator de ponderação 550 pode ser configurado para calcular uma medição de periodicidade de sinal de banda alta S30 de acordo com uma relação entre outro indicador de periodicidade do quadro atual, como ganho de pitch, e um valor limite. Em um exemplo, o calculador de fator de ponderação 550 é configurado para executar uma operação de autocorrelação no sinal de banda alta S30 somente se o ganho de pitch de quadro (por exemplo, o ganho de livro-código adaptável do residual de banda estreita) tiver um valor maior do que 0,5 (alternativamente, pelo menos 0,5). Em outro exemplo, o calculador de fator de ponderação 550 é configurado para executar uma operação de autocorrelação no sinal de banda alta S30 somente para quadros tendo estados específicos de modo de fala (por exemplo, somente para sinais com voz). Em tais casos, o calculador de fator de ponderação 550 pode ser configurado para atribuir um fator de ponderação default para quadros tendo outros estados de modo de fala e/ou valores menores de ganho de pitch.

[0140] As modalidades incluem implementações adicionais de calculador de fator de ponderação 550 que são configuradas para calcular fatores de ponderação de acordo com características diferentes de ou além da periodicidade. Por exemplo, tal implementação pode ser configurada para atribuir um valor mais elevado ao fator de ganho de ruído S190 para sinais de fala tendo um atraso de pitch grande do que para sinais de fala tendo um atraso de pitch pequeno. Outra implementação do calculador de fator de ponderação 550

é configurada para determinar uma medição de harmonicidade de sinal de fala de banda larga S10, ou de sinal de banda alta S30, de acordo com uma medição da energia do sinal em múltiplos da frequência fundamental em relação à energia do sinal nos outros componentes de frequência.

[0141] Algumas implementações do encodificador de fala de banda larga A100 são configuradas para transmitir uma indicação de periodicidade ou harmonicidade (por exemplo, um bit de sinalização único indicando se o quadro é harmônico ou não harmônico) com base no ganho de pitch e/ou outra medição de periodicidade ou harmonicidade como descrito aqui. Em um exemplo, um decodificador de fala de banda larga correspondente B100 utiliza essa indicação para configurar uma operação como um cálculo de fator de ponderação. Em outro exemplo, tal indicação é utilizada no encodificador e/ou decodificador no cálculo de um valor para um parâmetro de modo de fala.

[0142] Pode ser desejável que o gerador de excitação de banda alta A302 gere um sinal de excitação de banda alta S120 de tal modo que a energia do sinal de excitação seja substancialmente não afetada pelos valores específicos de fatores de ponderação S180 e S190. Em tal caso, o calculador de fator de ponderação 550 pode ser configurado para calcular um valor para o fator de ponderação de harmônica S180 ou para o fator de ponderação de ruído S190 (ou para receber tal valor a partir de armazenamento ou outro elemento de encodificador de banda alta A200) e derivar um valor para o outro fator de ponderação de acordo com a expressão a seguir:

$$(W_{\text{harmônica}})^2 + (W_{\text{ruído}})^2 = 1, \quad (2)$$

onde  $W_{\text{harmônica}}$  denota o fator de ponderação de harmônica S180 e  $W_{\text{ruído}}$  denota o fator de ponderação de ruído

S190. Alternativamente, o calculador de fator de ponderação 550 pode ser configurado para selecionar, de acordo com um valor de uma medição de periodicidade para o quadro ou subquadro atual, um fator correspondente entre uma pluralidade de pares de fatores de ponderação S180, S190, onde os pares são pré-calculados para satisfazer uma razão de energia constante tal como a expressão (2). Para uma implementação do calculador de fator de ponderação 550 no qual a expressão (2) é observada, valores típicos para fator de ponderação de harmônica S180 variam de aproximadamente 0,7 a aproximadamente 1,0, e valores típicos para fator de ponderação de ruído S190 variam de aproximadamente 0,1 a aproximadamente 0,7. Outras implementações do calculador de fator de ponderação 550 podem ser configuradas para operar de acordo com uma versão de expressão (2) que é modificada de acordo com uma ponderação de linha de base desejada entre o sinal estendido harmonicamente S160 e o sinal de ruído modulado S170.

[0143] Artefatos podem ocorrer em um sinal de fala sintetizado quando um livro-código disperso (em cujas entradas sejam na maioria valores zero) é utilizado para calcular a representação quantizada do residual. Dispersão de livro-código ocorre especialmente quando o sinal de banda estreita é codificado em uma taxa baixa de bits. Artefatos causados por dispersão de livro-código são tipicamente quase periódicos em tempo e ocorrem na maior parte acima de 3 kHz. Como o ouvido humano tem melhor resolução de tempo em frequências mais elevadas, esses artefatos podem ser mais perceptíveis na banda alta.

[0144] As modalidades incluem implementações do gerador de excitação de banda alta A300 que são configuradas para executar filtragem antidispersão. A FIGURA 18 mostra um diagrama em blocos de uma implementação A312 do gerador de

excitação de banda alta A302 que inclui um filtro antidispersão 600 disposto para filtrar o sinal de excitação de banda estreita desquantizado produzido pelo quantizador inverso 450. A FIGURA 19 mostra um diagrama em blocos de uma implementação A314 do gerador de excitação de banda alta A302 que inclui um filtro antidispersão 600 disposto para filtrar o sinal estendido espectralmente produzido pelo extensor de espectro A400. A FIGURA 20 mostra um diagrama em blocos de uma implementação A316 do gerador de excitação de banda alta A302 que inclui um filtro antidispersão 600 disposto para filtrar a saída do combinador 490 para produzir um sinal de excitação de banda alta S120. Evidentemente, implementações do gerador de excitação de banda alta A300 que combinam as características de quaisquer das implementações A304 e A306 com as características de quaisquer das implementações A312, A314, e A316 são consideradas e aqui expressamente reveladas. O filtro antidispersão 600 também pode ser disposto dentro do extensor de espectro A400: por exemplo, após qualquer um dos elementos 510, 520, 530 e 540 no extensor de espectro A402. É expressamente observado que o filtro antidispersão 600 também pode ser utilizado com implementações de extensor de espectro A400 que executam dobramento espectral, translação espectral ou extensão harmônica.

[0145] O filtro antidispersão 600 pode ser configurado para alterar a fase de seu sinal de entrada. Por exemplo, pode ser desejável que o filtro antidispersão 600 seja configurado e disposto de tal modo que a fase do sinal de excitação de banda alta S120 seja randomizada ou, de outro modo, mais uniformemente distribuída, com o passar do tempo. Também pode ser desejável que a resposta do filtro antidispersão 600 seja espectralmente plana, de tal modo que o espectro de magnitude do sinal filtrado não seja

apreciavelmente alterado. Em um exemplo, o filtro antidispersão 600 é implementado como um filtro de passagem total tendo uma função de transferência de acordo com a seguinte expressão:

$$H(z) = \frac{-0,7 + z^{-4}}{1 - 0,7z^{-4}} \cdot \frac{0,6 + z^{-6}}{1 + 0,6z^{-6}} \quad (3)$$

[0146] Um efeito de tal filtro pode ser espalhar a energia do sinal de entrada de modo que não mais seja concentrada somente em algumas amostras.

[0147] Artefatos causados por dispersão de livro-código são normalmente mais perceptíveis para sinais semelhantes a ruído, onde o residual inclui menos informação de pitch, e também para fala em ruído de fundo. Dispersão tipicamente causa um número menor de artefatos nos casos onde a excitação tem estrutura de longo prazo, e a modificação de fase pode realmente causar ruído em sinais vozeados. Desse modo, pode ser desejável configurar o filtro antidispersão 600 para filtrar sinais não-vozeados e passar pelo menos alguns sinais com voz sem alteração. Sinais não-vozeados são caracterizados por um ganho de pitch baixo (por exemplo, ganho de livro-código adaptável de banda estreita quantificado) e uma inclinação espectral (por exemplo, primeiro coeficiente de reflexão quantificado) que é próximo de zero ou positivo, indicando um envelope espectral que é plano ou inclinado para cima com frequência crescente. Implementações típicas de filtro antidispersão 600 são configuradas para filtrar sons não-vozeados (por exemplo, como indicado pelo valor da inclinação espectral), para filtrar sons vozeados quando o ganho de pitch está abaixo de um valor limite (alternativamente, não maior do que o valor limite) e, de outro modo, passar o sinal sem alteração.

[0148] Implementações adicionais de filtro antidispersão 600 incluem dois ou mais filtros que são configurados para ter ângulos de modificação de fase máxima diferentes (por exemplo, até 180 graus). Em tal caso, o filtro antidispersão 600 pode ser configurado para selecionar entre esses filtros de componentes de acordo com um valor do ganho de pitch (por exemplo, o ganho LTP ou livro-código adaptável quantificado) de tal modo que um ângulo de modificação de fase máxima maior seja utilizado para quadros tendo valores de ganho pitch mais baixos. Uma implementação de filtro antidispersão 600 pode incluir também filtros de componentes diferentes que são configurados para modificar a fase sobre mais ou menos do espectro de frequência, de tal modo que um filtro configurado para modificar a fase sobre uma faixa de frequência mais larga do sinal de entrada seja utilizado para quadros tendo valores de ganho de pitch mais baixos.

[0149] Para reprodução precisa do sinal de fala codificado, pode ser desejável que a razão entre os níveis das partes de banda alta e banda estreita do sinal de fala de banda larga sintetizado S100 seja similar àquela no sinal de fala de banda larga original S10. Além de um envelope espectral como representado por parâmetros de codificação de banda alta S60a, o encodificador de banda alta A200 pode ser configurado para caracterizar o sinal de banda alta S30 especificando um envelope de ganho ou temporal. Como mostrado na FIGURA 10, o encodificador de banda alta A202 inclui um calculador de fator de ganho de banda alta A230 que é configurado e disposto para calcular um ou mais fatores de ganho de acordo com uma relação entre o sinal de banda alta S30 e o sinal de banda alta sintetizado S130, como uma diferença ou razão entre as energias dos dois sinais sobre um quadro ou alguma parte do mesmo. Em outras implementações

do encodificador de banda alta A202, o calculador de ganho de banda alta A230 pode ser configurado de modo semelhante, porém disposto em vez de calcular o envelope de ganho de acordo com tal relação de variação de tempo entre o sinal de banda alta S30 e o sinal de excitação de banda estreita S80 ou sinal de excitação de banda alta S120.

[0150] Os envelopes temporais do sinal de excitação de banda estreita S80 e o sinal de banda alta S30 provavelmente são similares. Portanto, a encodificação de um envelope de ganho que se baseia em uma relação entre o sinal de banda alta S30 e o sinal de excitação de banda estreita S80 (ou um sinal derivado a partir do mesmo, como sinal de excitação de banda alta S120 ou sinal de banda alta sintetizado S130) será geralmente mais eficiente do que a codificação de um envelope de ganho com base somente no sinal de banda alta S30. Em uma implementação típica, o encodificador de banda alta A202 é configurado para transmitir um índice quantizado de oito a doze bits que especifica cinco fatores de ganho para cada quadro.

[0151] O calculador de fator de ganho de banda alta A230 pode ser configurado para executar o cálculo de fator de ganho como uma tarefa que inclui uma ou mais séries de subtarefas. A FIGURA 21 mostra um fluxograma de um exemplo T200 de tal tarefa que calcula um valor de ganho para um subquadro correspondente de acordo com as energias relativas do sinal de banda alta S30 e o sinal de banda alta sintetizado S130. As tarefas 220a e 220b calculam as energias dos subquadros correspondentes dos sinais respectivos. Por exemplo, tarefas 220a e 220b podem ser configuradas para calcular a energia como uma soma dos quadrados das amostras do subquadro respectivo. A tarefa T230 calcula um fator de ganho para o subquadro como a raiz quadrada da razão destas energias. Nesse exemplo, a tarefa T230 calcula o fator de

ganho como a raiz quadrada da razão da energia de sinal de banda alta S30 para a energia de sinal de banda alta sintetizado S130 sobre o subquadro.

[0152] Pode ser desejável que o calculador de fator de ganho de banda alta A230 seja configurado para calcular as energias de subquadro de acordo com uma função de janelamento. A FIGURA 22 mostra um fluxograma de tal implementação T210 da tarefa de cálculo de fator de ganho T200. A tarefa T215a aplica uma função de janelamento no sinal de banda alta S30, e a tarefa T215b aplica a mesma função de janelamento no sinal de banda alta sintetizado S130. Implementações 222a e 222b das tarefas 220a e 220b calculam as energias das janelas respectivas e a tarefa T230 calcula um fator de ganho para o subquadro como a raiz quadrada da razão das energias.

[0153] Pode ser desejável aplicar uma função de janelamento que sobreponha subquadros adjacentes. Por exemplo, uma função de janelamento que produz fatores de ganho, que podem ser aplicados em um modo de sobrepor-adicionar, pode ajudar a reduzir ou evitar descontinuidade entre subquadros. Em um exemplo, o calculador de fator de ganho de banda alta A230 é configurado para aplicar uma função de janelamento trapezoidal como mostrado na FIGURA 23a, na qual a janela sobrepõe cada um dos dois subquadros adjacentes por um milissegundo. A FIGURA 23b mostra uma aplicação dessa função de janelamento em cada um dos cinco subquadros de um quadro de 20 milissegundos. Outras implementações do calculador de fator de ganho de banda alta A230 podem ser configuradas para aplicar funções de janelamento tendo diferentes períodos de sobreposição e/ou diferentes formatos de janela (por exemplo, retangular, Hamming) que podem ser simétricos ou assimétricos. Também é possível para uma implementação de calculador de fator de

ganho de banda alta A230 ser configurado para aplicar diferentes funções de janelamento em diferentes subquadros em um quadro e/ou para um quadro incluir subquadros de comprimentos diferentes.

[0154] Sem limitação, os seguintes valores são apresentados como exemplos para implementações específicas. Um quadro de 20 ms é assumido para esses casos, embora qualquer outra duração possa ser utilizada. Para um sinal de banda alta amostrado em 7 kHz, cada quadro tem 140 amostras. Se tal quadro for dividido em cinco subquadros de comprimento igual, cada subquadro terá 28 amostras, e a janela como mostrado na FIGURA 23a terá 42 amostras de largura. Para um sinal de banda alta amostrado em 8 kHz, cada quadro tem 160 amostras. Se tal quadro for dividido em cinco subquadros de comprimento igual, cada subquadro terá 32 amostras e a janela como mostrado na FIGURA 23a terá 48 amostras de largura. Em outras implementações, os subquadros de qualquer largura podem ser utilizados, e é até mesmo possível para uma implementação de calculador de ganho de banda alta A230 ser configurado para produzir um fator de ganho diferente para cada amostra de um quadro.

[0155] A FIGURA 24 mostra um diagrama em blocos de uma implementação B202 do decodificador de banda alta B200. O decodificador de banda alta B202 inclui um gerador de excitação de banda alta B300 que é configurado para produzir o sinal de excitação de banda alta S120 com base no sinal de excitação de banda estreita S80. Dependendo das escolhas de projeto de sistema específicas, o gerador de excitação de banda alta B300 pode ser implementado de acordo com qualquer uma das implementações do gerador de excitação de banda alta A300 como descrito aqui. Tipicamente, é desejável implementar o gerador de excitação de banda alta B300 para ter a mesma resposta que o gerador de excitação de

banda alta do codificador de banda alta do sistema de codificação específico. Como o decodificador de banda estreita B110 executará tipicamente a desquantização do sinal de excitação de banda estreita encodificado S50, entretanto, na maioria dos casos o gerador de excitação de banda alta B300 pode ser implementado para receber o sinal de excitação de banda estreita S80 a partir do decodificador de banda estreita B110 e não necessita incluir um quantizador inverso configurado para desquantizar um sinal de excitação de banda estreita encodificado S50. É também possível que o decodificador de banda estreita B110 seja implementado para incluir uma ocorrência de filtro antidispersão 600 disposto para filtrar o sinal de excitação de banda estreita desquantizado antes de ser inserido em um filtro de síntese de banda estreita como filtro 330.

[0156] O quantizador inverso 560 é configurado para desquantizar parâmetros de filtro de banda alta S60a (nesse exemplo, para um conjunto de LSFs) e transformação de coeficiente de LSF para filtro LP 570 é configurado para transformar os LSFs em um conjunto de coeficientes de filtro (por exemplo, como descrito acima com referência ao quantificador inverso 240 e transformação 250 do codificador de banda estreita A122). Em outras implementações, como mencionado acima, conjuntos de coeficientes diferentes (por exemplo, coeficientes cepstrais) e/ou representações de coeficientes (por exemplo, ISPs) podem ser utilizados. O filtro de síntese de banda alta B200 é configurado para produzir um sinal de banda alta sintetizado de acordo com o sinal de excitação de banda alta S120 e o conjunto de coeficientes de filtro. Para um sistema no qual o codificador de banda alta inclui um filtro de síntese (por exemplo, como no exemplo do codificador A202 descrito acima), pode ser desejável implementar o filtro de

síntese de banda alta B200 para ter a mesma resposta (por exemplo, a mesma função de transferência) que a do filtro de síntese.

[0157] O decodificador de banda alta B202 também inclui um quantizador inverso 580 configurado para desquantizar fatores de ganho de banda alta S60b, e um elemento de controle de ganho 590 (por exemplo, um multiplicador ou amplificador) configurado e disposto para aplicar os fatores de ganho desquantizados ao sinal de banda alta sintetizado para produzir um sinal de banda alta S100. Para um caso no qual o envelope de ganho de um quadro é especificado por mais de um fator de ganho, o elemento de controle de ganho 590 pode incluir lógica configurada para aplicar os fatores de ganho aos respectivos subquadros, possivelmente de acordo com uma função de janelamento que pode ser a mesma ou uma função de janelamento diferente como aplicado por um calculador de ganho (por exemplo, calculador de ganho de banda alta A230) do encodificador de banda alta correspondente. Em outras implementações do decodificador de banda alta B202, o elemento de controle de ganho 590 é similarmente configurado, porém é disposto, em vez disso, para aplicar os fatores de ganho desquantizados no sinal de excitação de banda estreita S80 ou no sinal de excitação de banda alta S120.

[0158] Como mencionado acima, pode ser desejável obter o mesmo estado no encodificador de banda alta e decodificador de banda alta (por exemplo, utilizando valores desquantificados durante encodificação). Desse modo, pode ser desejável em um sistema de codificação de acordo com tal implementação assegurar o mesmo estado para geradores de ruído correspondentes em geradores de excitação de banda alta A300 e B300. Por exemplo, os geradores de excitação de banda alta A300 e B300 de tal implementação podem ser

configurados de tal modo que o estado do gerador de ruído seja uma função determinística de informações já codificadas dentro do mesmo quadro (por exemplo, parâmetros de filtro de banda estreita S40 ou uma parte dos mesmos e/ou sinal de excitação de banda estreita encodificado S50 ou uma parte do mesmo).

[0159] Um ou mais dos quantificadores dos elementos descritos aqui (por exemplo, quantificador 230, 420 ou 430) podem ser configurados para executar quantização de vetor classificada. Por exemplo, tal quantizador pode ser configurado para selecionar um dentre um conjunto de livros-código com base em informações que já foram codificadas no mesmo quadro no canal de banda estreita e/ou no canal de banda alta. Tal técnica provê tipicamente eficiência aumentada de codificação com o auxílio da armazenagem adicional de livro-código.

[0160] Como discutido acima com referência, por exemplo, às FIGURAS 8 e 9, uma quantidade considerável de estrutura periódica pode permanecer no sinal residual após remoção do envelope espectral grosseiro a partir do sinal de fala de banda estreita S20. Por exemplo, o sinal residual pode conter uma seqüência de pulsos ou pontas aproximadamente periódicos através do tempo. Tal estrutura, que é tipicamente relacionada a pitch, provavelmente ocorre especialmente em sinais de fala vozeados. O cálculo de uma representação quantizada do sinal residual de banda estreita pode incluir a codificação dessa estrutura de pitch de acordo com um modelo de periodicidade de longo prazo como representado, por exemplo, por um ou mais livros-código.

[0161] A estrutura de pitch de um sinal residual efetivo pode não associar exatamente com o modelo de periodicidade. Por exemplo, o sinal residual pode incluir pequenas flutuações na regularidade dos locais dos pulsos de

pitch, de tal modo que as distâncias entre pulsos sucessivos de pitch em um quadro não são exatamente iguais e a estrutura não é bem regular. Essas irregularidades tendem a reduzir a eficiência de codificação.

[0162] Algumas implementações de encodificador de banda estreita A120 são configuradas para executar uma regularização da estrutura de pitch pela aplicação de uma distorção de tempo adaptável para o residual antes ou durante quantização, ou por incluir, de outro modo, uma distorção de tempo adaptável no sinal de excitação codificado. Por exemplo, tal encodificador pode ser configurado para selecionar ou, de outro modo, calcular um grau de distorção no tempo (por exemplo, de acordo com um ou mais critérios de ponderação perceptível e/ou minimização de erro) de tal modo que o sinal de excitação resultante adapte de forma ótima ao modelo de periodicidade de longo prazo. A regularização de estrutura de pitch é executada por um subconjunto de encodificadores CELP denominados encodificadores de Predição Linear Excitada por Código de Relaxamento (RCELP).

[0163] Um encodificador RCELP é tipicamente configurado para executar a distorção de tempo como um deslocamento de tempo adaptável. Esse deslocamento de tempo pode ser um retardo que varia de alguns milisegundos negativos para alguns milisegundos positivos e, normalmente, varia suavemente para evitar descontinuidades audíveis. Em algumas implementações, tal encodificador é configurado para aplicar a regularização em um modelo de segmento, onde cada quadro ou subquadro é distorcido por um deslocamento de tempo fixo correspondente. Em outras implementações, o encodificador é configurado para aplicar a regularização como uma função de distorção contínua, de tal modo que um quadro ou subquadro é distorcido de acordo com um contorno de pitch (também denominada uma trajetória de pitch). Em

alguns casos (por exemplo, como descrito na publicação do pedido de patente US 2004/0098255), o codificador é configurado para incluir uma distorção de tempo no sinal de excitação encodificado pela aplicação do deslocamento em um sinal de entrada ponderado de forma perceptível que é utilizado para calcular o sinal de excitação codificado.

[0164] O codificador calcula um sinal de excitação encodificado que é regularizado e quantizado, e o decodificador desquantiza o sinal de excitação encodificado para obter um sinal de excitação que é utilizado para sintetizar o sinal de fala decodificado. O sinal de saída decodificado exibe, desse modo, o mesmo retardo de variação que foi incluído no sinal de excitação encodificado pela regularização. Tipicamente, nenhuma informação especificando as quantidades de regularização é transmitida para o decodificador.

[0165] A regularização tende a tornar o sinal residual mais fácil de encodificar, o que melhora o ganho de codificação a partir do preditor de longo prazo e, desse modo, intensifica a eficiência geral de codificação genericamente sem gerar artefatos. Pode ser desejável executar regularização apenas em quadros que são vozeados. Por exemplo, o codificador de banda estreita A124 pode ser configurado para deslocar somente aqueles quadros ou subquadros tendo uma estrutura de longo prazo, como sinais vozeados. Pode ser até mesmo desejável executar regularização somente em subquadros que incluem energia de pulso de pitch. Várias implementações de codificação RCELP são descritas nas patentes US nos. 5.704.003 (Kleijn e outros) e 6.879.955 (Rao) e na publicação do pedido de patente US 2004/0098255 (Kovesi e outros). Implementações existentes de codificadores RCELP incluem o Codec de Taxa Variável Intensificada (EVRC), como descrito na

Telecommunications Industry Association (TIA) IS-127, e Vocoder de Modo Seleccionável (SMV) do Projeto de Parceria da terceira Geração - 2 (3GPP2).

[0166] Infelizmente, a regularização pode causar problemas para um codificador de fala de banda larga no qual a excitação de banda alta é derivada do sinal de excitação de banda estreita encodificado (como um sistema que inclui encodificador de fala de banda larga A100 e decodificador de fala de banda larga B100). Devido à sua derivação a partir de um sinal distorcido de tempo, o sinal de excitação de banda alta terá geralmente um perfil de tempo que é diferente daquele do sinal de fala de banda alta original. Em outras palavras, o sinal de excitação de banda alta não mais será síncrono com o sinal de fala de banda alta original.

[0167] Um desalinhamento de tempo entre o sinal de excitação de banda alta distorcido e o sinal de fala de banda alta original pode causar vários problemas. Por exemplo, o sinal de excitação de banda alta distorcido não mais pode fornecer uma excitação de fonte apropriada para um filtro de síntese que é configurado de acordo com os parâmetros de filtro extraídos do sinal de fala de banda alta original. Como resultado, o sinal de banda alta sintetizado pode conter artefatos audíveis que reduzem a qualidade percebida do sinal de fala de banda larga decodificado.

[0168] O desalinhamento de tempo também pode causar ineficiências na codificação de envelope de ganho. Como mencionado acima, uma correlação existe provavelmente entre os envelopes temporais do sinal de excitação de banda estreita S80 e o sinal de banda alta S30. Pela encodificação do envelope de ganho do sinal de banda alta de acordo com uma relação entre esses dois envelopes temporais, um aumento

na eficiência de codificação pode ser realizado em comparação com a encodificação do envelope de ganho diretamente. Quando o sinal de excitação de banda estreita encodificado é regularizado, entretanto, essa correlação pode ser enfraquecida. O desalinhamento de tempo entre o sinal de excitação de banda estreita S80 e o sinal de banda alta S30 pode fazer com que flutuações apareçam em fatores de ganho de banda alta S60b, e a eficiência de codificação pode cair.

[0169] As modalidades incluem métodos de encodificação de fala de banda larga que executam distorção de tempo de um sinal de fala de banda alta de acordo com uma distorção de tempo incluído em um sinal de excitação de banda estreita encodificado correspondente. Vantagens em potencial de tais métodos incluem melhora da qualidade de um sinal de fala de banda larga decodificado e/ou melhora da eficiência de codificação de um envelope de ganho de banda alta.

[0170] A FIGURA 25 mostra um diagrama em blocos de uma implementação AD10 do encodificador de fala de banda larga A100. Encodificador AD10 inclui uma implementação A124 do encodificador de banda estreita A120 que é configurado para executar regularização durante o cálculo do sinal de excitação de banda estreita encodificado S50. Por exemplo, o encodificador de banda estreita A124 pode ser configurado de acordo com uma ou mais das implementações RCELP discutidas acima.

[0171] Encodificador de banda estreita A124 é também configurado para transmitir um sinal de dados de regularização SD10 que especifica o grau de distorção de tempo aplicado. Para vários casos nos quais o encodificador de banda estreita A124 é configurado para aplicar um deslocamento de tempo fixo em cada quadro ou subquadro, o sinal de dados de regularização SD10 pode incluir uma série de valores indicando cada quantidade de deslocamento de tempo

como um valor inteiro ou não inteiro em termos de amostras, milissegundos ou algum outro incremento de tempo. Para um caso no qual o encodificador de banda estreita A124 é configurado para, de outro modo, modificar a escala de tempo de um quadro ou outra seqüência de amostras (por exemplo, pela compressão de uma parte e expansão de outra parte), o sinal de informação de regularização SD10 pode incluir uma descrição correspondente da modificação, como um conjunto de parâmetros de função. Em um exemplo, específico, o encodificador de banda estreita A124 é configurado para dividir um quadro em três subquadros e calcular um deslocamento de tempo fixo para cada subquadro, de tal modo que o sinal de dados de regularização SD10 indique três quantidades de deslocamento de tempo para cada quadro regularizado do sinal de banda estreita encodificado.

[0172] O encodificador de fala de banda larga AD10 inclui uma linha de retardo D120 configurada para avançar ou retardar partes e sinal de fala de banda alta S30, de acordo com quantidades de retardo indicadas por um sinal de entrada, para produzir sinal de fala de banda alta com distorção de tempo S30a. No exemplo mostrado na FIGURA 25, a linha de retardo D120 é configurada para desviar no tempo o sinal de fala de banda alta S30 de acordo com a distorção indicada pelo sinal de dados de regularização SD10. Desse modo, a mesma quantidade de distorção de tempo que foi incluída no sinal de excitação de banda estreita encodificado S50 é também aplicada na parte correspondente do sinal de fala de banda alta S30 antes da análise. Embora esse exemplo mostre a linha de retardo D120 como um elemento separado a partir do encodificador de banda alta A200, em outras implementações a linha de retardo D120 é disposta como parte do encodificador de banda alta.

[0173] Implementações adicionais do codificador de banda alta A200 podem ser configuradas para executar análise espectral (por exemplo, análise LPC) do sinal de fala de banda alta não-distorcido S30 e executar distorção de tempo do sinal de fala de banda alta S30 antes do cálculo dos parâmetros de ganho de banda alta S60b. Tal codificador pode incluir, por exemplo, uma implementação de linha de retardo D120 disposta para executar a distorção de tempo. Em tais casos, entretanto, os parâmetros de filtro de banda alta S60a com base na análise de sinal não distorcido S30 podem descrever um envelope espectral que é desalinhado no tempo com o sinal de excitação de banda alta S120.

[0174] A linha de retardo D120 pode ser configurada de acordo com qualquer combinação de elementos de lógica e elementos de armazenamento apropriados para aplicar as operações desejadas de distorção de tempo no sinal de fala de banda alta S30. Por exemplo, a linha de retardo D120 pode ser configurada para ler o sinal de fala de banda alta S30 a partir de um buffer de acordo com os deslocamentos de tempo desejados. A FIGURA 26a mostra um diagrama esquemático de uma implementação D122 da linha de retardo D120 que inclui um registro de deslocamento SR1. O registro de deslocamento SR1 é um buffer de algum comprimento  $m$  que é configurado para receber e armazenar as amostras mais recentes  $m$  de sinal de fala de banda alta S30. O valor  $m$  é igual a pelo menos a soma dos deslocamentos de tempo positivo (ou "avançar") e negativo (ou "retardar") máximos a serem suportados. Pode ser conveniente para o valor  $m$  ser igual ao comprimento de um quadro ou subquadro de sinal de banda alta S30.

[0175] A linha de retardo D122 é configurada para transmitir o sinal de banda alta distorcido no tempo

S30a a partir de uma localização offset OL do registro de deslocamento SR1. A posição da localização offset OL varia em torno de uma posição de referência (deslocamento de tempo zero) de acordo com o deslocamento de tempo atual como indicado, por exemplo, pelo sinal de dados de regularização SD10. A linha de retardo D122 pode ser configurada para suportar limites de avanço e retardo iguais ou, alternativamente, um limite maior do que o outro de tal modo que um deslocamento maior possa ser executado em uma direção do que na outra. A FIGURA 26a mostra um exemplo específico que suporta um deslocamento maior de tempo positivo do que negativo. A linha de retardo D122 pode ser configurada para transmitir uma ou mais amostras em um tempo (dependendo de uma largura de barramento de saída, por exemplo).

[0176] Um deslocamento de tempo de regularização tendo uma magnitude de mais de alguns milisegundos pode causar artefatos audíveis no sinal decodificado. Tipicamente, a magnitude de um deslocamento de tempo de regularização como executado por um encodificador de banda estreita A124 não excederá alguns milisegundos, de tal modo que os deslocamentos de tempo indicados pelo sinal de dados de regularização SD10 serão limitados. Entretanto, pode ser desejável em tais casos que a linha de retardo D122 seja configurada para impor um limite máximo sobre os deslocamentos de tempo na direção positiva e/ou negativa (por exemplo, para observar um limite mais perto do que àquele imposto pelo encodificador de banda estreita).

[0177] A FIGURA 26b mostra um diagrama esquemático de uma implementação D124 da linha de retardo D122 que inclui uma janela de deslocamento SW. Nesse exemplo, a posição da localização offset OL é limitada pela janela de deslocamento SW. Embora a FIGURA 26b mostre um caso no qual o comprimento de buffer  $m$  é maior do que a largura da janela

de deslocamento SW, a linha de retardo D124 pode ser também implementada de tal modo que a largura da janela de deslocamento SW seja igual a  $m$ .

[0178] Em outras implementações, a linha de retardo D120 é configurada para gravar um sinal de fala de banda alta S30 para um buffer de acordo com os deslocamentos de tempo desejados. A FIGURA 27 mostra um diagrama esquemático de tal implementação D130 da linha de retardo D120 que inclui dois registros de deslocamento SR2 e SR3 configurados para receber e armazenar sinal de fala de banda alta S30. A linha de retardo D130 é configurada para gravar um quadro ou subquadro a partir do registro de deslocamento SR2 para o registro de deslocamento SR3 de acordo com um deslocamento de tempo como indicado, por exemplo, pelo sinal de dados de regularização SD10. O registro de deslocamento SR3 é configurado para um buffer FIFO disposto para transmitir um sinal de banda alta distorcido no tempo S30.

[0179] No exemplo específico mostrado na FIGURA 27, o registro de deslocamento SR2 inclui uma parte de buffer de quadro FB1 e uma parte de buffer de retardo DB, e o registro de deslocamento SR3 inclui uma parte de buffer de quadro FB2, uma parte de buffer de avanço AB e uma parte de buffer de retardo RB. Os comprimentos do buffer de avanço AB e buffer de retardo RB podem ser iguais, ou um pode ser maior do que o outro, de tal modo que um deslocamento maior em uma direção é mais suportado do que na outra. O buffer de retardo DB e a parte de buffer de retardo RB podem ser configurados para ter o mesmo comprimento. Alternativamente, o buffer de retardo DB pode ser mais curto do que o buffer de retardo RB para considerar um intervalo de tempo necessário para transferir amostras do buffer de quadro FB1 para o registro de deslocamento SR3, que pode incluir outras operações de

processamento como distorção de amostras antes de armazenamento para o registro de deslocamento SR3.

[0180] No exemplo da FIGURA 27, o buffer de quadro FB1 é configurado para ter um comprimento igual ao de um quadro de sinal de banda alta S30. Em outro exemplo, o buffer de quadro FB1 é configurado para ter um comprimento igual ao de um subquadro de sinal de banda alta S30. Em tal caso, a linha de retardo D130 pode ser configurada para incluir lógica para aplicar o mesmo retardo (por exemplo, uma média) em todos os subquadros de um quadro a ser deslocado. A linha de retardo D130 pode também incluir lógica para valores médios do buffer de quadro FB1 com valores a serem sobregravados em buffer de retardo RB ou buffer de avanço AB. Em um exemplo adicional, o registro de deslocamento SR3 pode ser configurado para receber valores de sinal de banda alta S30 somente através do buffer de quadro FB1, e em tal caso a linha de retardo D130 pode incluir lógica para interpolar através de folgas entre quadros ou subquadros sucessivos gravados para o registro de deslocamento SR3. Em outras implementações, a linha de retardo D130 pode ser configurada para executar uma operação de distorção nas amostras a partir do buffer de quadro FB1 antes de gravar os mesmos para o registro de deslocamento SR3 (por exemplo, de acordo com uma função descrita pelo sinal de dados de regularização SD10).

[0181] Pode ser desejável que a linha de retardo D120 aplique uma distorção de tempo que se baseia na, porém não é idêntica à distorção especificada por um sinal de dados de regularização SD10. A FIGURA 28 mostra um diagrama em blocos de uma implementação AD12 do codificador de fala de banda larga AD10 que inclui um mapeador de valor de retardo D110. O mapeador de valor de retardo D110 é configurado para mapear a distorção indicada por um sinal de dados de

regularização SD10 em valores de retardo mapeados SD10a. A linha de retardo D120 é disposta para produzir um sinal de fala de banda alta distorcido no tempo S30a de acordo com a distorção indicada por valores de retardo mapeados SD10a.

[0182] Pode-se esperar que o deslocamento de tempo aplicado pelo codificador de banda estreita se desenvolva suavemente com o passar do tempo. Portanto, é tipicamente suficiente computar o deslocamento de tempo de banda estreita média aplicado nos subquadros durante um quadro de fala, e deslocar um quadro correspondente de sinal de fala de banda alta S30 de acordo com essa média. Em tal exemplo, o mapeador de valor de retardo D110 é configurado para calcular uma média dos valores de retardo de subquadro para cada quadro, e a linha de retardo D120 é configurada para aplicar a média calculada a um quadro correspondente do sinal de banda alta S30. Em outros exemplos, uma média durante um período mais curto (como dois subquadros, ou metade de um quadro) ou um período mais longo (como dois quadros) pode ser calculada e aplicada. Em um caso onde a média é um valor não inteiro de amostras, o mapeador de valor de retardo D110 pode ser configurado para arredondar o valor para um número inteiro de amostras antes de transmitir o mesmo para a linha de retardo D120.

[0183] O codificador de banda estreita A124 pode ser configurado para incluir um deslocamento de tempo de regularização de um número não inteiro de amostras no sinal de excitação de banda estreita codificado. Em tal caso, pode ser desejável que o mapeador de valor de retardo D110 seja configurado para arredondar o deslocamento de tempo de banda estreita para um número inteiro de amostras e para que a linha de retardo D120 aplique o deslocamento de tempo arredondado para o sinal de fala de banda alta S30.

[0184] Em algumas implementações do encodificador de fala de banda larga AD10, as taxas de amostragem do sinal de fala de banda estreita S20 e o sinal de fala de banda alta S30 podem diferir. Em tais casos, o mapeador de valor de retardo D110 pode ser configurado para ajustar quantidades de deslocamento de tempo indicadas no sinal de dados de regularização SD10 para considerar uma diferença entre as taxas de amostragem de sinal de fala de banda estreita S20 (ou sinal de excitação de banda estreita S80) e sinal de fala de banda alta S30. Por exemplo, o mapeador de valor de retardo D110 pode ser configurado para escalar as quantidades de deslocamento de tempo de acordo com uma razão das taxas de amostragem. Em um exemplo específico como mencionado acima, o sinal de fala de banda estreita S20 é amostrado em 8 kHz, e o sinal de fala de banda alta S30 é amostrado em 7 kHz. Neste caso, o mapeador de valor de retardo D110 é configurado para multiplicar cada quantidade de deslocamento por 7/8. As implementações do mapeador de valor de retardo D110 também podem ser configuradas para executar uma operação de escala juntamente com uma operação de arredondamento de número inteiro e/ou uma mediação de deslocamento de tempo como descrito aqui.

[0185] Em implementações adicionais, a linha de retardo D120 é configurada para modificar, de outro modo, a escalar de tempo de um quadro ou outra seqüência de amostras (por exemplo, pela compressão de uma parte e expansão de outra parte). Por exemplo, o encodificador de banda estreita A124 pode ser configurado para executar a regularização de acordo com uma função como uma trajetória ou contorno de pitch. Em tal caso, o sinal de dados de regularização SD10 pode incluir uma descrição correspondente da função, como um conjunto de parâmetros e a linha de retardo D120 pode incluir lógica configurada para desviar quadros ou subquadros de

sinal de fala de banda alta S30 de acordo com a função. Em outras implementações, o mapeador de valor de retardo D110 é configurado para mediar, escalar, e/ou arredondar a função antes de ser aplicado no sinal de fala de banda alta S30 pela linha de retardo D120. Por exemplo, o mapeador de valor de retardo D110 pode ser configurado para calcular um ou mais valores de retardo de acordo com a função, cada valor de retardo indicando um número de amostras, as quais são então aplicadas pela linha de retardo D120 para desviar em tempo um ou mais quadros ou subquadros correspondentes do sinal de fala de banda alta S30.

[0186] A FIGURA 29 mostra um fluxograma para um método MD100 de distorção de tempo, um sinal de fala de banda alta de acordo com uma distorção de tempo incluída em um sinal de excitação de banda estreita encodificado correspondente. A tarefa TD100 processa um sinal de fala de banda larga para obter um sinal de fala de banda estreita e um sinal de fala de banda alta. Por exemplo, a tarefa TD100 pode ser configurada para filtrar o sinal de fala de banda larga utilizando um banco de filtros tendo filtros passa-baixa e passa-alta, como uma implementação de banco de filtros A110. A tarefa TD200 encodifica o sinal de fala de banda estreita em pelo menos um sinal de excitação de banda estreita encodificado e uma pluralidade de parâmetros de filtro de banda estreita. O sinal de excitação de banda estreita encodificado e/ou parâmetros de filtro podem ser quantizados, e o sinal de fala de banda estreita encodificado também pode incluir outros parâmetros como um parâmetro de modo de fala. A tarefa TD200 também inclui uma distorção de tempo no sinal de excitação de banda estreita encodificado.

[0187] A tarefa TD300 gera um sinal de excitação de banda alta com base em um sinal de excitação de banda estreita. Nesse caso, o sinal de excitação de banda estreita

se baseia no sinal de excitação de banda estreita encodificado. De acordo com pelo menos o sinal de excitação de banda alta, a tarefa TD400 encodifica o sinal de fala de banda alta em pelo menos uma pluralidade de parâmetros de filtro de banda alta. Por exemplo, a tarefa TD400 pode ser configurada para encodificar o sinal de fala de banda alta em uma pluralidade de LSFs quantizados. A tarefa TD500 aplica um deslocamento de tempo para o sinal de fala de banda alta que se baseia em informações referentes a uma distorção de tempo incluída no sinal de excitação de banda estreita encodificado.

[0188] A tarefa TD400 pode ser configurada para executar análise espectral (como uma análise LPC) no sinal de fala de banda alta, e/ou calcular um envelope de ganho do sinal de fala de banda alta. Em tais casos, a tarefa TD500 pode ser configurada para aplicar o deslocamento de tempo no sinal de fala de banda alta antes da análise e/ou do cálculo de envelope de ganho.

[0189] Outras implementações do encodificador de fala de banda larga A100 são configuradas para inverter uma distorção de tempo do sinal de excitação de banda alta S120 causado por uma distorção de tempo incluída no sinal de excitação de banda estreita encodificado. Por exemplo, o gerador de excitação de banda alta A300 pode ser implementado para incluir uma implementação da linha de retardo D120 que é configurada para receber sinal de dados de regularização SD10 ou valores de retardo mapeados SD10a, e aplicar um deslocamento de tempo reverso correspondente no sinal de excitação de banda estreita S80, e/ou em um sinal subsequente com base no mesmo como sinal harmonicamente estendido S160 ou sinal de excitação de banda alta S120.

[0190] Implementações de encodificador de fala de banda larga adicionais podem ser configuradas para

codificar o sinal de fala de banda estreita S20 e o sinal de fala de banda alta S30 independentemente de um para o outro, de tal modo que o sinal de fala de banda alta S30 seja codificado como uma representação de um envelope espectral de banda alta e um sinal de excitação de banda alta. Tal implementação pode ser configurada para executar distorção de tempo do sinal residual de banda alta, ou de outro modo incluir distorção de tempo em um sinal de excitação de banda alta encodificado, de acordo com informações referentes a uma distorção de tempo incluída no sinal de excitação de banda estreita encodificado. Por exemplo, o codificador de banda alta pode incluir uma implementação da linha de retardo D120 e/ou mapeador de valor de retardo D110 como descrito aqui que são configurados para aplicar uma distorção de tempo no sinal residual de banda alta. Vantagens em potencial de tal operação incluem codificação mais eficiente do sinal residual de banda alta e melhor associação entre os sinais de fala de banda alta e banda estreita sintetizados.

[0191] Como mencionado acima, as modalidades como descritas aqui incluem implementações que podem ser utilizadas para executar codificação incorporada, suportando compatibilidade com sistemas de banda estreita e evitando necessidade de transcodificação. O suporte para codificação de banda alta também pode servir para diferenciar uma base de custo entre chips, chipsets, dispositivos, e/ou redes tendo suporte de banda larga com compatibilidade retroativa, e aqueles tendo somente suporte de banda estreita. O suporte para codificação de banda alta como descrito aqui também pode ser utilizado em combinação com uma técnica para suportar codificação de banda baixa, e um sistema, método, ou equipamento de acordo com tal modalidade pode suportar codificação de componentes de frequência de, por exemplo, aproximadamente 50 ou 100 Hz até aproximadamente 7 ou 8 kHz.

[0192] Como mencionado acima, a adição de suporte de banda alta em um codificador de fala pode melhorar a inteligibilidade, especialmente em relação à diferenciação de fricativos. Embora tal diferenciação possa normalmente ser derivada por um ouvinte humano a partir do contexto específico, o suporte de banda alta pode servir como uma característica de habilitação no reconhecimento de fala e outras aplicações de interpretação de máquina, como sistemas para navegação de menu de voz automatizada e/ou processamento de chamada automática.

[0193] Um equipamento de acordo com uma modalidade pode ser incorporado em um dispositivo portátil para comunicações sem fio, tal como um telefone celular ou assistente digital pessoal (PDA). Alternativamente, tal equipamento pode ser incluído em outro dispositivo de comunicações, tal como um aparelho de telefone VoIP, um computador pessoal configurado para suportar comunicações VoIP, ou um dispositivo de rede configurado para rotear comunicações telefônicas ou VoIP. Por exemplo, um equipamento de acordo com uma modalidade pode ser implementado em um chip ou chipset para um dispositivo de comunicações. Dependendo da aplicação específica, tal dispositivo também pode incluir tais características como conversão de analógico em digital e/ou digital em analógico de um sinal de fala, conjunto de circuitos para executar amplificação e/ou outras operações de processamento de sinal em um sinal de fala, e/ou conjunto de circuitos de radiofrequência para transmissão e/ou recepção do sinal de fala codificado.

[0194] É explicitamente considerado e revelado que modalidades podem incluir e/ou ser utilizadas com qualquer uma ou mais das outras características reveladas nos pedidos de patentes provisionais US números 60/667,901

e 60/673,965, dos quais esse pedido reivindica benefício. Tais características incluem remoção de rajadas de energia alta de curto prazo que ocorrem na banda alta e estão substancialmente ausentes a partir da banda estreita. Tais características incluem suavização adaptativa ou fixa de representações de coeficiente como LSFs de banda alta. Tais características incluem formatação adaptável ou fixa de ruído associado à quantização de representações de coeficientes como LSFs. Tais características também incluem suavização adaptativa ou fixa de um envelope de ganho, e atenuação adaptativa de um envelope de ganho.

[0195] A apresentação acima das modalidades descritas é fornecida para permitir que qualquer pessoa versada na técnica faça ou utilize a presente invenção. Várias modificações nestas modalidades são possíveis, e os princípios gerais apresentados aqui podem ser aplicados também em outras modalidades. Por exemplo, uma modalidade pode ser implementada em parte ou integralmente como um circuito permanente (hardwired), como uma configuração de circuito fabricada em um circuito integrado de aplicação específica, ou como um programa de firmware carregado em armazenamento não-volátil ou um programa de software carregado de, ou para, um meio de armazenagem de dados como código legível por máquina, tal código tendo instruções executáveis por uma disposição de elementos de lógica como um microprocessador ou outra unidade de processamento de sinais digitais. O meio de armazenagem de dados pode ser uma disposição de elementos de armazenagem, tais como memória de semicondutor (que pode incluir sem limitação, RAM dinâmica ou estática (memória de acesso aleatório), ROM (memória somente de leitura), e/ou RAM flash), ou memória ferroelétrica, magnetorresistiva, ovônica, polimérica, ou de alteração de fase; ou um meio de disco como um disco

magnético ou óptico. O termo "software" deve ser entendido como incluindo código fonte, código de linguagem assembly, código de máquina, código binário, firmware, macrocódigo, microcódigo, qualquer um ou mais conjuntos ou seqüências de instruções executáveis por uma disposição de elementos de lógica, e qualquer combinação de tais exemplos.

[0196] Os vários elementos de implementações de geradores de excitação de banda alta A300 e B300, codificador de banda alta A100, decodificador de banda alta B200, codificador de fala de banda larga A100, e decodificador de fala de banda larga B100 podem ser implementados como dispositivos eletrônicos e/ou ópticos residindo, por exemplo, no mesmo chip ou entre dois ou mais chips em um chipset, embora outros arranjos sem essa limitação também sejam considerados. Um ou mais elementos de tal equipamento pode ser implementado integralmente ou em parte como um ou mais conjuntos de instruções dispostos para executar em uma ou mais disposições fixas ou programáveis de elementos de lógica (por exemplo, transistores, portas), tais como microprocessadores, processadores incorporados, núcleos IP, processadores de sinais digitais, FPGAs (matrizes de porta programáveis em campo), ASSPs (produtos padrão de aplicação específica) e ASICs (circuitos integrados de aplicação específica). É também possível que um ou mais de tais elementos tenham estrutura em comum (por exemplo, um processador utilizado para executar partes de código correspondendo a diferentes elementos em tempos diferentes, um conjunto de instruções executadas para realizar tarefas correspondendo a diferentes elementos em diferentes tempos, ou um arranjo de dispositivos eletrônicos e/ou ópticos executando operações para diferentes elementos em diferentes tempos). Além disso, é possível que um ou mais destes elementos sejam utilizados para executar tarefas ou

executar outros conjuntos de instruções que não sejam diretamente relacionadas a uma operação do equipamento, como uma tarefa referente à outra operação de um dispositivo ou sistema no qual o equipamento seja incorporado.

[0197] A FIGURA 30 mostra um fluxograma de um método M100, de acordo com uma modalidade, para encodificar uma parte de banda alta de um sinal de fala tendo uma parte de banda estreita e uma parte de banda alta. A tarefa X100 calcula um conjunto de parâmetros de filtro que caracteriza um envelope espectral da parte de banda alta. A tarefa X200 calcula um sinal estendido espectralmente pela aplicação de uma função não-linear a um sinal derivado a partir da parte de banda estreita. A tarefa X300 gera um sinal de banda alta sintetizado de acordo com (A) o conjunto de parâmetros de filtro e (B) um sinal de excitação de banda alta com base no sinal estendido espectralmente. A tarefa X400 calcula um envelope de ganho com base em uma relação entre (C) energia da parte de banda alta e (D) energia de um sinal derivado a partir da parte de banda estreita.

[0198] A FIGURA 31a mostra um fluxograma de um método M200 para gerar um sinal de excitação de banda alta de acordo com uma modalidade. A tarefa Y100 calcula um sinal harmonicamente estendido pela aplicação de uma função não-linear em um sinal de excitação de banda estreita derivado de uma parte de banda estreita de um sinal de fala. A tarefa Y200 mistura o sinal harmonicamente estendido com um sinal de ruído modulado para gerar um sinal de excitação de banda alta. A FIGURA 31b mostra um fluxograma de um método M210 para gerar um sinal de excitação de banda alta de acordo com outra modalidade incluindo tarefas Y300 e Y400. A tarefa Y300 calcula um envelope de domínio de tempo de acordo com a energia no tempo de um entre o sinal de excitação de banda estreita e o sinal harmonicamente estendido. A tarefa Y400

modula um sinal de ruído de acordo com o envelope de domínio de tempo para produzir o sinal de ruído modulado.

[0199] A FIGURA 32 mostra um fluxograma de um método M300 de acordo com uma modalidade, para decodificar uma parte de banda alta de um sinal de fala tendo uma parte de banda estreita e a parte de banda alta. A tarefa Z100 recebe um conjunto de parâmetros de filtro que caracteriza um envelope espectral da parte de banda alta e um conjunto de fatores de ganho que caracterizam um envelope temporal da parte de banda alta. A tarefa Z200 calcula um sinal espectralmente estendido pela aplicação de uma função não linear em um sinal derivado a partir da parte de banda estreita. A tarefa Z300 gera um sinal de banda alta sintetizado de acordo com (A) o conjunto de parâmetros de filtro e (B) um sinal de excitação de banda alta com base no sinal estendido espectralmente. A tarefa Z400 modula um envelope de ganho do sinal de banda alta sintetizado com base no conjunto de fatores de ganho. Por exemplo, a tarefa Z400 pode ser configurada para modular o envelope de ganho do sinal de banda alta sintetizado pela aplicação do conjunto de fatores de ganho em um sinal de excitação derivado da parte de banda estreita, ao sinal estendido espectralmente, ao sinal de excitação de banda alta ou ao sinal de banda alta sintetizado.

[0200] As modalidades também incluem métodos adicionais de codificação de fala, encodificação e decodificação, como são expressamente revelados aqui, por exemplo, por descrições de modalidades estruturais configuradas para executar tais métodos. Cada um destes métodos também pode ser incorporado de forma tangível (por exemplo, em um ou mais meios de armazenamento de dados como listado acima) como um ou mais conjuntos de instruções legíveis e/ou executáveis por máquina incluindo uma

disposição de elementos de lógica (por exemplo, um processador, microprocessador, microcontrolador, ou outra máquina de estado finito). Desse modo, a presente invenção não pretende ser limitada às modalidades mostradas acima, porém, em vez disso, deve ser acordado o escopo mais amplo compatível com os princípios e características novas reveladas em qualquer modo aqui, inclusive nas reivindicações em anexo como depositadas, que fazem parte da descrição original.

REIVINDICAÇÕES

## 1. Equipamento, compreendendo:

- um banco de filtros (A110) tendo

(A) um percurso de processamento de banda baixa configurado para receber um sinal de fala de banda larga e para produzir um sinal de fala de banda baixa com base em uma parte de baixa frequência do sinal de fala de banda larga, e

(B) um percurso de processamento de banda alta configurado para receber o sinal de fala de banda larga e para produzir um sinal de fala de banda alta com base em uma parte de alta frequência do sinal de fala de banda larga, em que uma banda passante do percurso de processamento de banda baixa sobrepõe uma banda passante do percurso de processamento de banda alta, em que a sobreposição é considerada como a distância do ponto no qual a resposta de frequência do filtro de banda alta salta para -20 dB a partir do ponto no qual a resposta de frequência do filtro de banda baixa salta para -20 dB;

- um primeiro encodificador de fala (A120) configurado para encodificar o sinal de fala de banda baixa em pelo menos um sinal de excitação de banda baixa encodificado e uma pluralidade de parâmetros de filtro de banda baixa; e

- um segundo encodificador de fala (A200) configurado para gerar um sinal de excitação de banda alta com base no sinal de excitação de banda baixa encodificado, e encodificar o sinal de banda alta, de acordo com o sinal de excitação de banda alta, em pelo menos uma pluralidade de parâmetros de filtro de banda alta,

o equipamento **caracterizado pelo** fato de que a banda passante do percurso de processamento de banda baixa

sobrepõe a banda passante do percurso de processamento de banda alta por 400 a 1000 Hz.

2. Equipamento, de acordo com a reivindicação 1, **caracterizado pelo** fato de que o segundo encodificador de fala (A200) é configurado para gerar o sinal de excitação de banda alta pela aplicação de uma função não-linear a um sinal que é baseado no sinal de excitação de banda baixa encodificado para gerar um sinal estendido espectralmente, e

em que o sinal de excitação de banda alta é baseado no sinal estendido espectralmente.

3. Equipamento, de acordo com a reivindicação 1, **caracterizado pelo** fato de que o segundo encodificador de fala (A200) é configurado para encodificar um envelope de ganho do sinal de banda alta.

4. Equipamento, de acordo com a reivindicação 3, **caracterizado pelo** fato de que o segundo encodificador de fala (A200) é configurado para gerar um sinal de banda alta sintetizado de acordo com o sinal de excitação de banda alta e a pluralidade de parâmetros de filtro de banda alta, e

em que o segundo encodificador de fala (A200) é configurado para encodificar o envelope de ganho com base no sinal de banda alta sintetizado.

5. Equipamento, de acordo com a reivindicação 4, **caracterizado pelo** fato de que o segundo encodificador (A200) é configurado para encodificar o envelope de ganho com base em uma relação entre o sinal de banda alta e o sinal de banda alta sintetizado.

6. Equipamento, de acordo com a reivindicação 1, **caracterizado pelo** fato de que a banda passante do percurso de processamento de banda baixa sobrepõe a banda passante do percurso de processamento de banda alta em 500 Hz.

7. Equipamento, de acordo com a reivindicação 1, **caracterizado pelo** fato de que a banda passante do percurso de processamento de banda baixa sobrepõe a banda passante do percurso de processamento de banda alta em 400 a 600 Hz.

8. Equipamento, de acordo com a reivindicação 1, **caracterizado pelo** fato de que a sobreposição inclui pelo menos uma parte da faixa de frequência de 2000 a 5000 Hz.

9. Equipamento, de acordo com a reivindicação 1, **caracterizado pelo** fato de que a sobreposição inclui pelo menos uma parte da faixa de frequência de 3000 a 4000 Hz.

10. Equipamento, de acordo com a reivindicação 1, **caracterizado pelo** fato de que o sinal de fala de banda baixa e o sinal de fala de banda alta têm diferentes taxas de amostragem.

11. Equipamento, de acordo com a reivindicação 1, **caracterizado pelo** fato de que uma soma das taxas de amostragem do sinal de fala de banda baixa e o sinal de fala de banda alta não é maior do que a taxa de amostragem do sinal de banda larga.

12. Equipamento, de acordo com a reivindicação 1, **caracterizado pelo** fato de que compreende um telefone celular.

13. Equipamento, de acordo com a reivindicação 1, **caracterizado pelo** fato de que compreende um dispositivo configurado para transmitir uma pluralidade de pacotes em conformidade com uma versão do Protocolo Internet, em que a pluralidade de pacotes descreve o sinal de excitação de banda baixa encodificado, a pluralidade de parâmetros de filtro de banda baixa, e a pluralidade de parâmetros de filtro de banda alta.

14. Método para processar sinais, o método compreendendo:

- produzir um sinal de fala de banda baixa com base em um sinal de fala de banda larga tendo conteúdo de frequência entre pelo menos 1000 e 6000 Hz;

- encodificar o sinal de fala de banda baixa;

- produzir um sinal de fala de banda alta com base no sinal de fala de banda larga; e

- encodificar o sinal de fala de banda alta;

em que produzir um sinal de fala de banda baixa inclui a produção do sinal de fala de banda baixa com base em

(A) uma primeira parte do conteúdo de frequência do sinal de banda larga, a primeira parte incluindo a parte do sinal de banda larga entre 1000 e 2000 Hz, e

(B) uma terceira parte do conteúdo de frequência do sinal de banda larga,

o método **caracterizado pelo** fato de que a terceira parte inclui uma parte do sinal de banda larga entre 2000 e 5000 Hz que tem uma largura de pelo menos 400 Hz a 1000 Hz em -20dB, e

em que produzir um sinal de fala de banda alta inclui a produção do sinal de fala de banda alta com base em

(C) uma segunda parte do conteúdo de frequência do sinal de banda larga, a segunda parte incluindo a parte do sinal de banda larga entre 5000 e 6000 Hz, e

(D) a terceira parte do conteúdo de frequência do sinal de banda larga.

15. Método, de acordo com a reivindicação 14, **caracterizado pelo** fato de que o sinal de fala de banda baixa inclui conteúdo de frequência da primeira parte e conteúdo de frequência da terceira parte, e

em que o sinal de fala de banda alta inclui conteúdo de frequência da segunda parte e conteúdo de frequência da terceira parte.

16. Método, de acordo com a reivindicação 14, **caracterizado pelo** fato de que o sinal de fala de banda baixa e o sinal de fala de banda alta têm diferentes taxas de amostragem.

17. Método, de acordo com a reivindicação 14, **caracterizado pelo** fato de que uma soma das taxas de amostragem do sinal de fala de banda baixa e o sinal de fala de banda alta não é maior do que a taxa de amostragem do sinal de banda larga.

18. Método, de acordo com a reivindicação 14, **caracterizado pelo** fato de que o primeiro encodificador de fala é configurado para encodificar o sinal de fala de banda baixa em pelo menos um sinal de excitação de banda baixa encodificado e uma pluralidade de parâmetros de filtro de banda baixa, e

em que o segundo encodificador de fala é configurado para gerar um sinal de excitação de banda alta com base no sinal de excitação de banda baixa encodificado, e encodificar o sinal de banda alta, de acordo com o sinal de excitação de banda alta, em pelo menos uma pluralidade de parâmetros de filtro de banda alta.

19. Método, de acordo com a reivindicação 14, **caracterizado pelo** fato de que o segundo encodificador de fala é configurado para encodificar o sinal de banda alta em pelo menos uma pluralidade de parâmetros de filtro de banda alta e uma pluralidade de fatores de ganho.

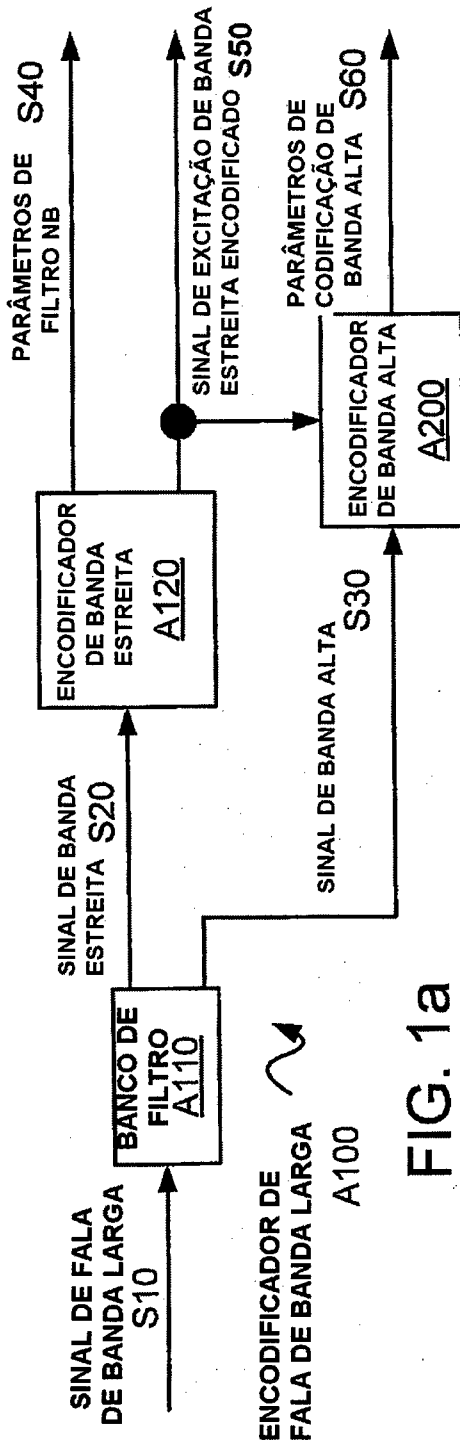


FIG. 1a

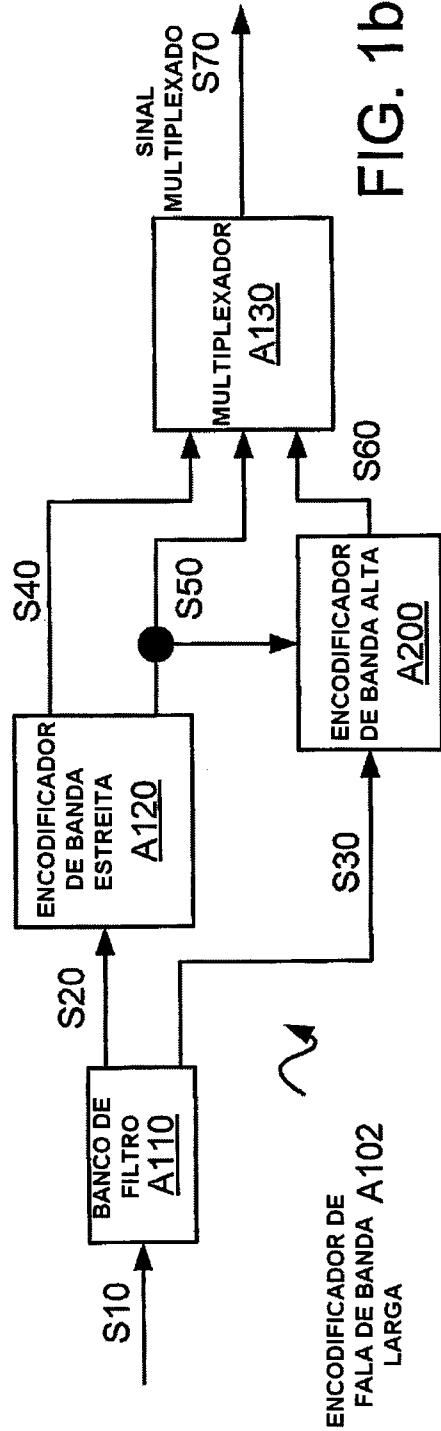


FIG. 1b

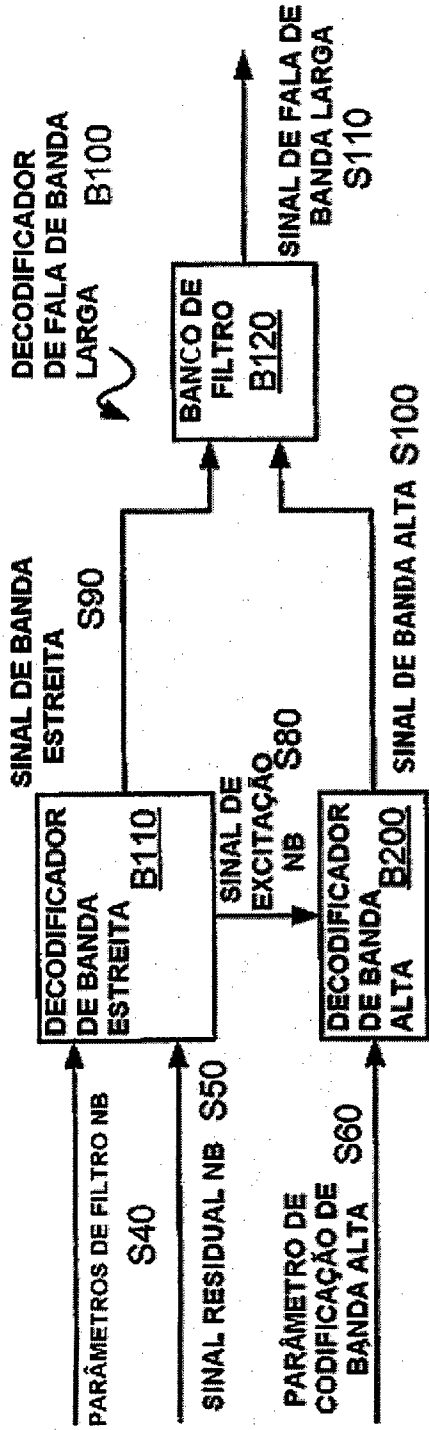


FIG. 2a

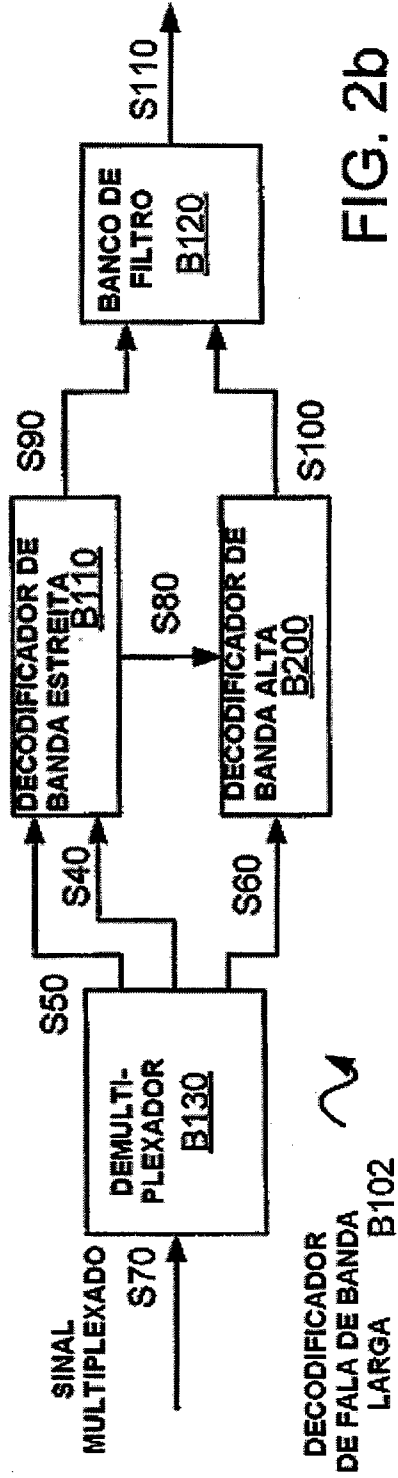


FIG. 2b

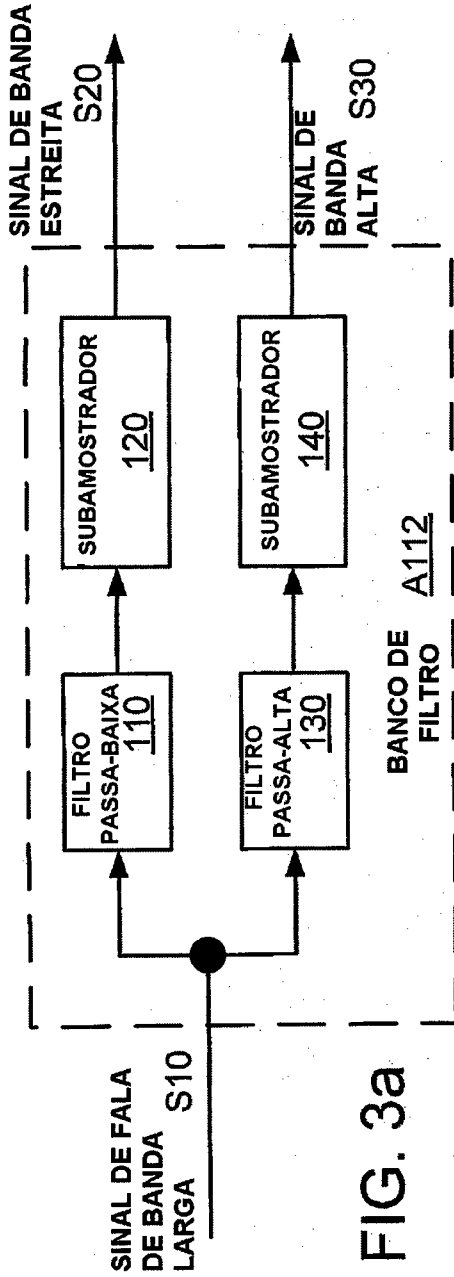


FIG. 3a

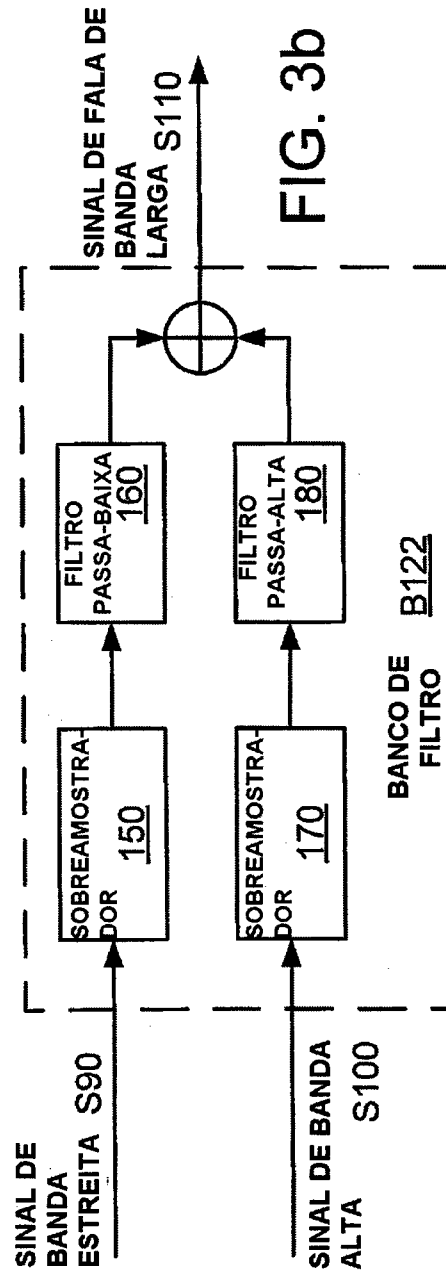


FIG. 3b

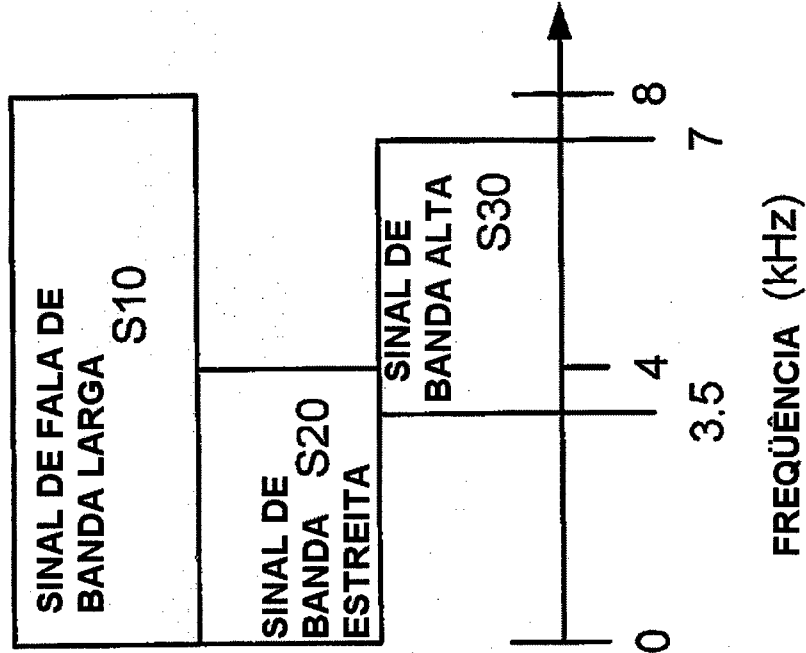


FIG. 4a

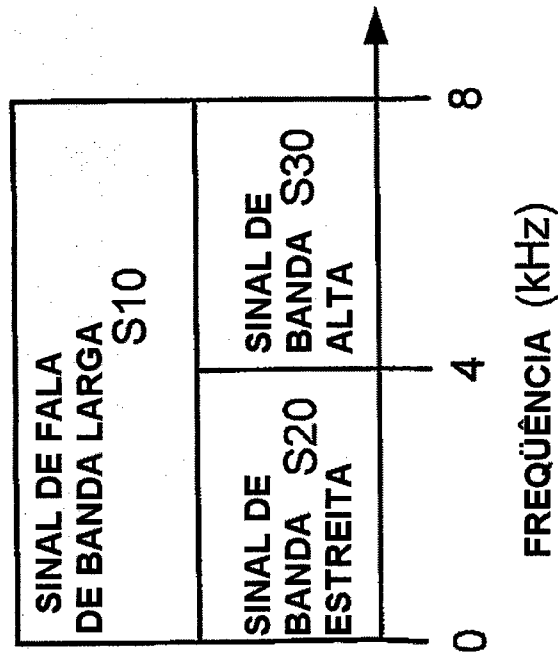


FIG. 4b

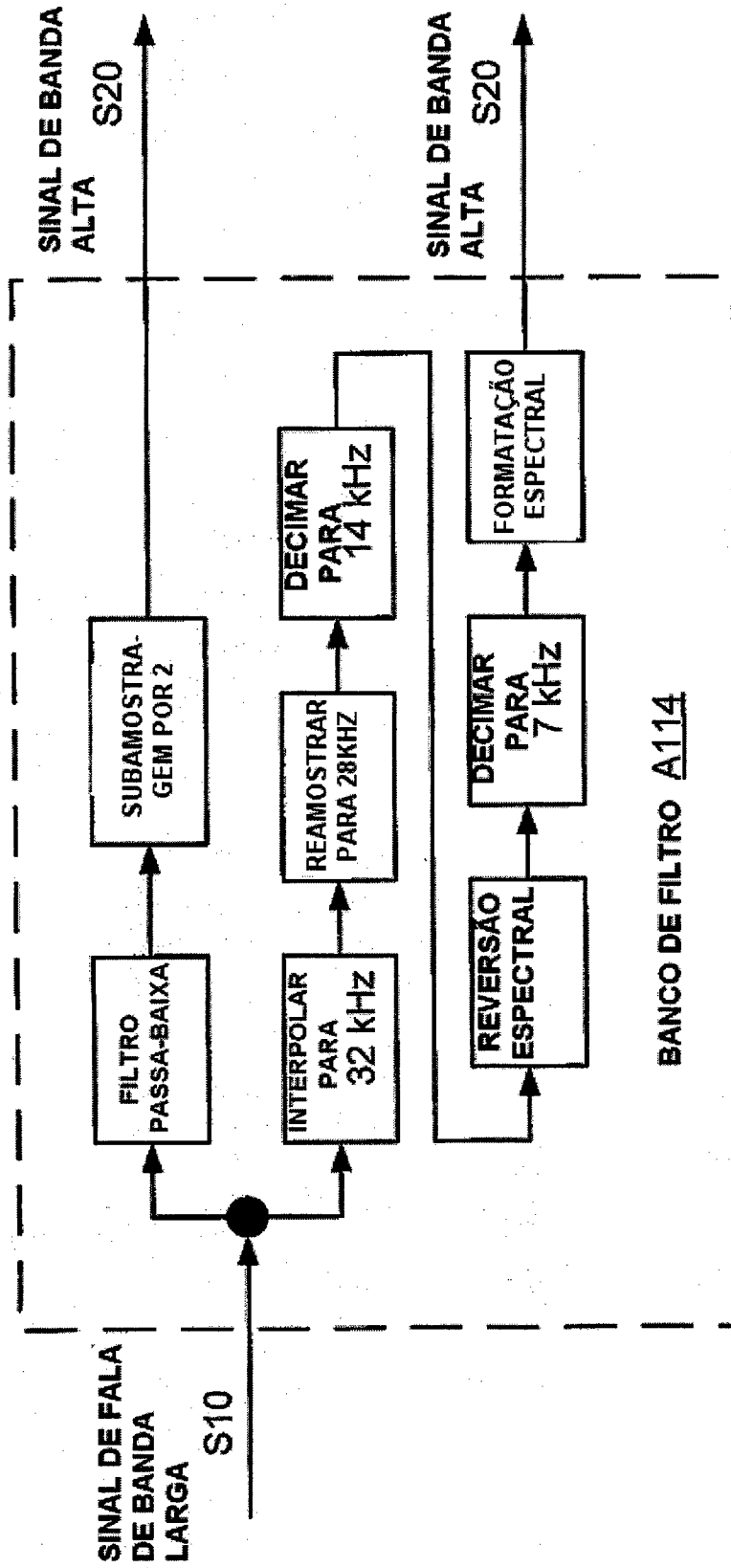


FIG. 4C

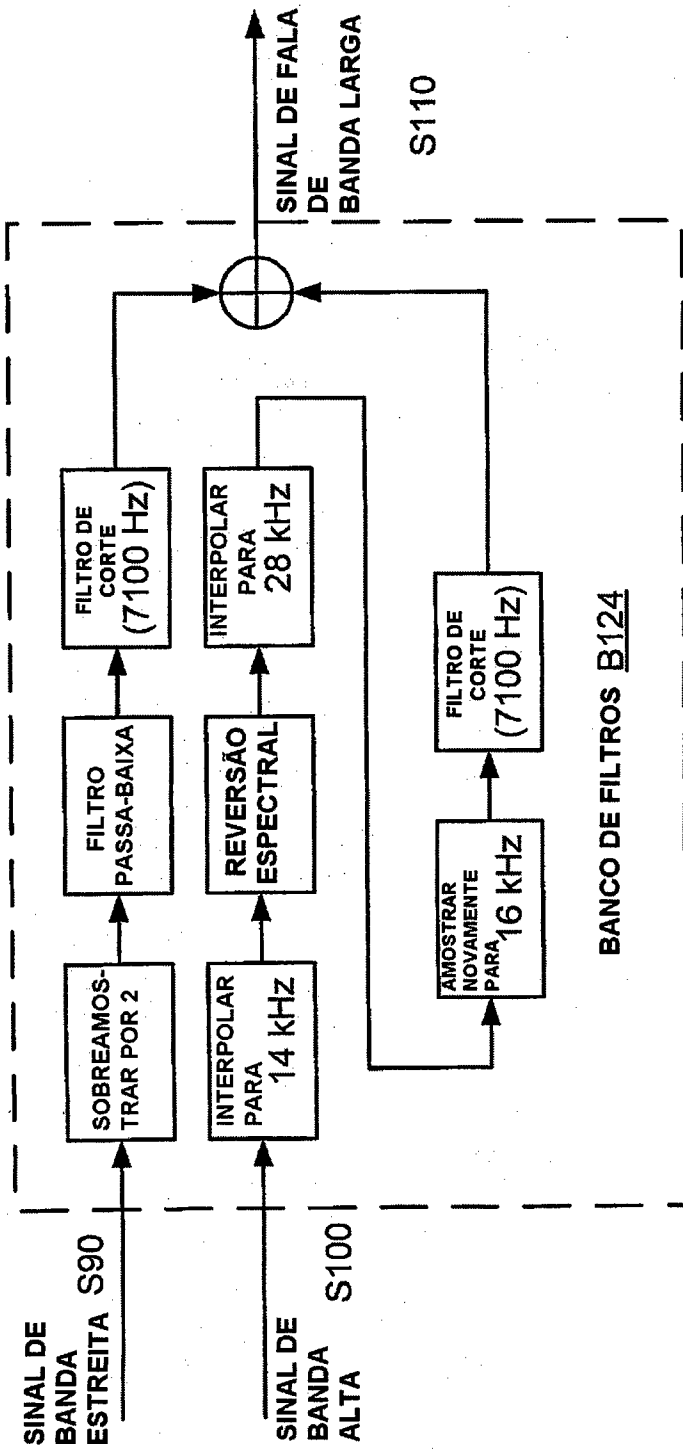
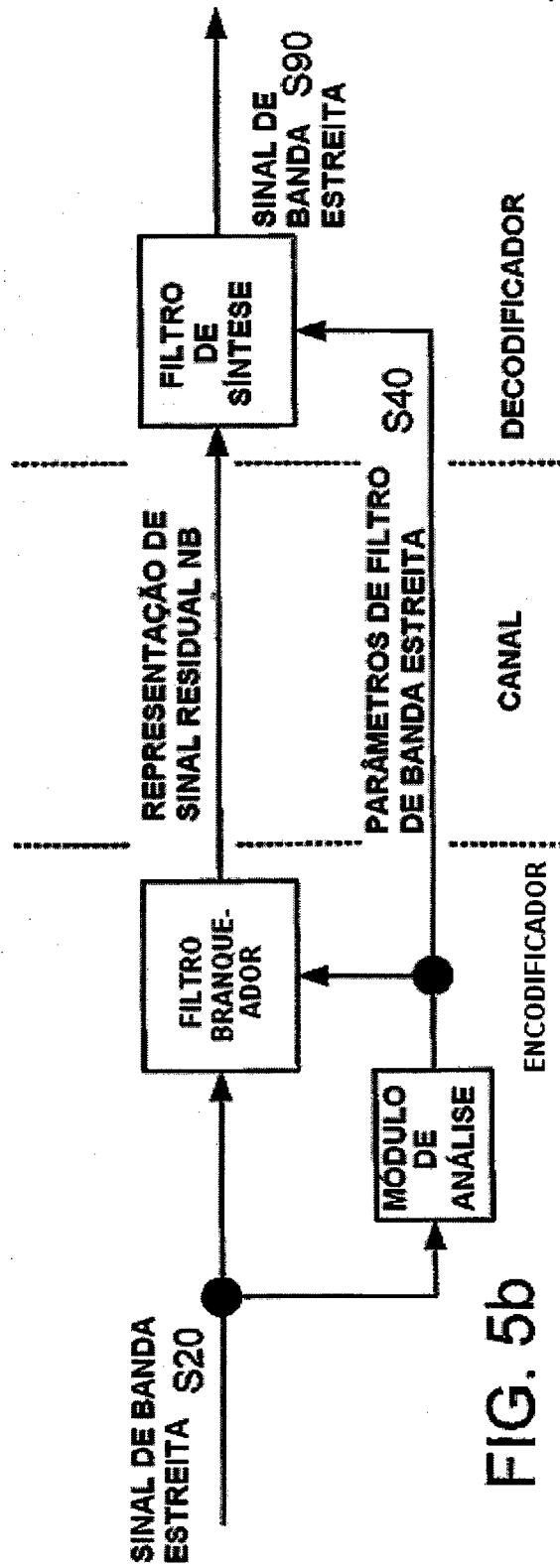
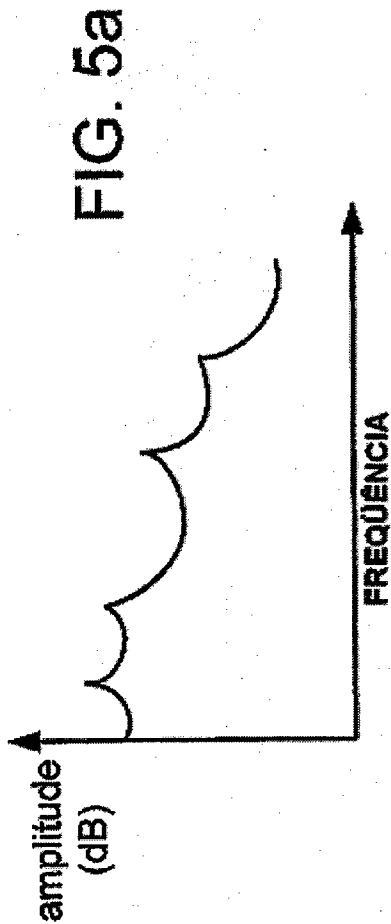


FIG. 4d



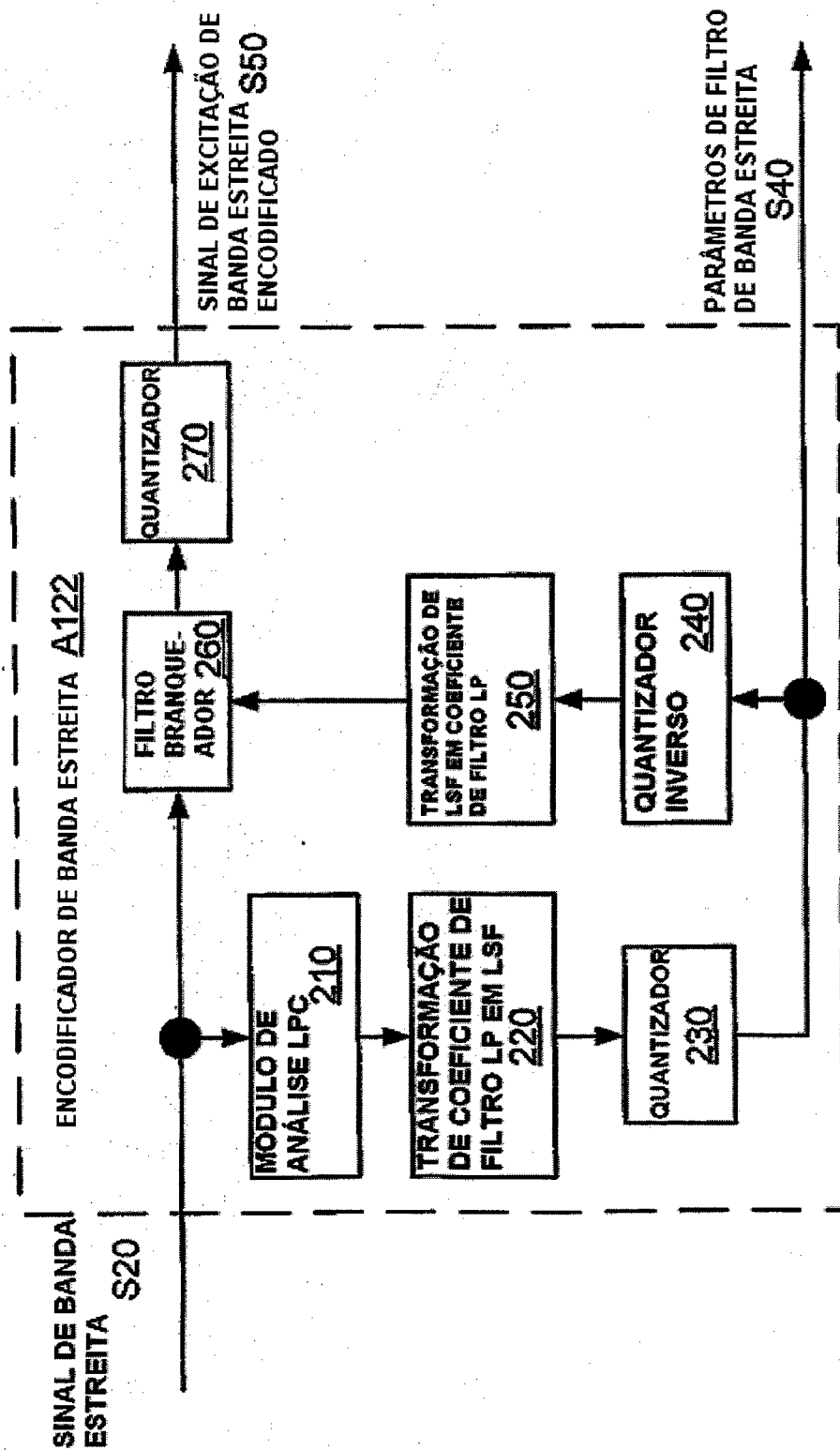


FIG. 6

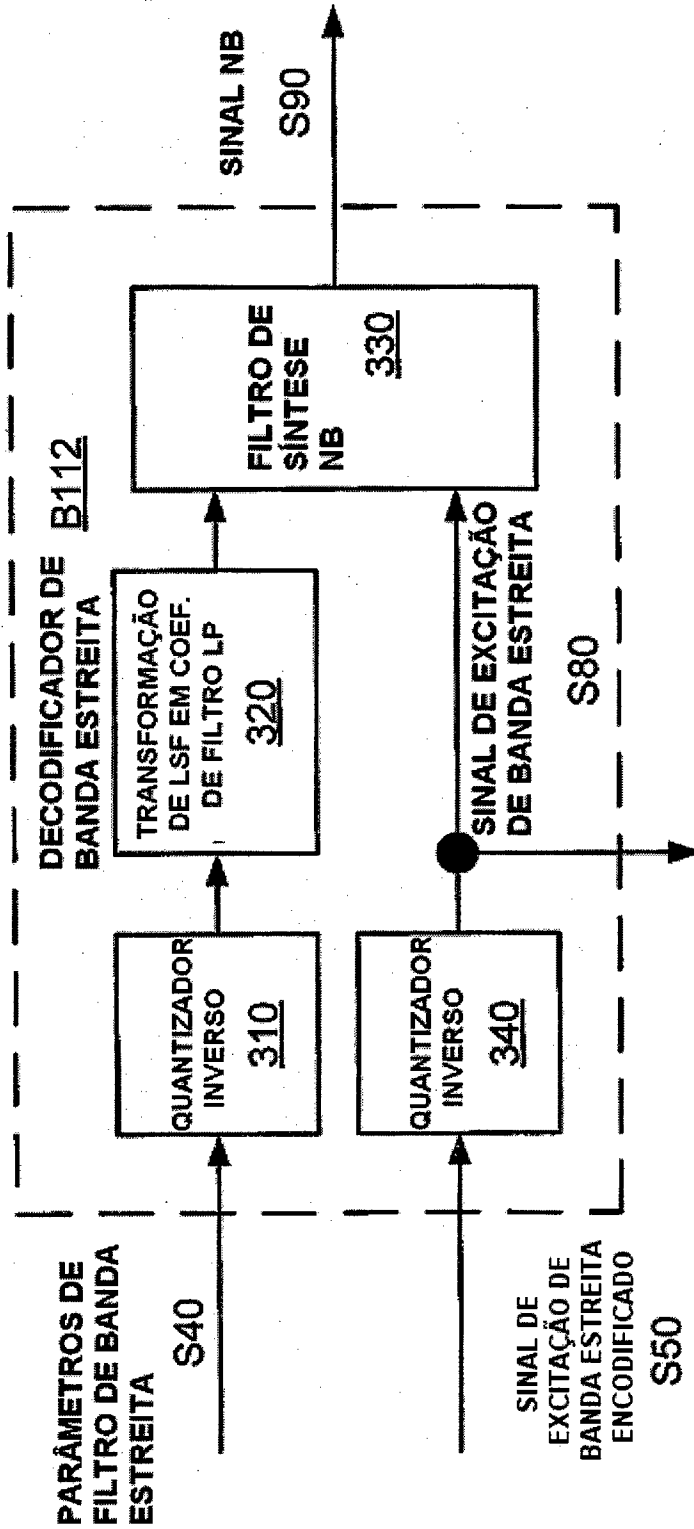
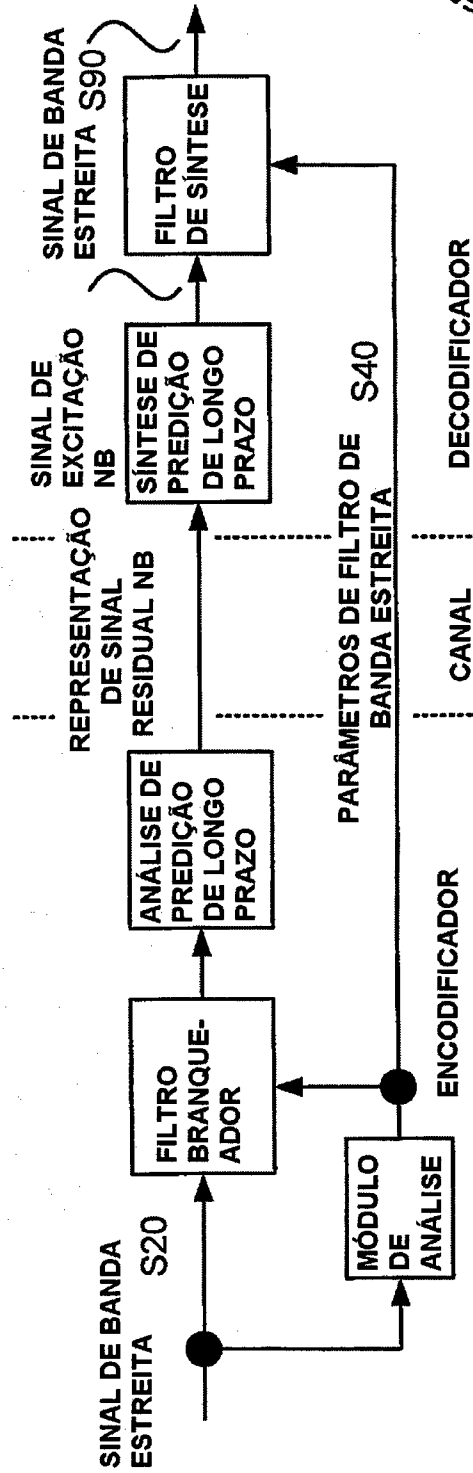
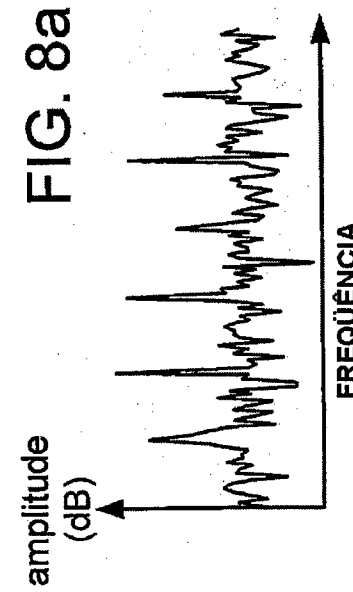
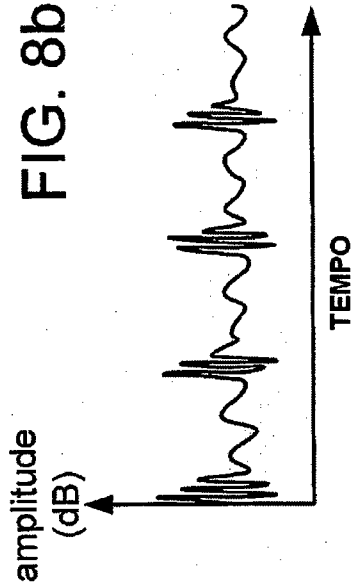


FIG. 7



**FIG. 9**

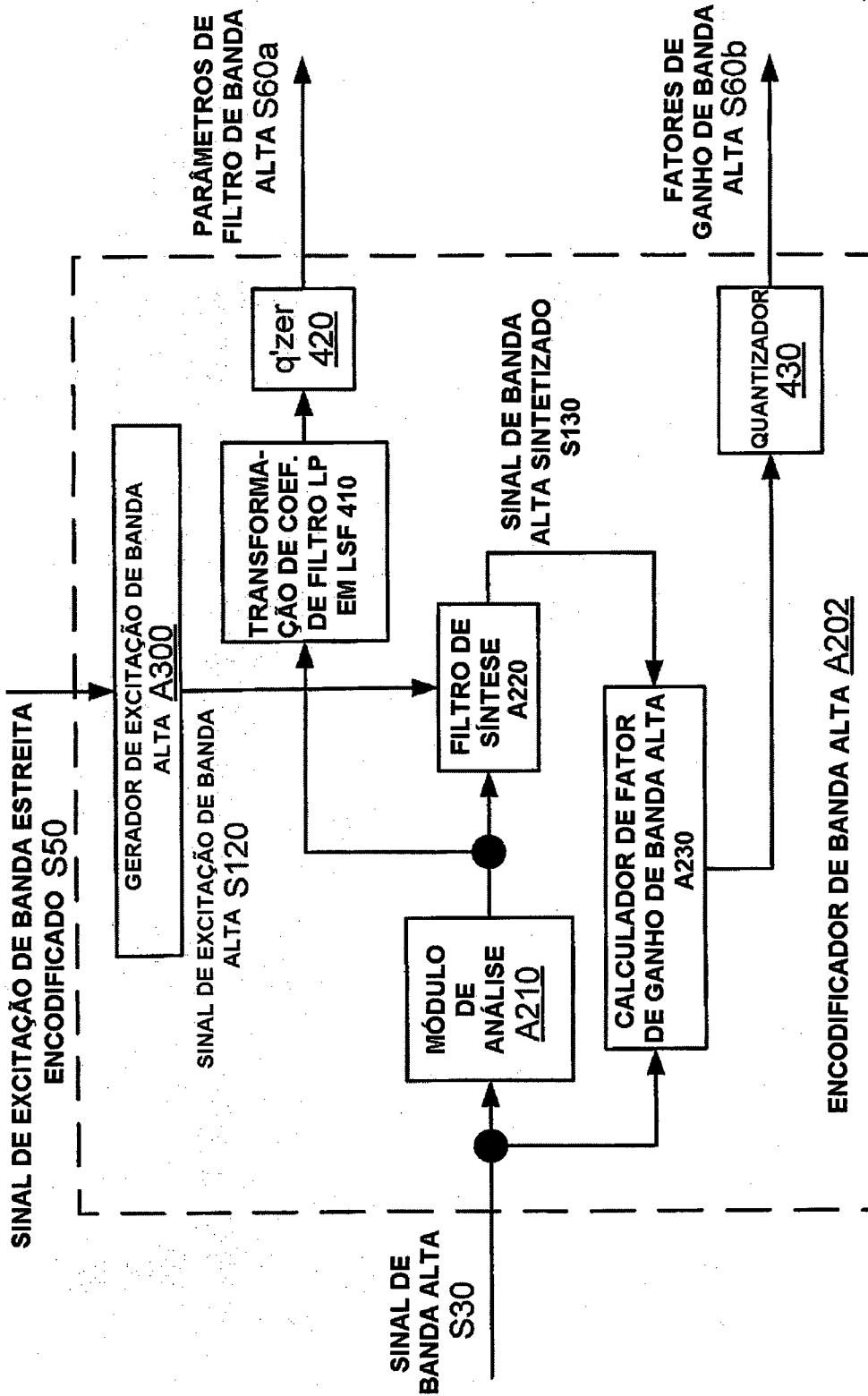


FIG. 10

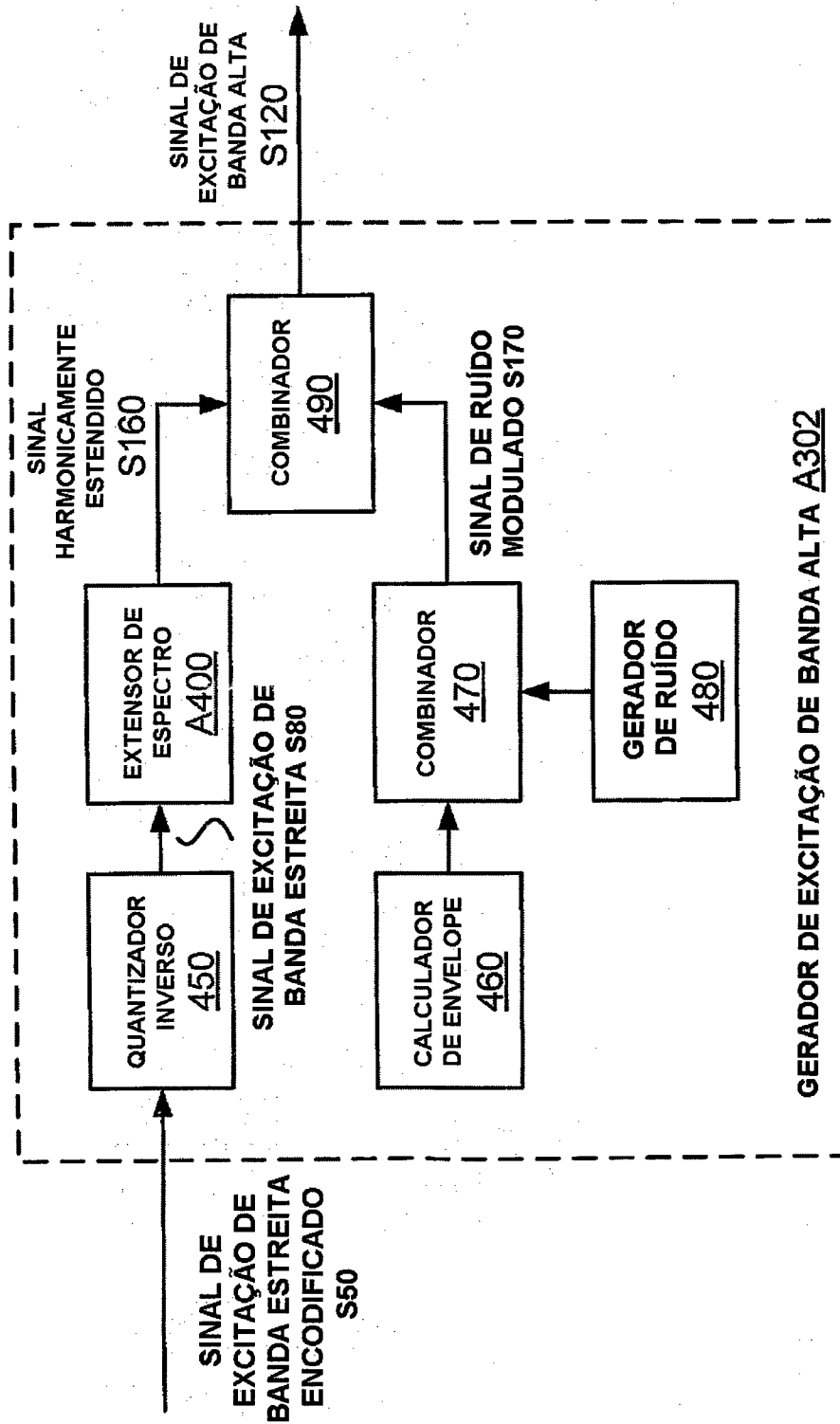


FIG. 11

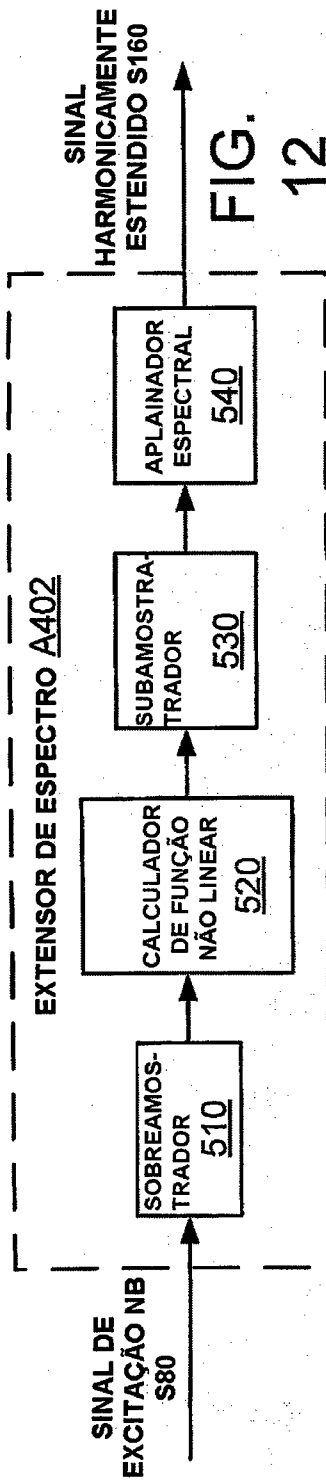


FIG. 12

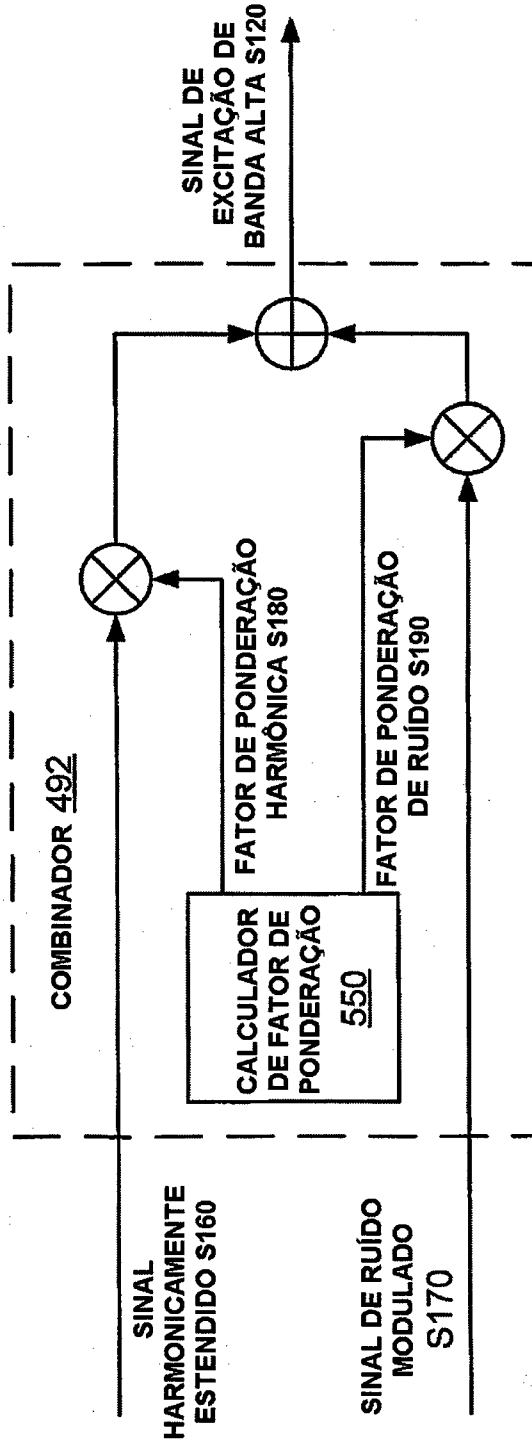


FIG. 16

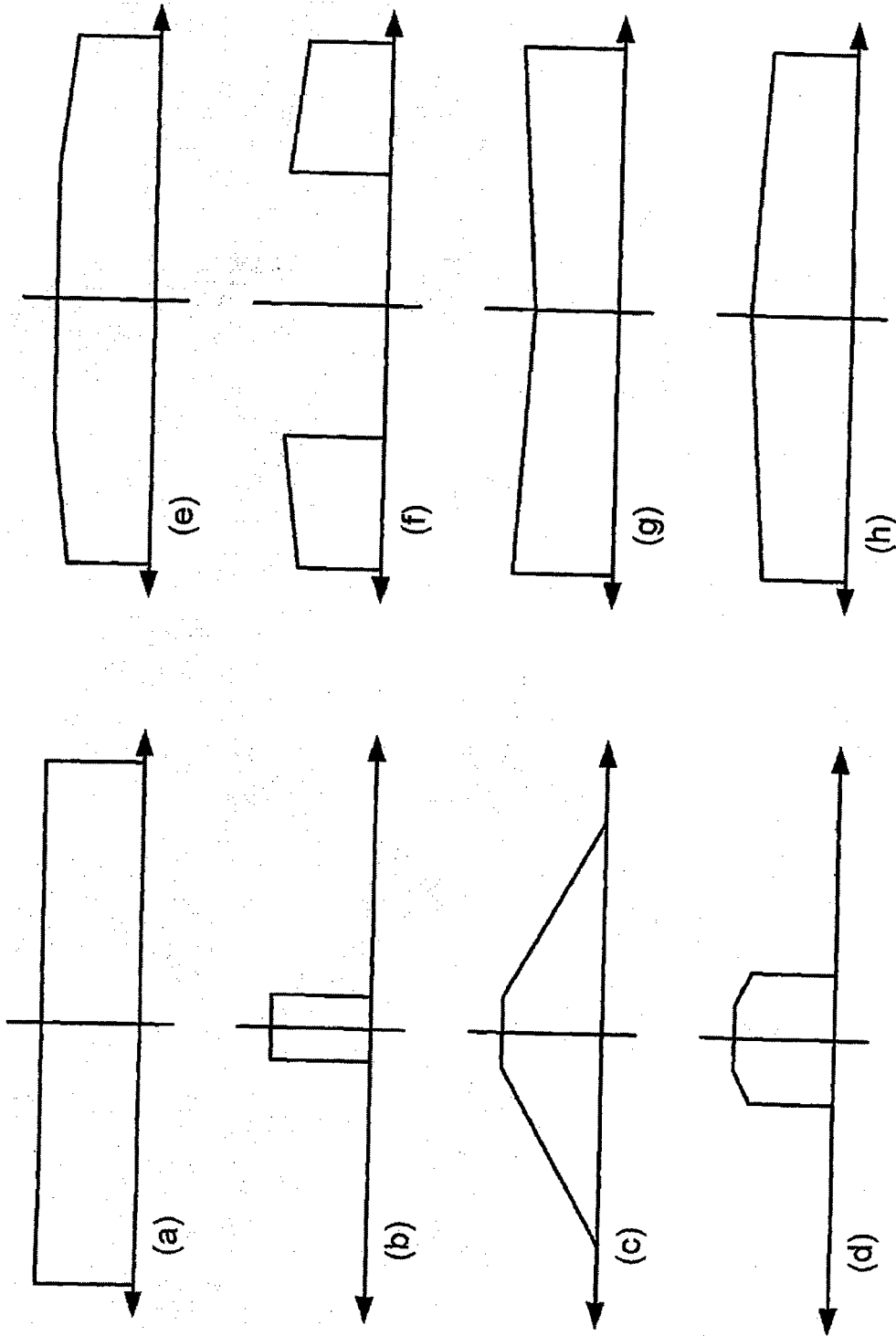


FIG. 12a

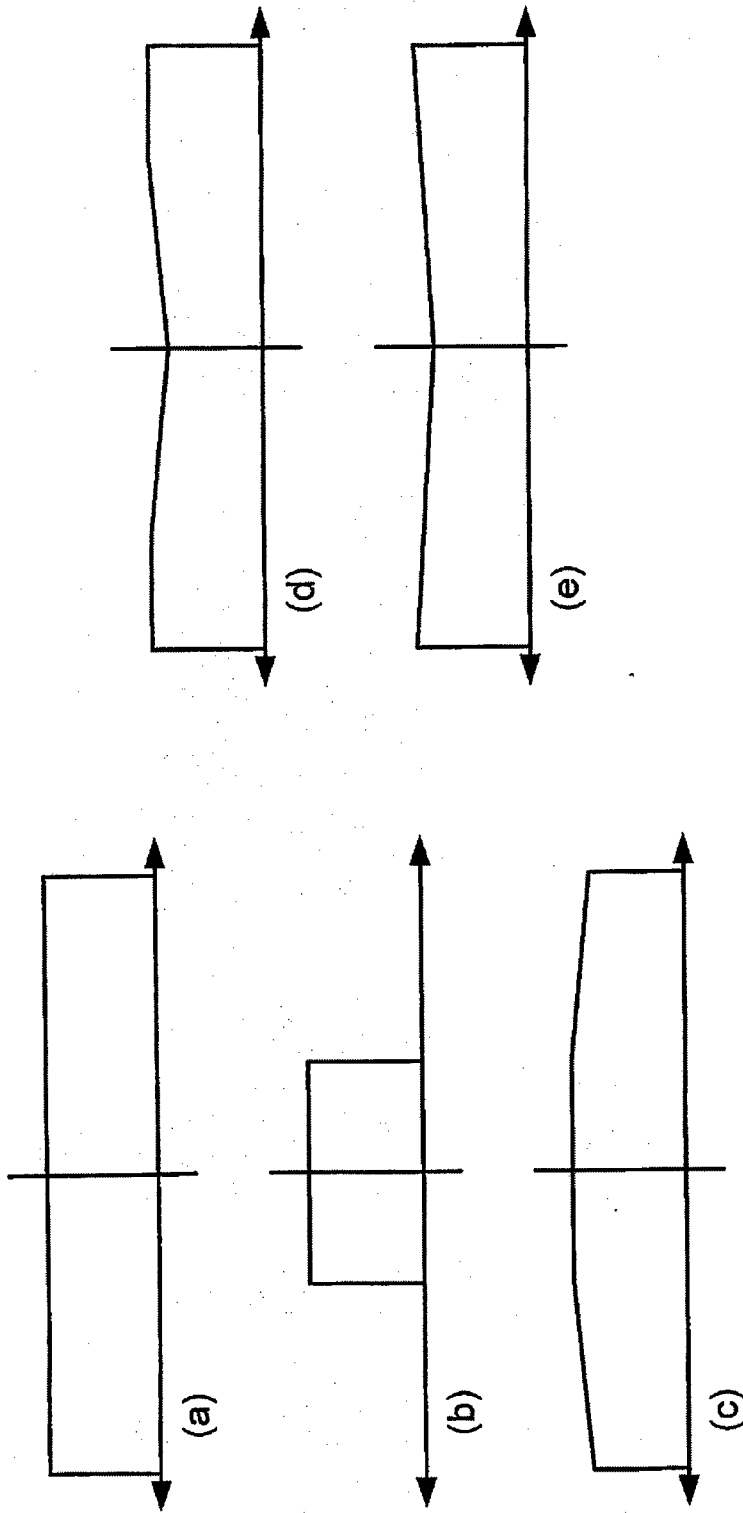


FIG. 12b

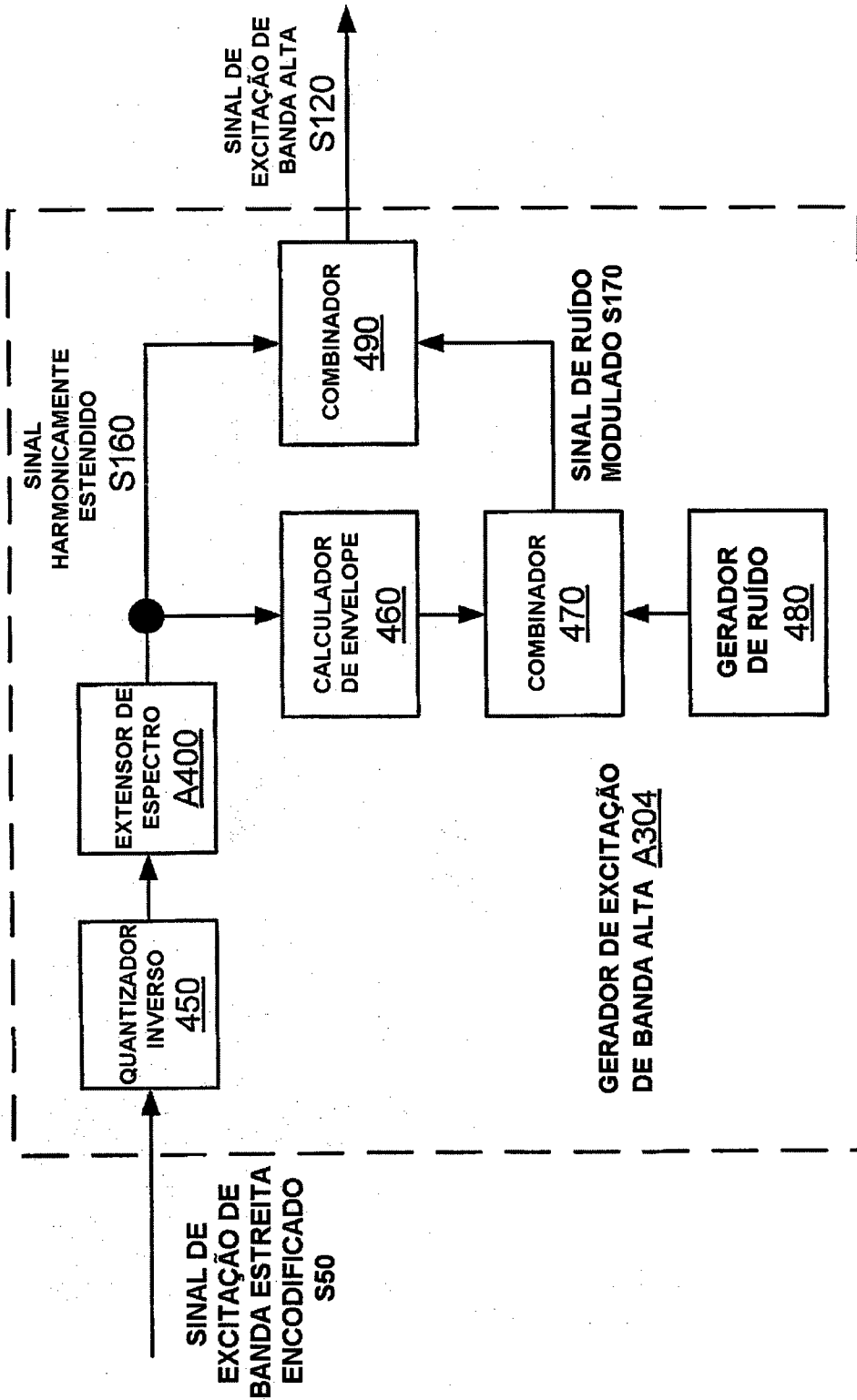


FIG. 13

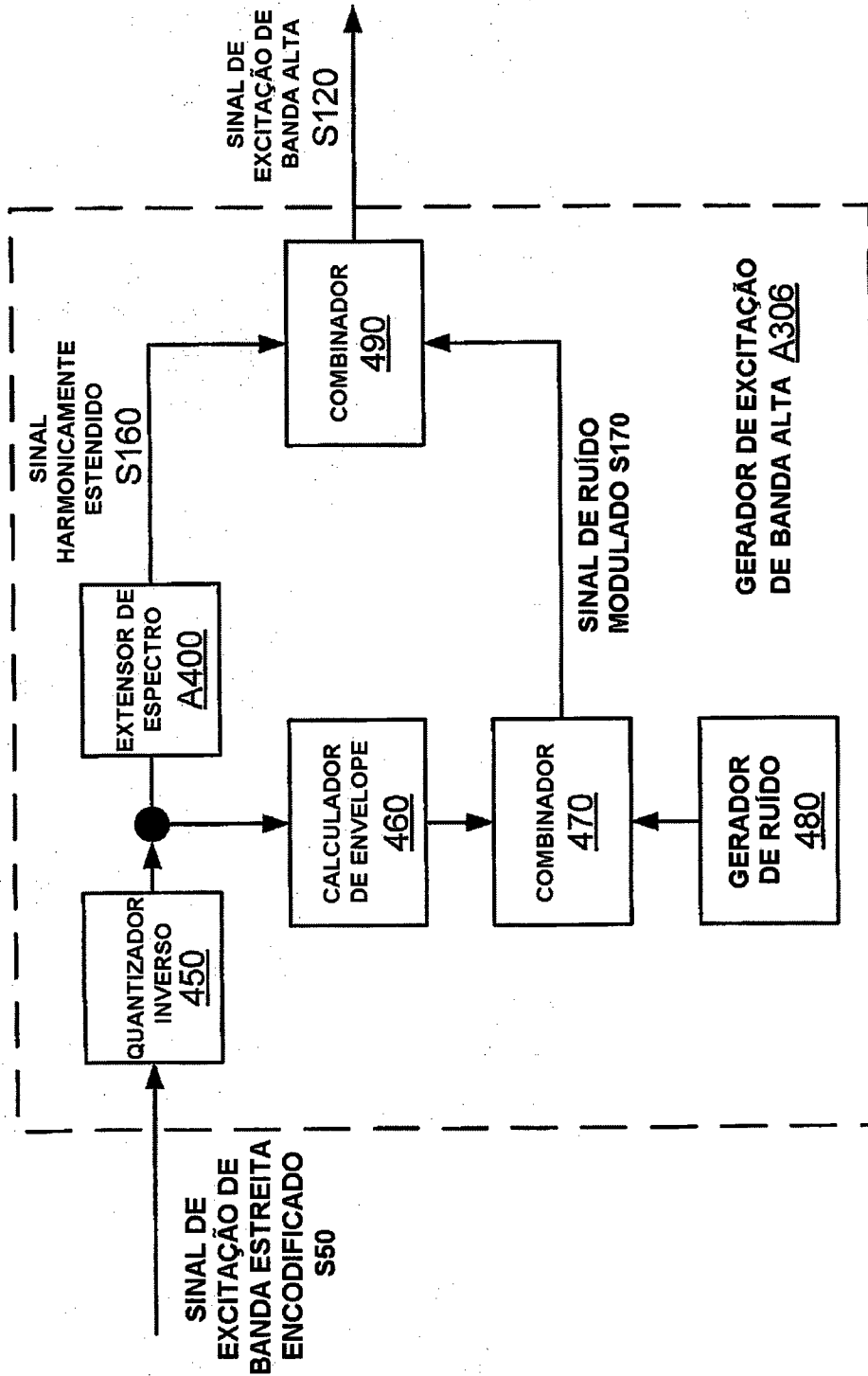


FIG. 14

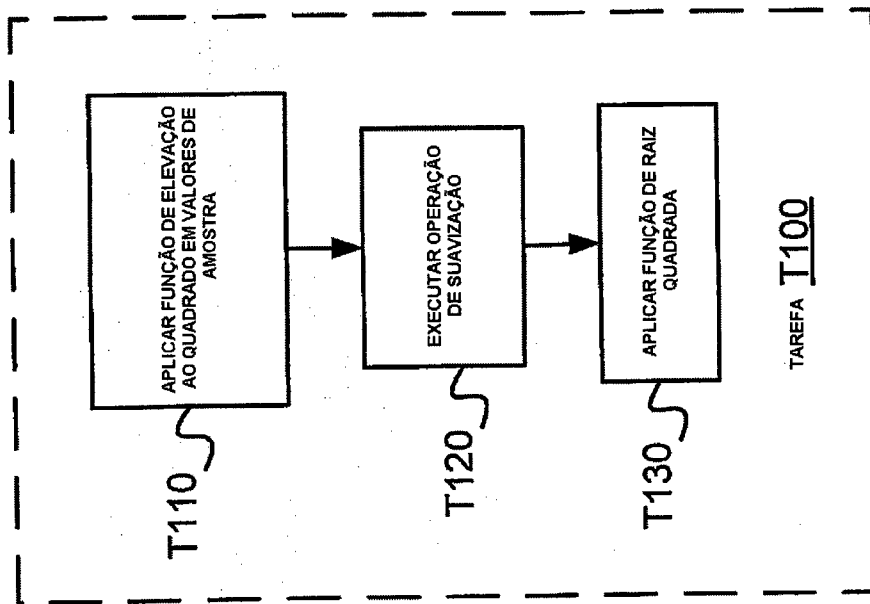


FIG. 15

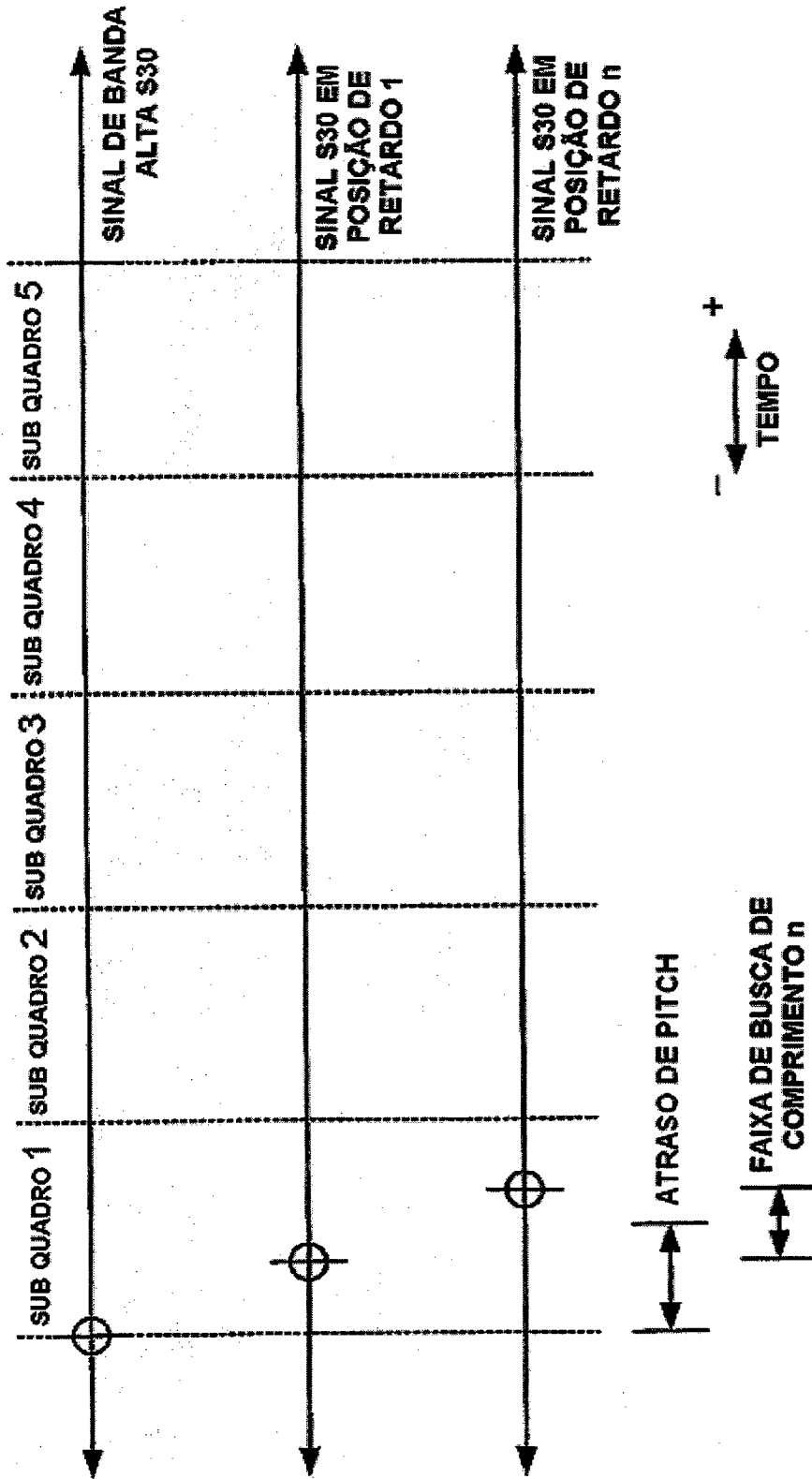


FIG. 17

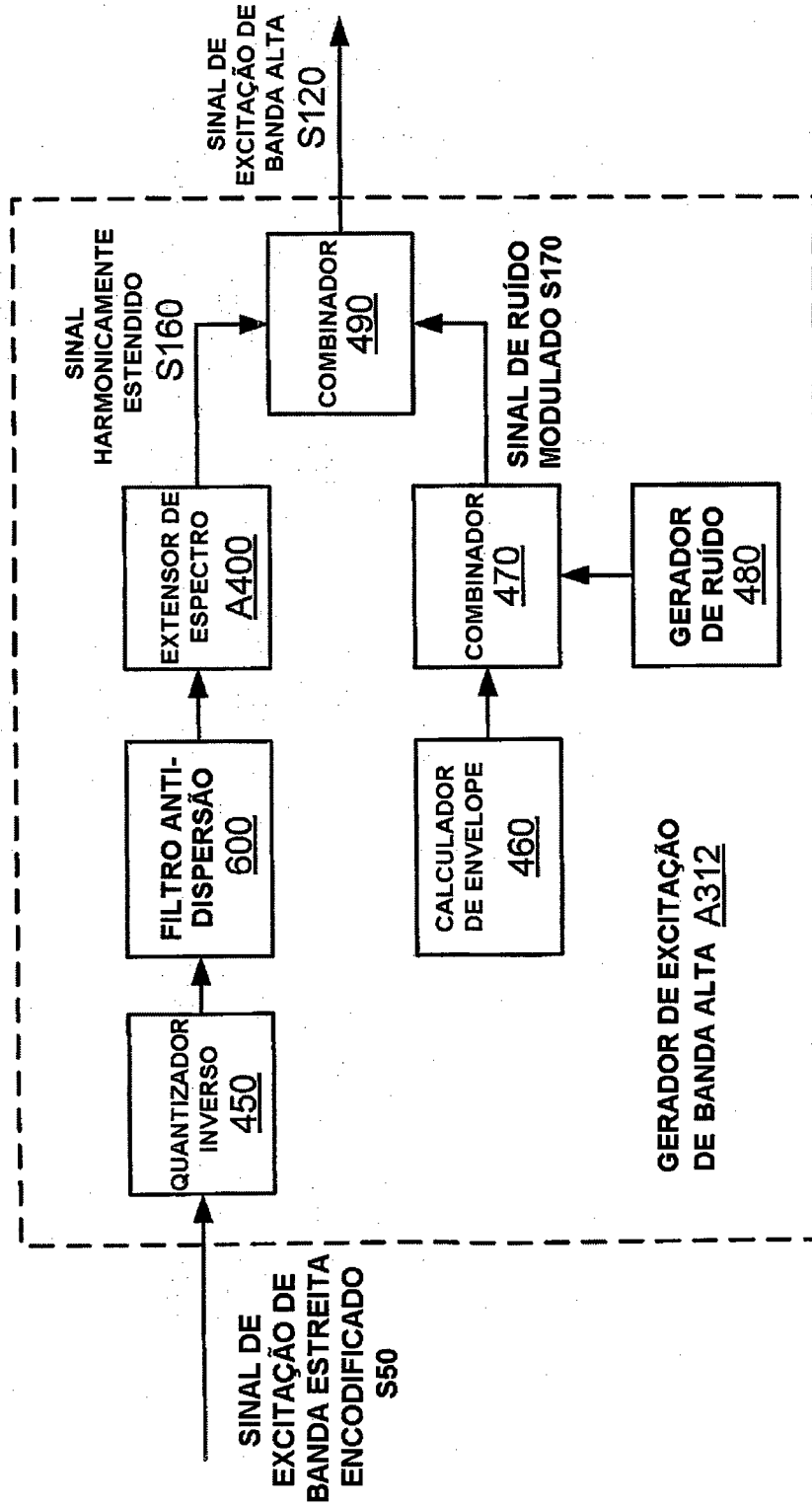


FIG. 18

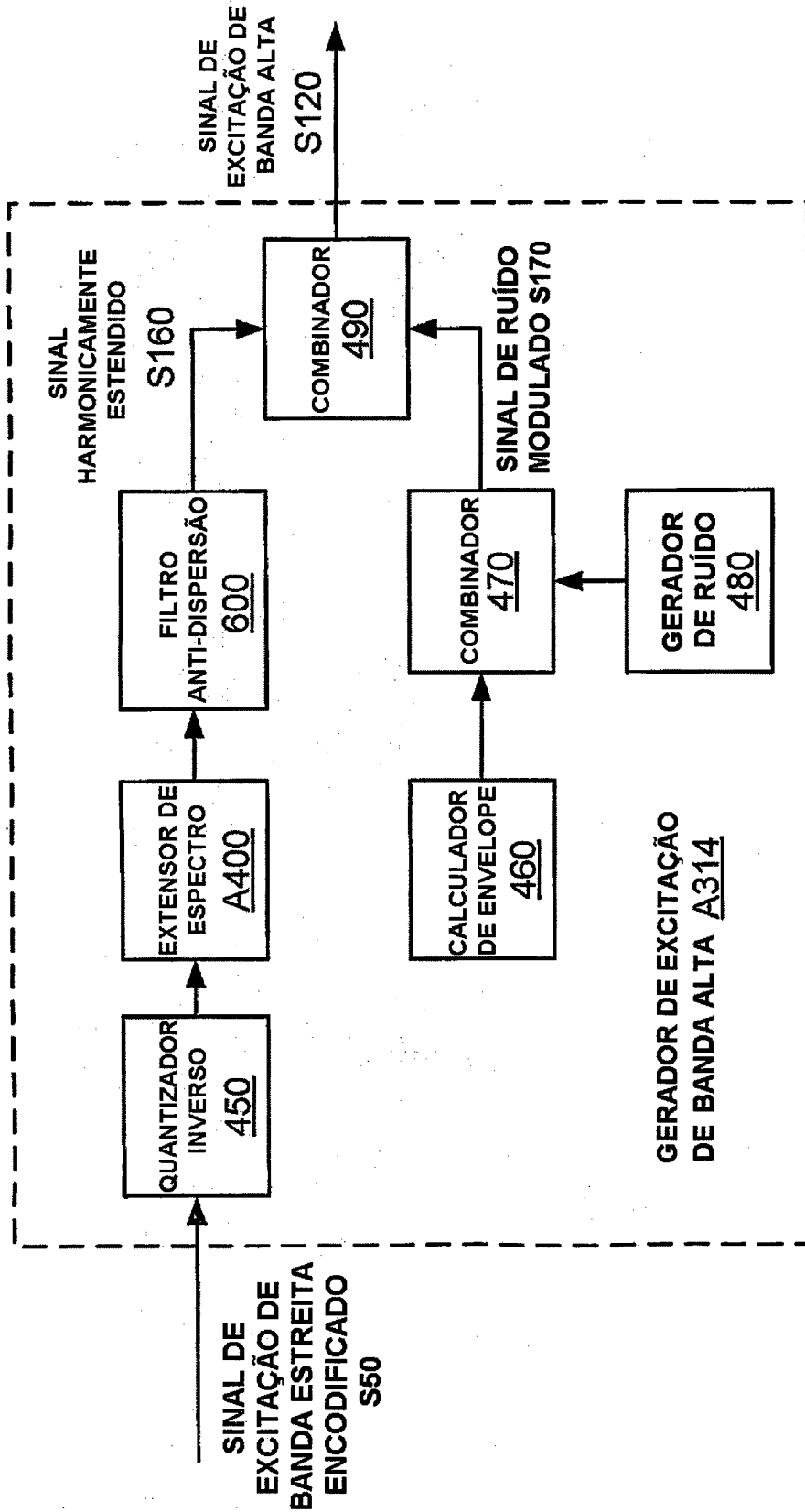


FIG. 19

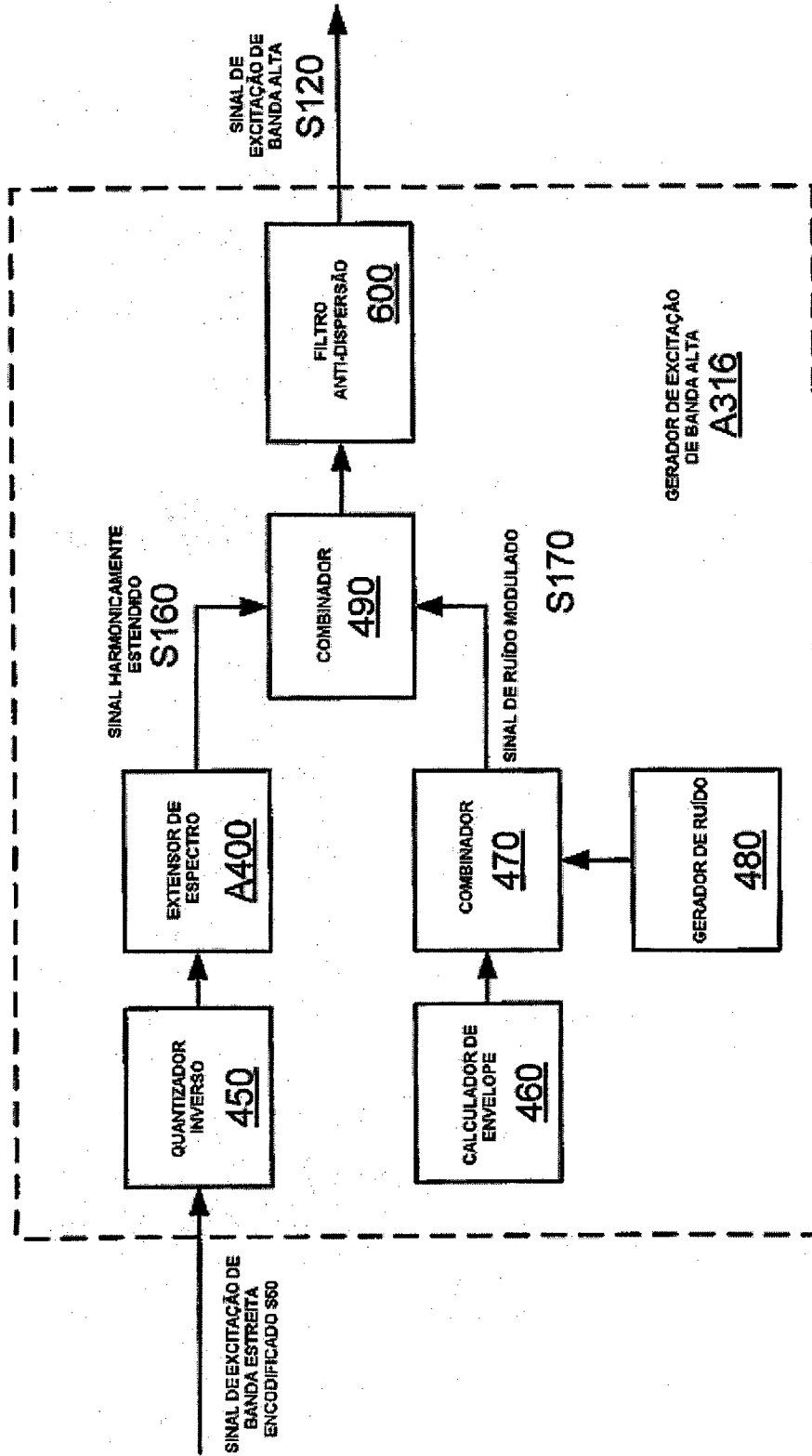


FIG. 20

I. Nacional da Propriedade Industrial - SFP/DOC -  
Fls. 284  
Rub: de

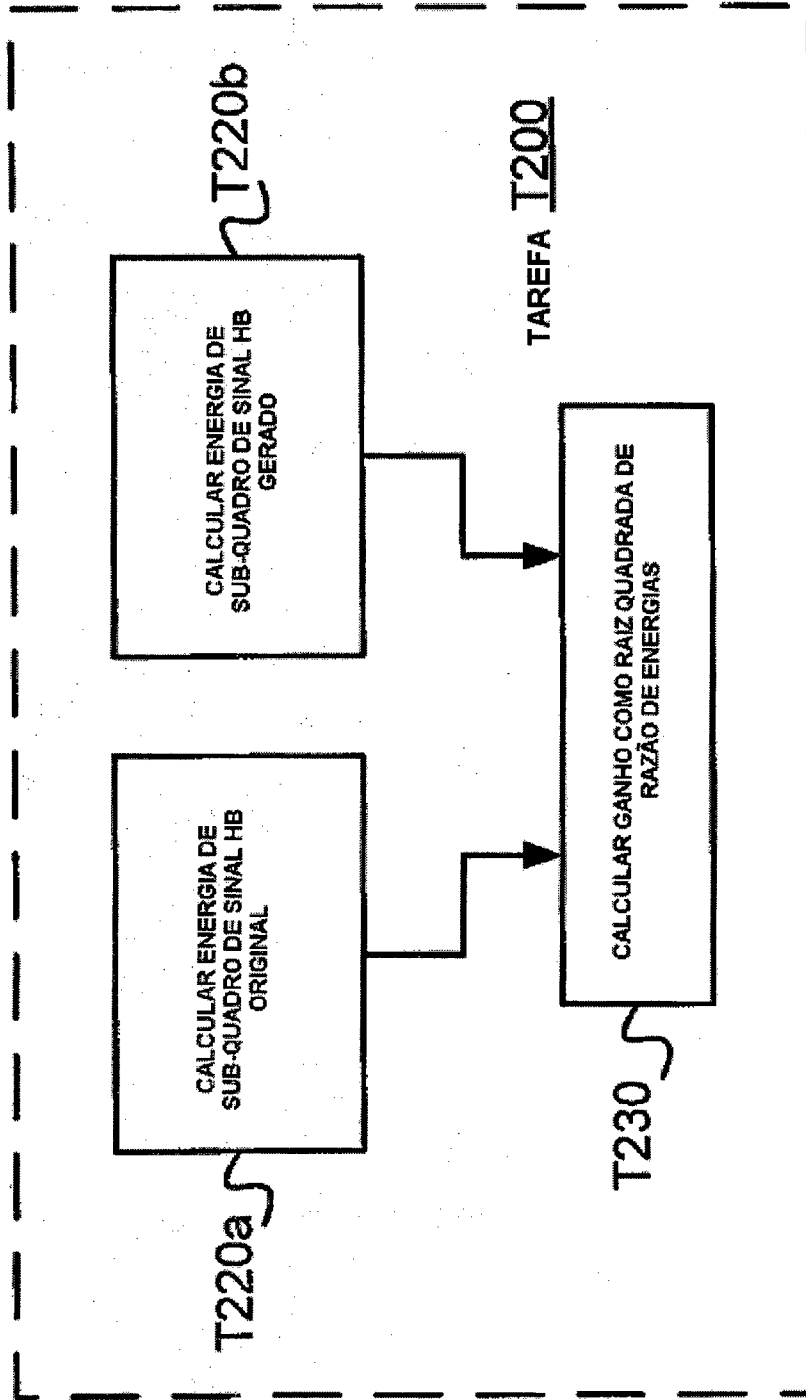


FIG. 21

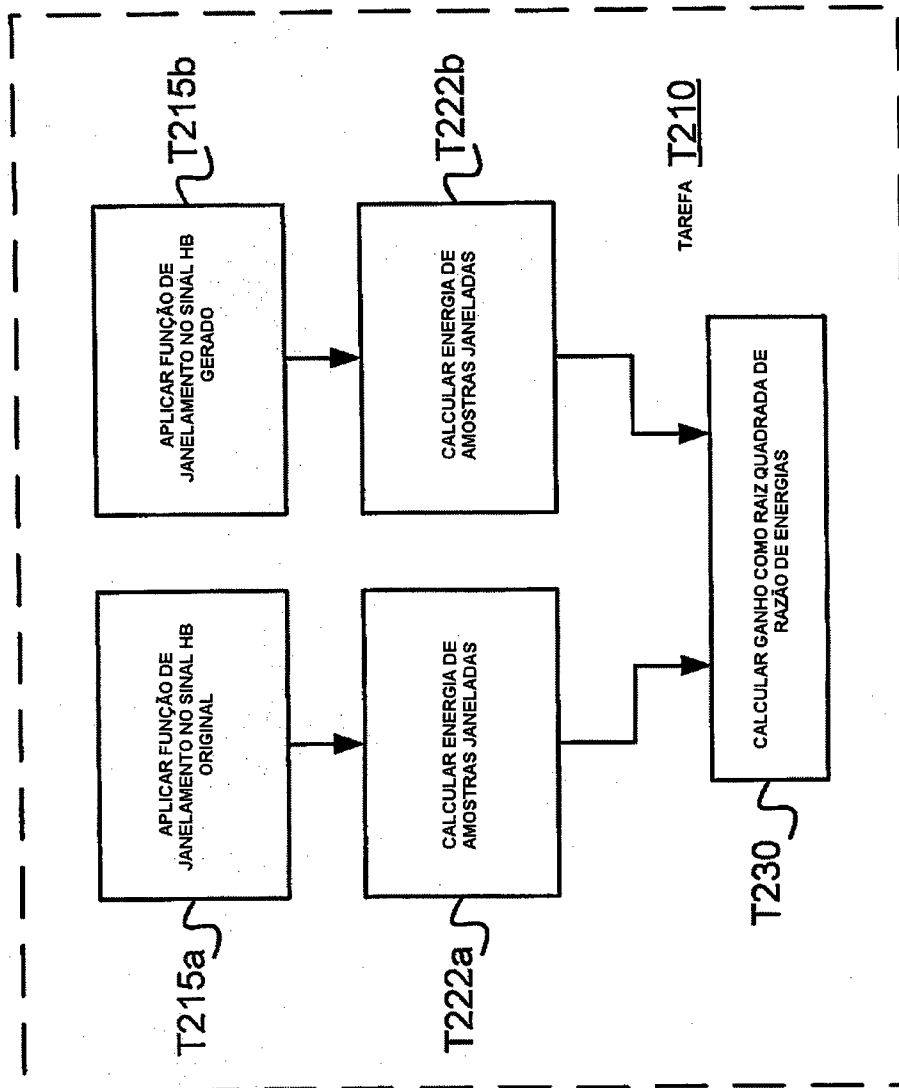
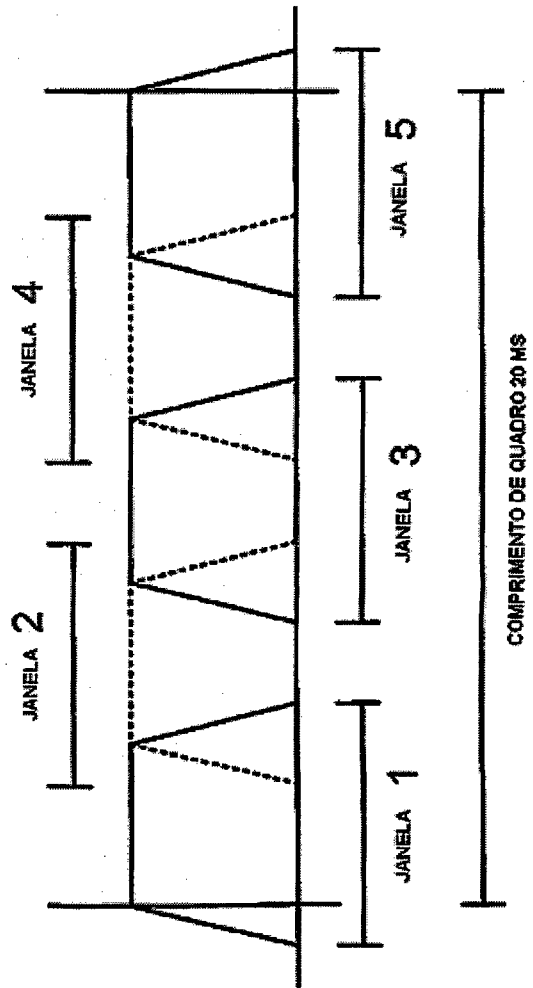
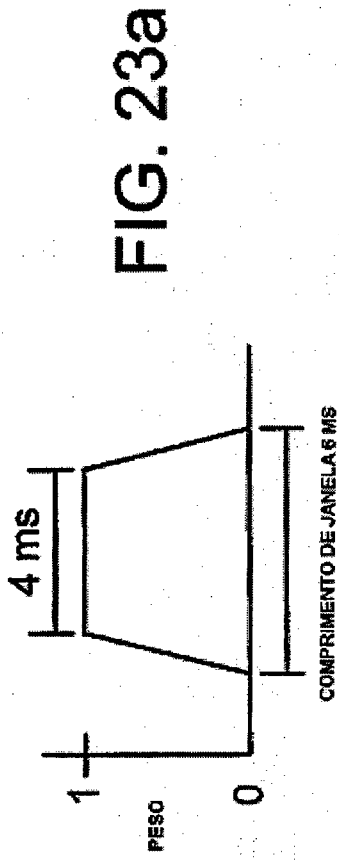
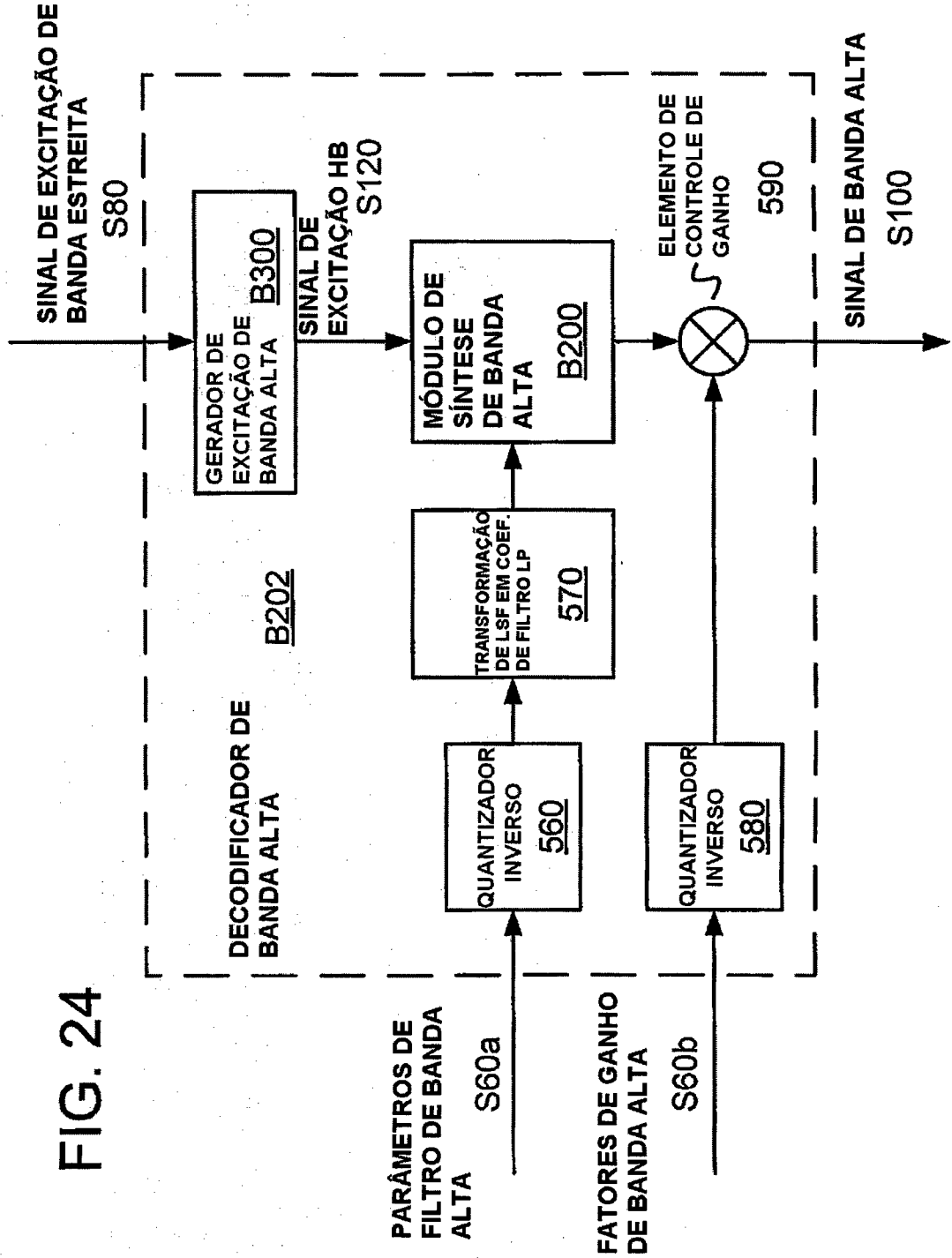


FIG. 22





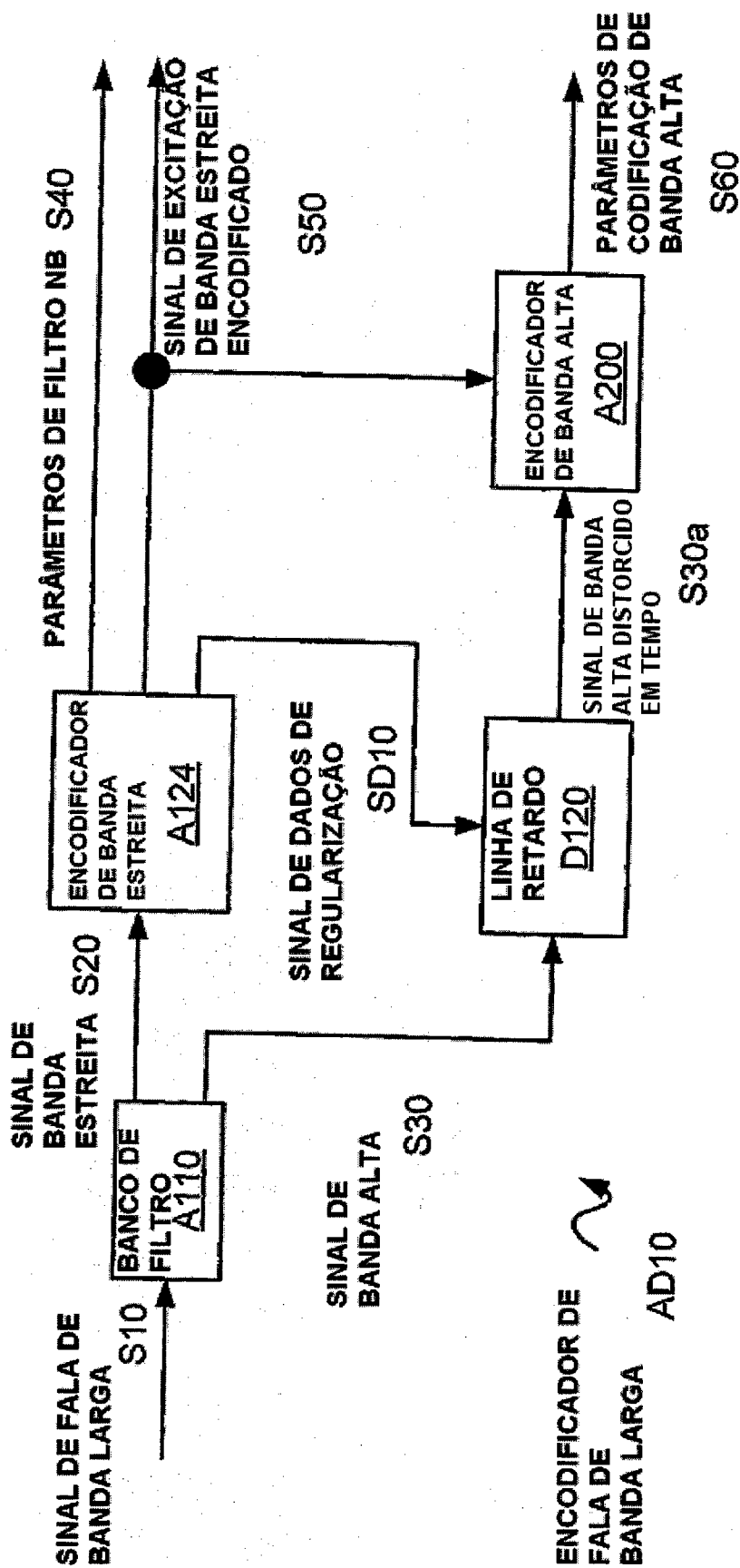
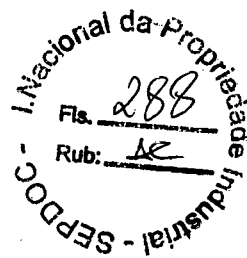


FIG. 25



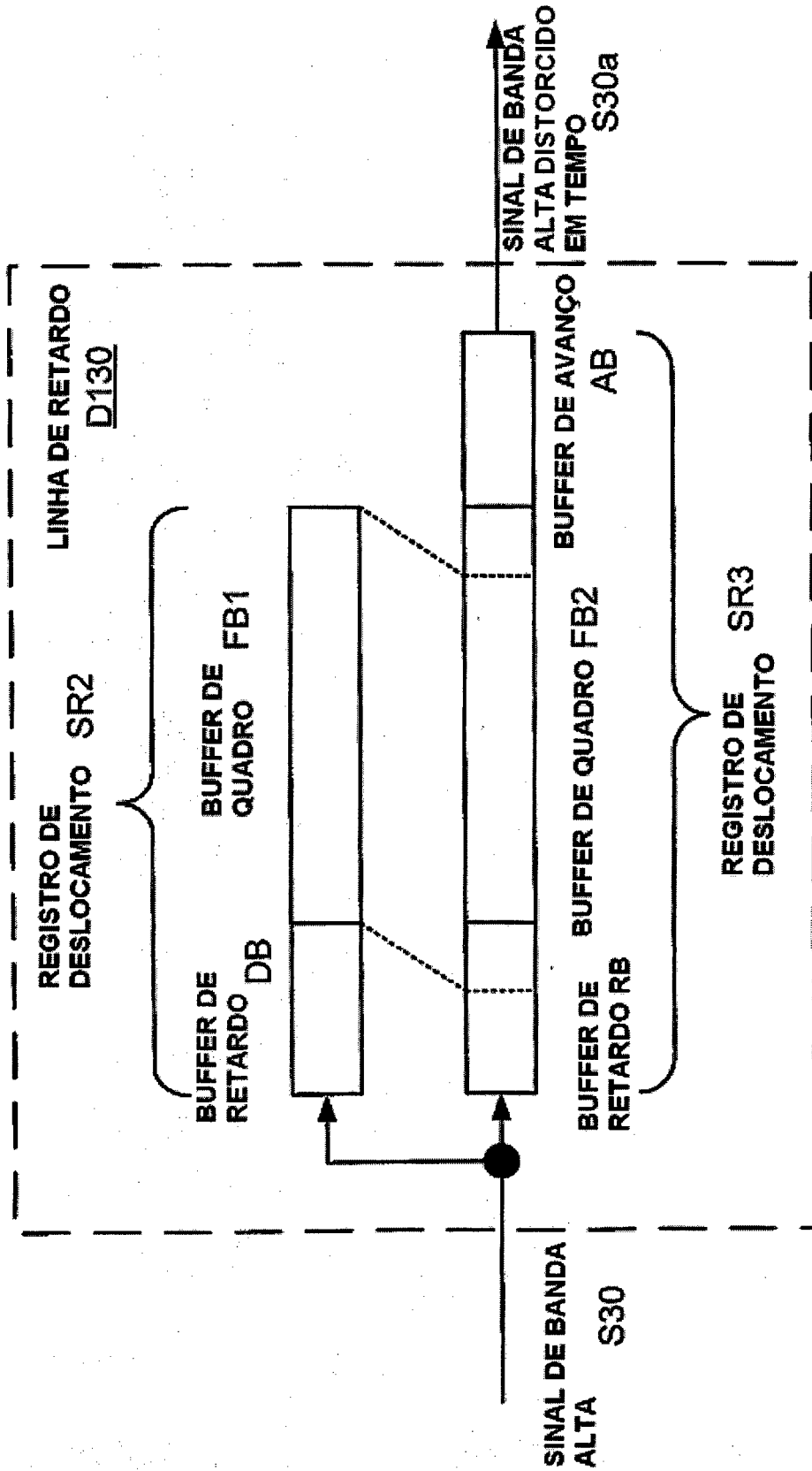


FIG. 27

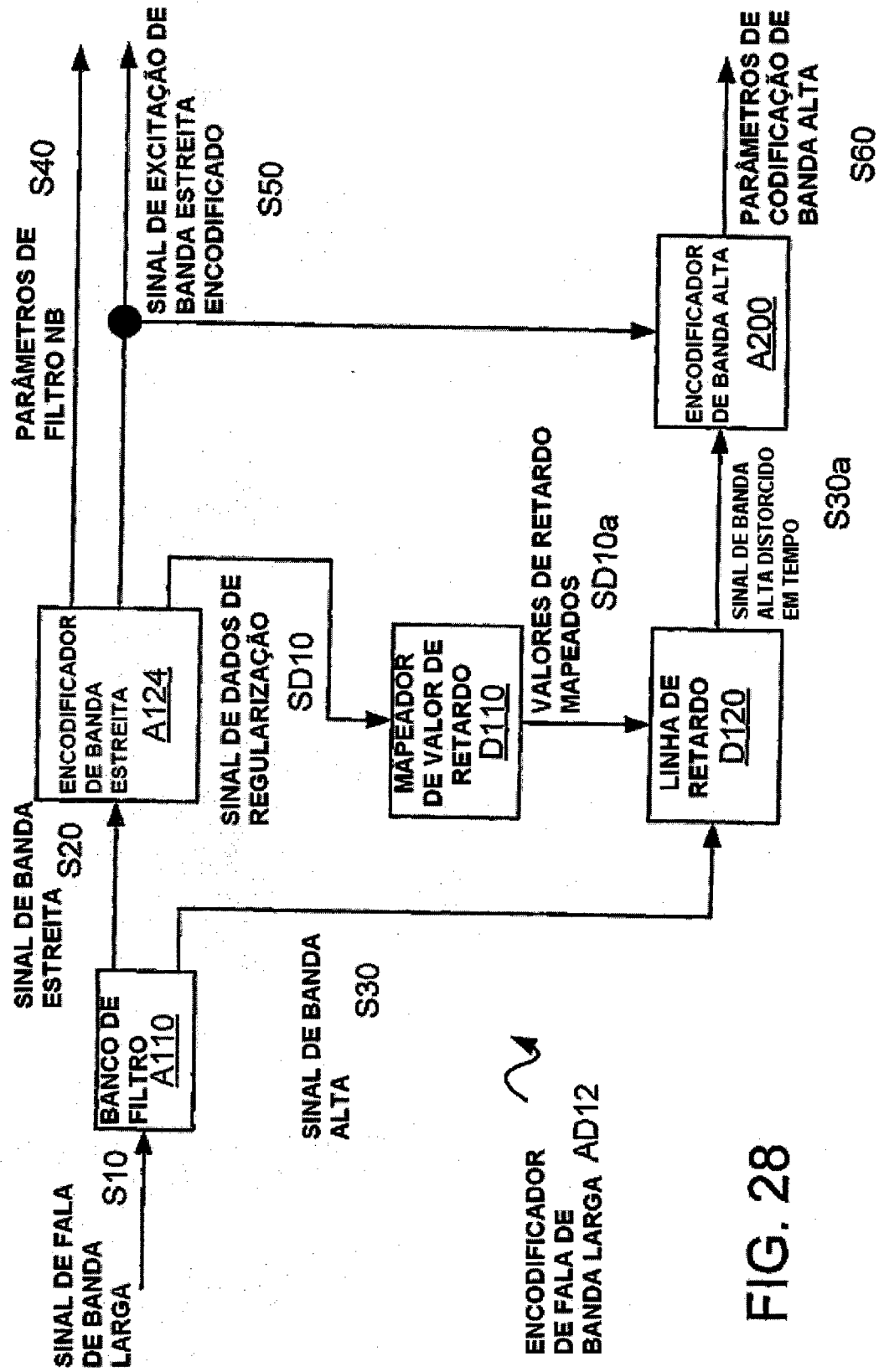


FIG. 28

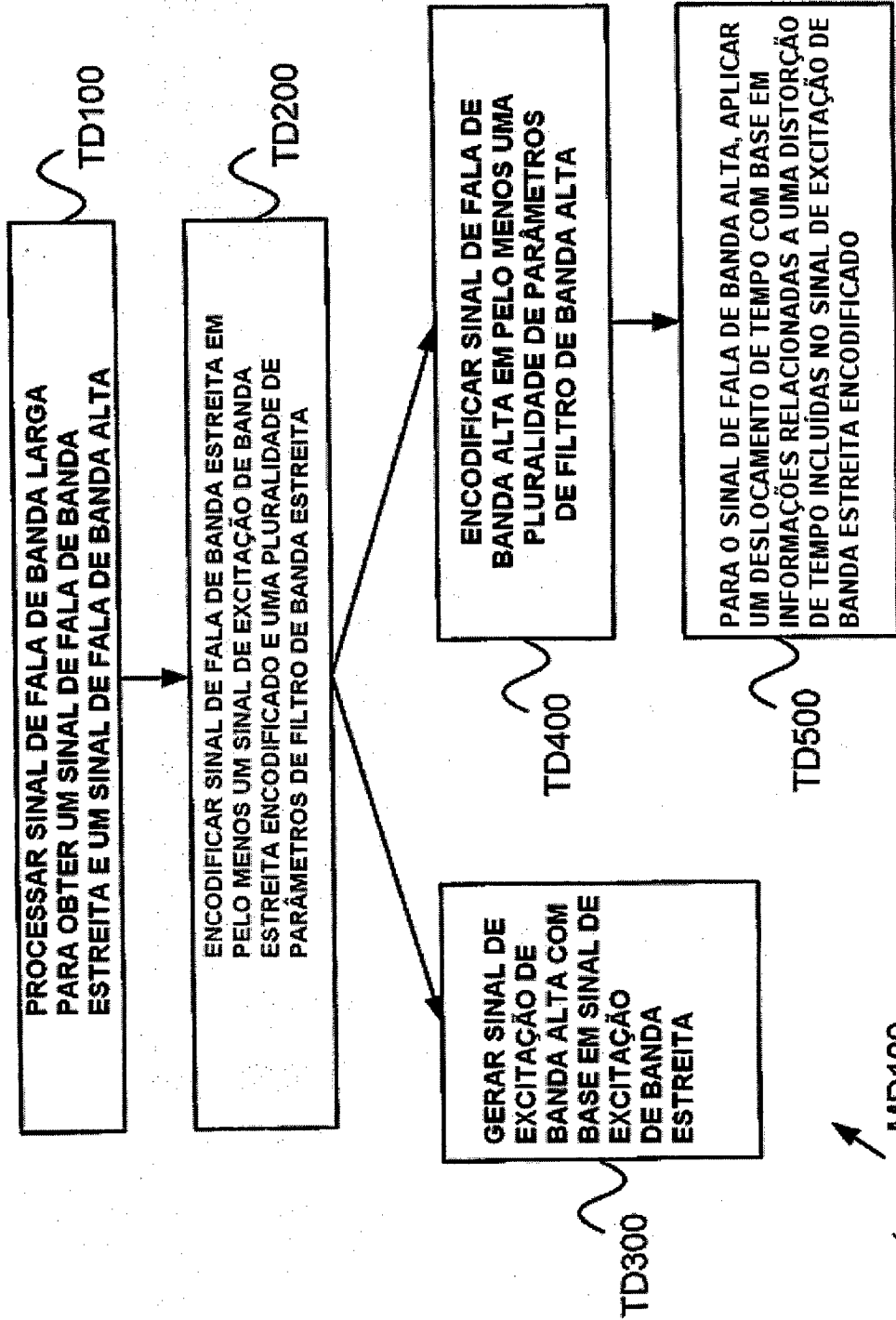


FIG. 29

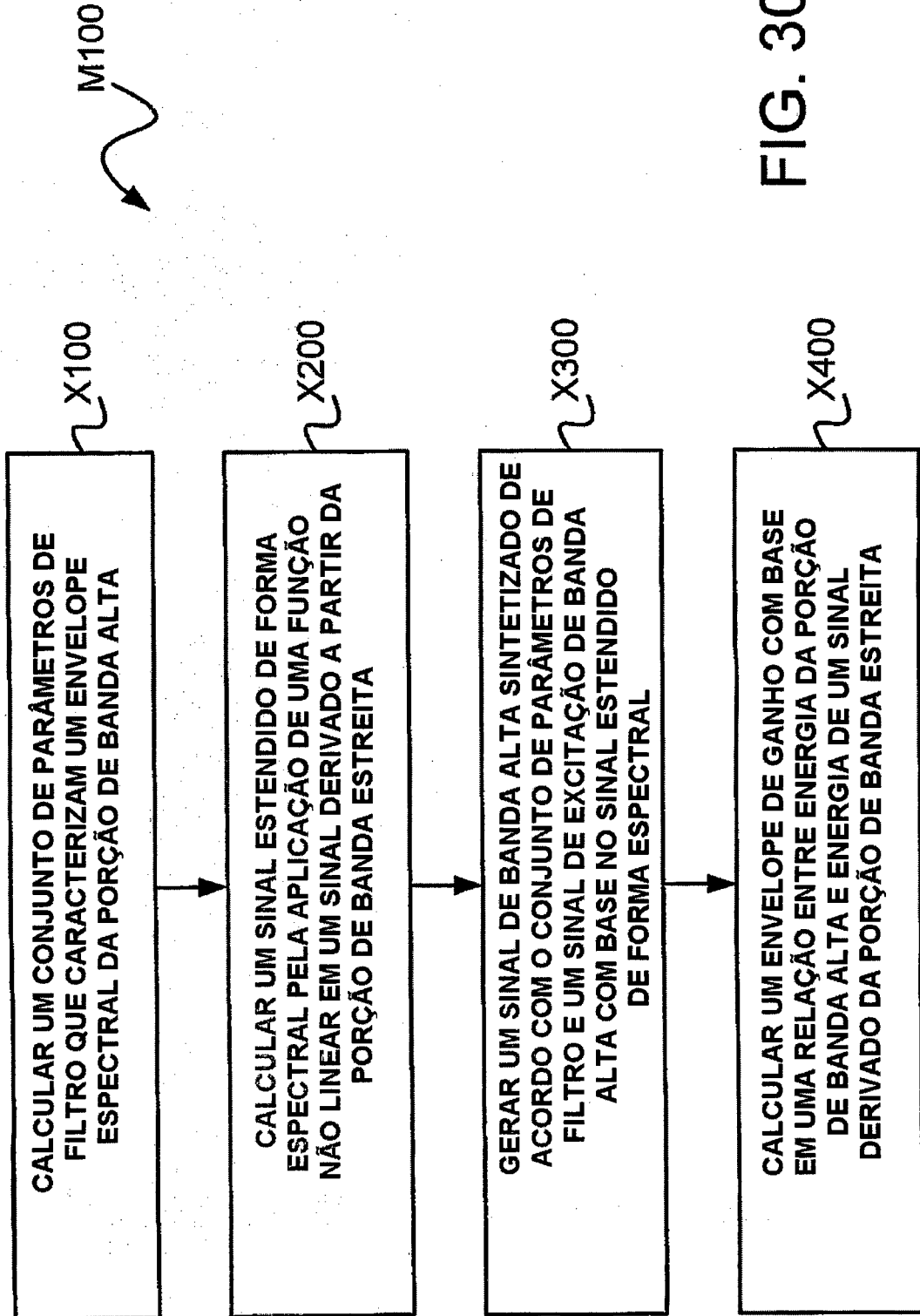


FIG. 30

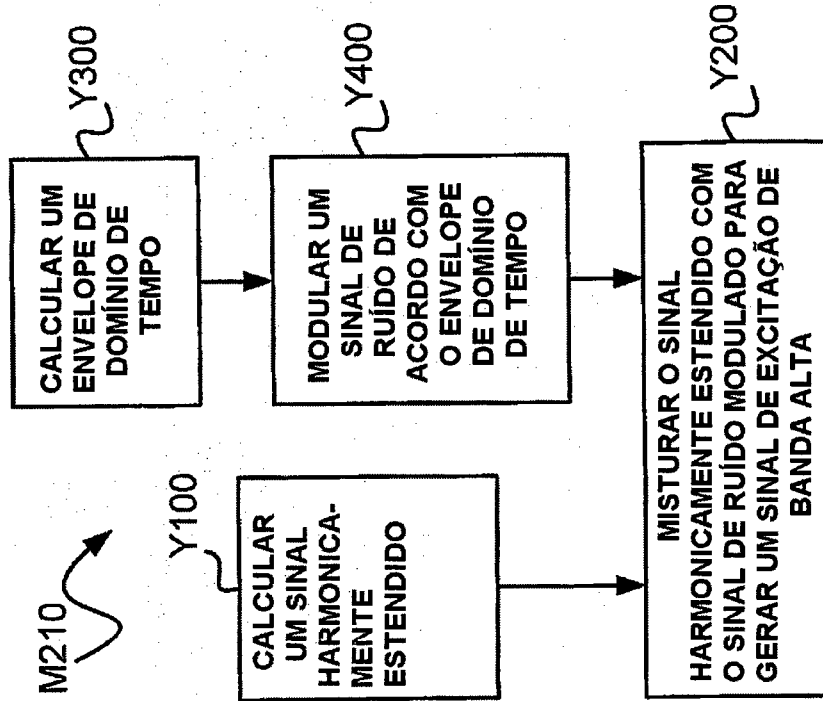


FIG. 31b

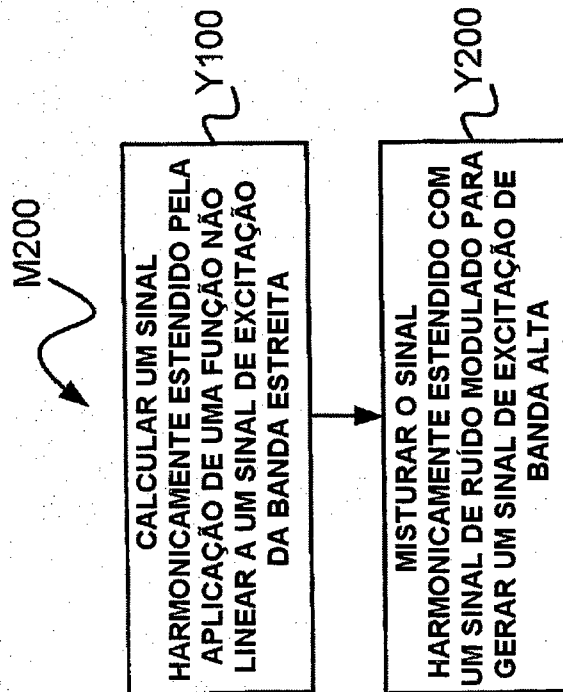


FIG. 31a

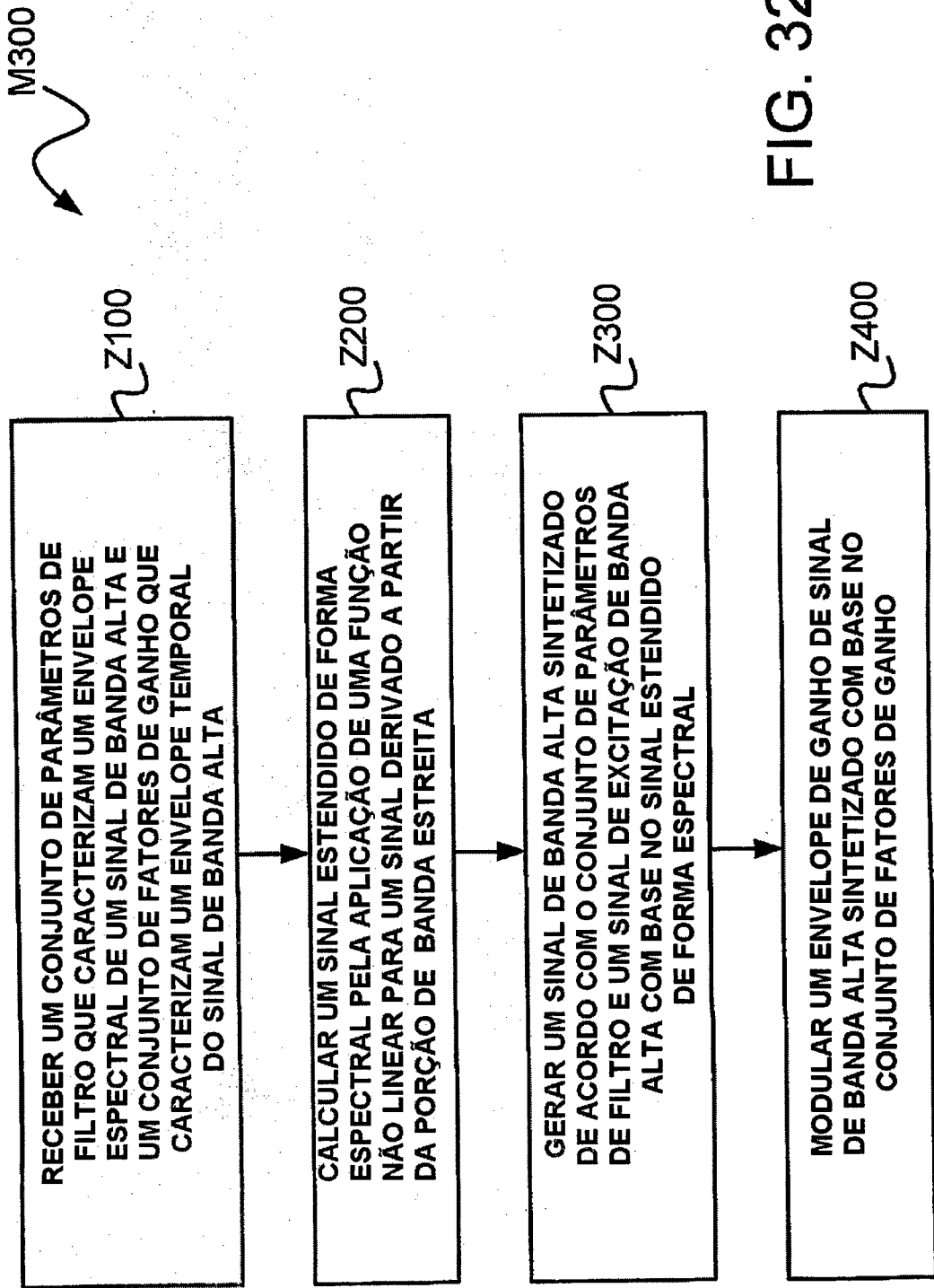
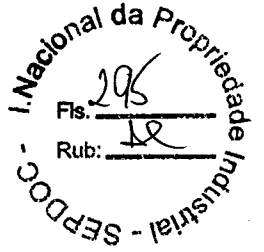


FIG. 32

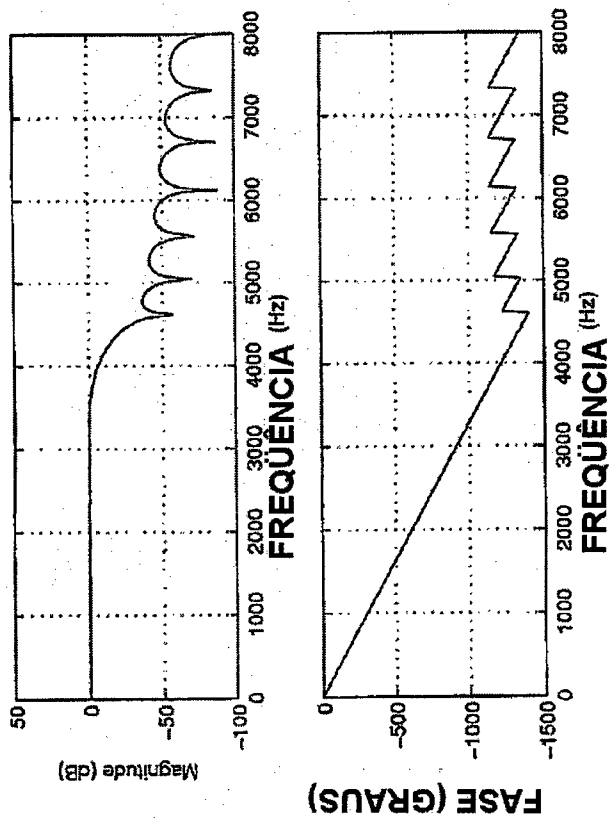
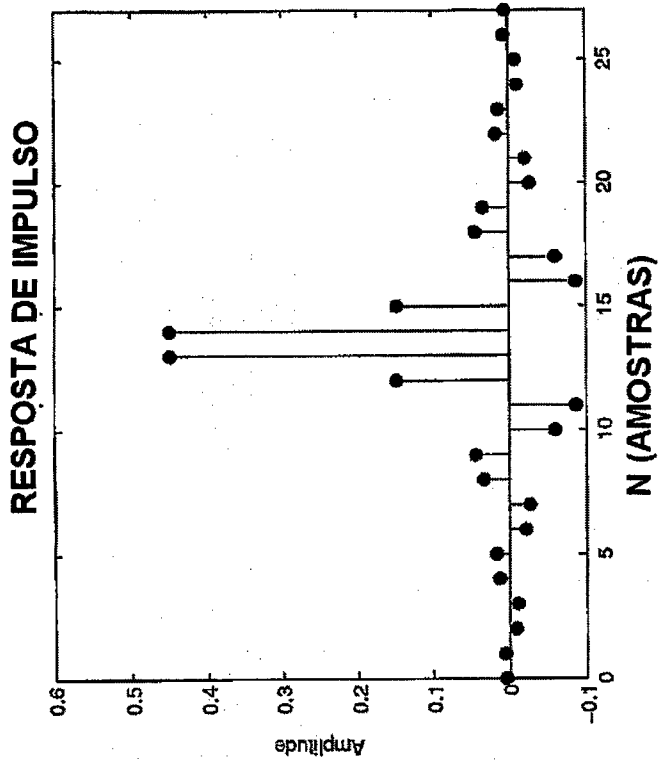


FIG. 33

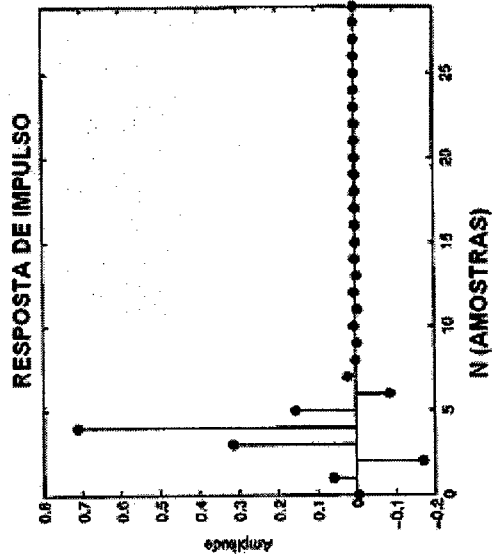
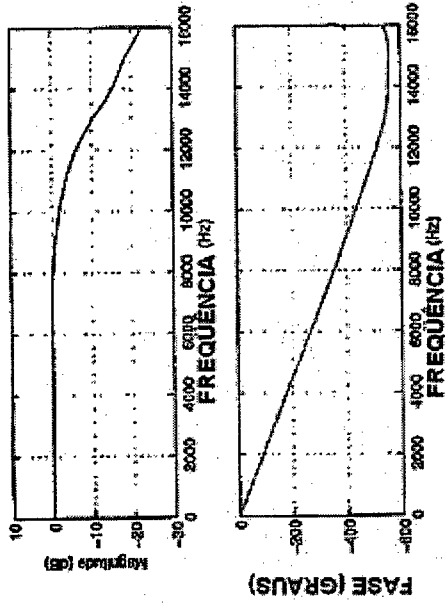


FIG. 34b

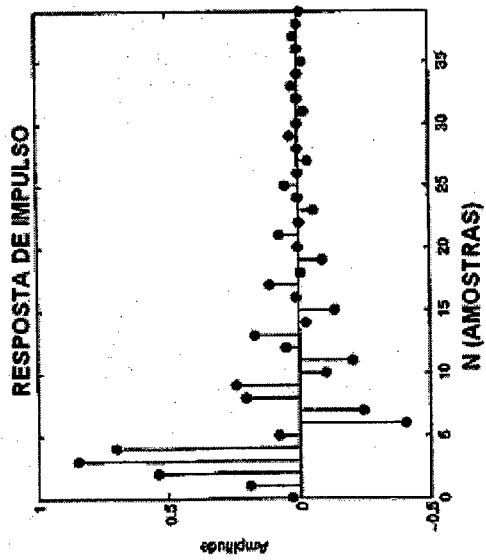
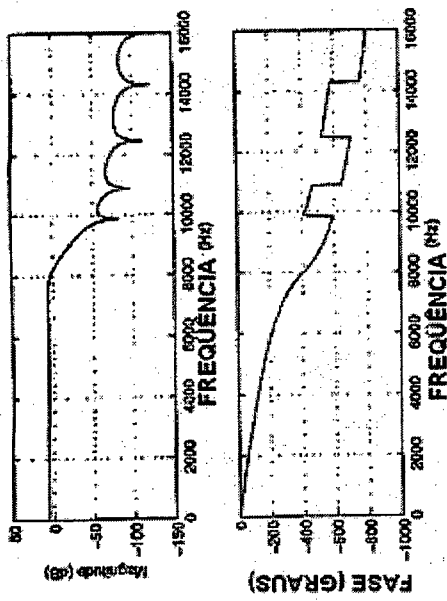


FIG. 34a

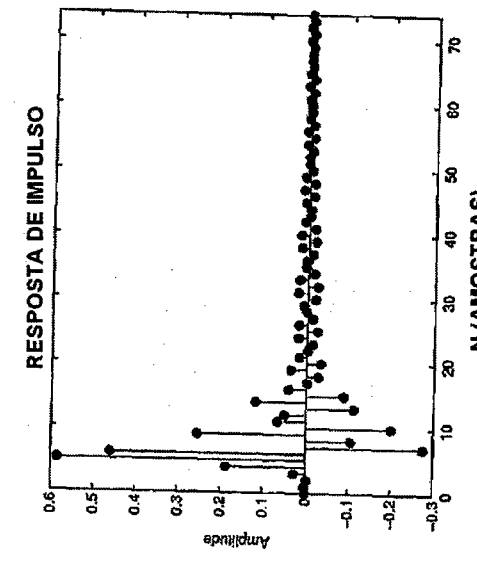
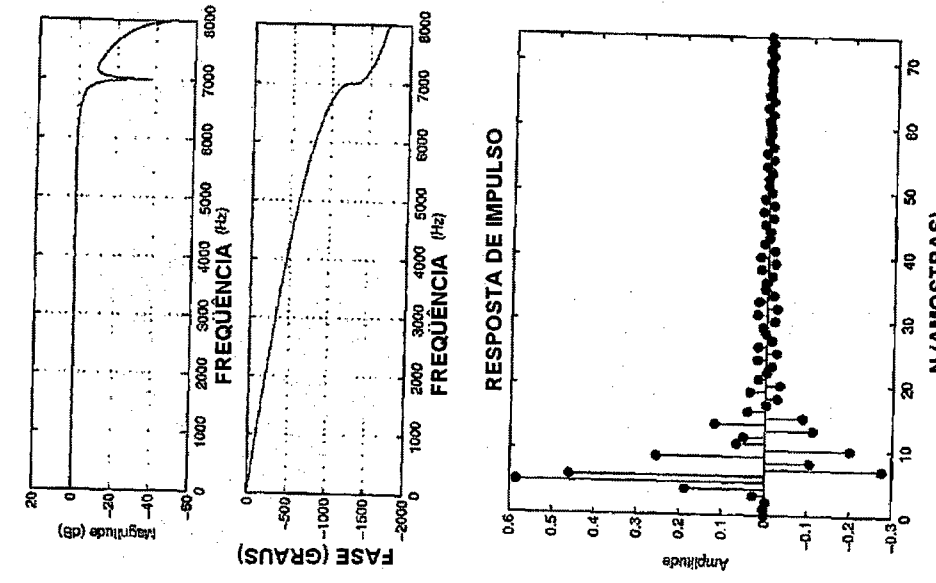


FIG. 35b

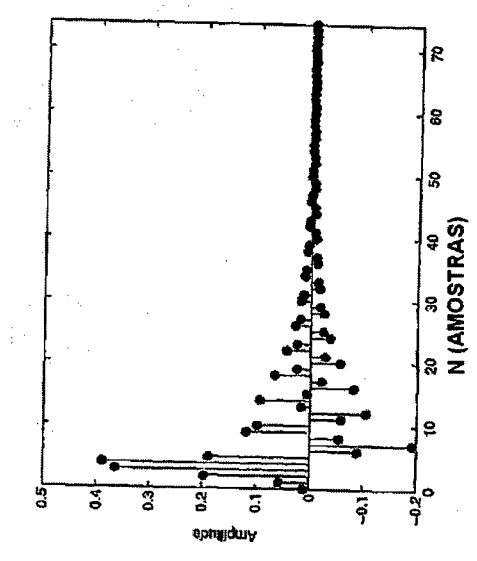
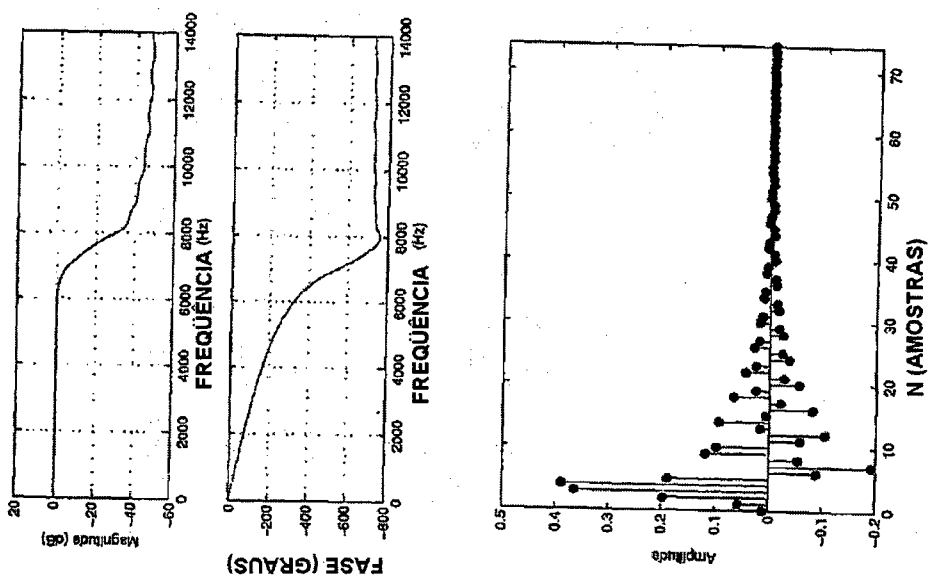


FIG. 35a

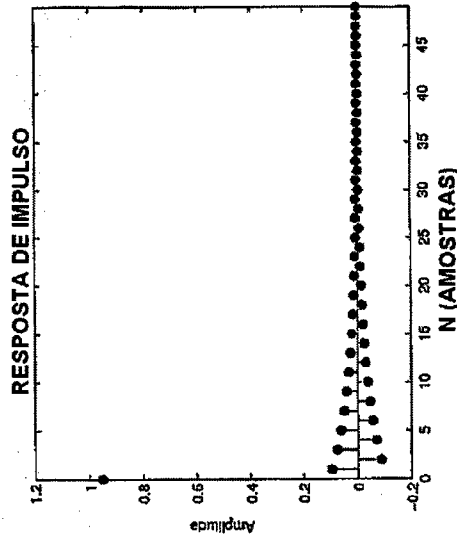
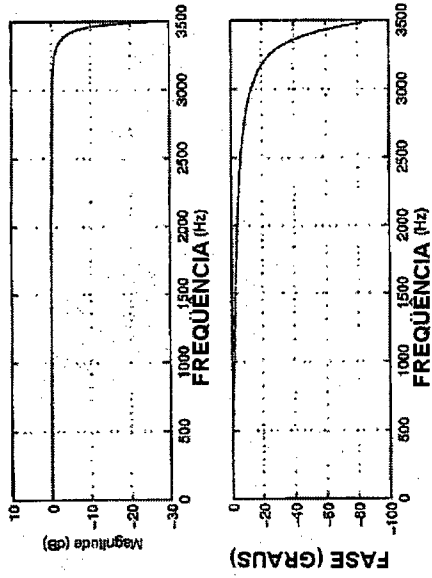


FIG. 36b

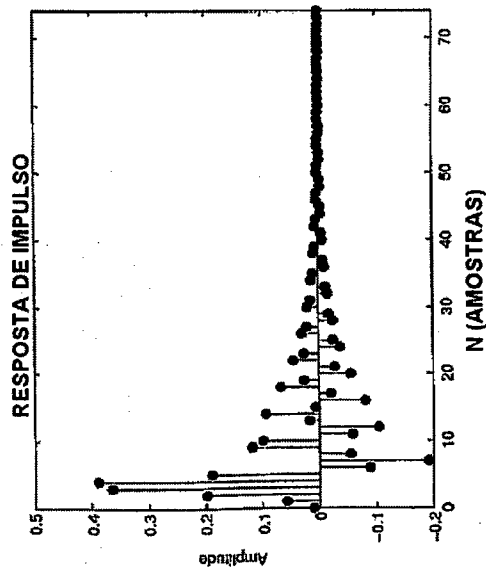
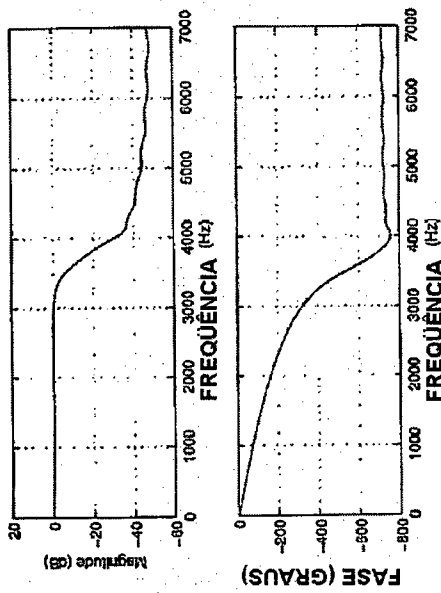


FIG. 36a

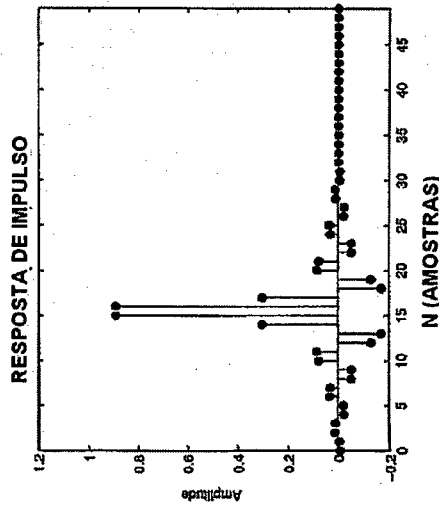
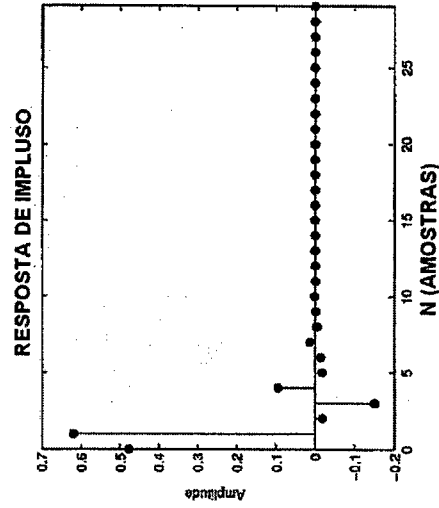
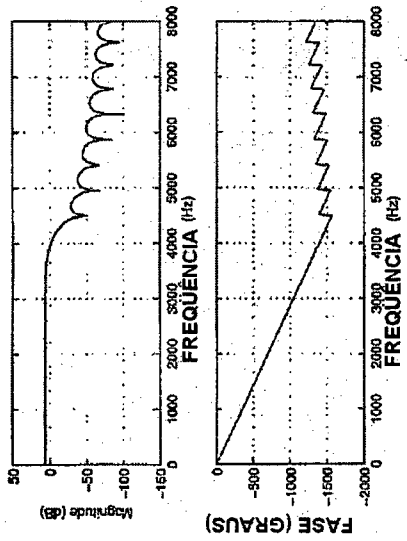
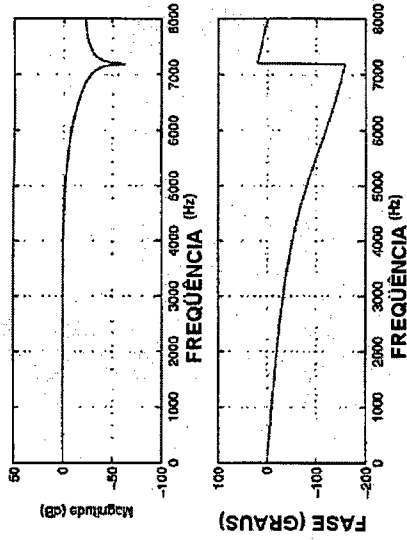


FIG. 37b

FIG. 37a

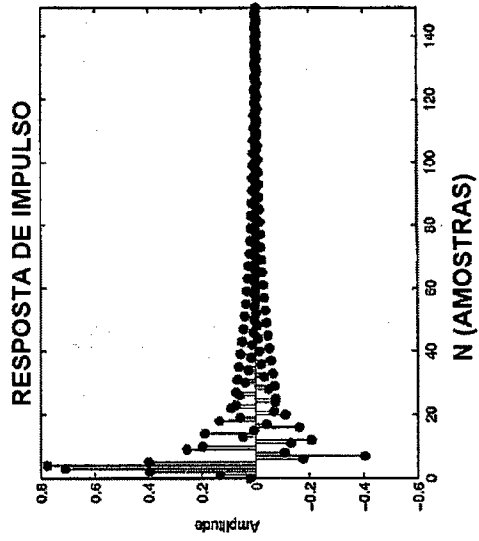
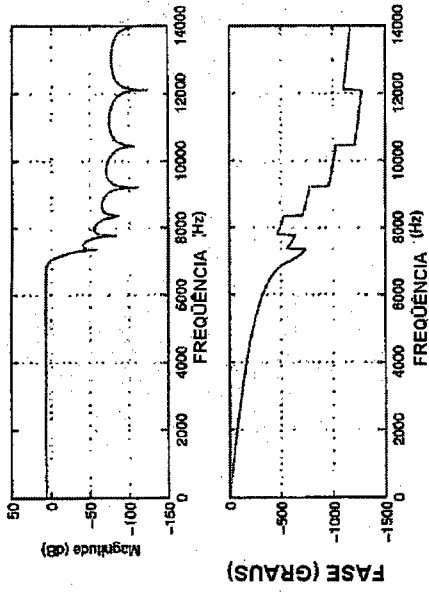
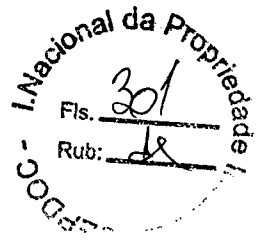


FIG. 38b

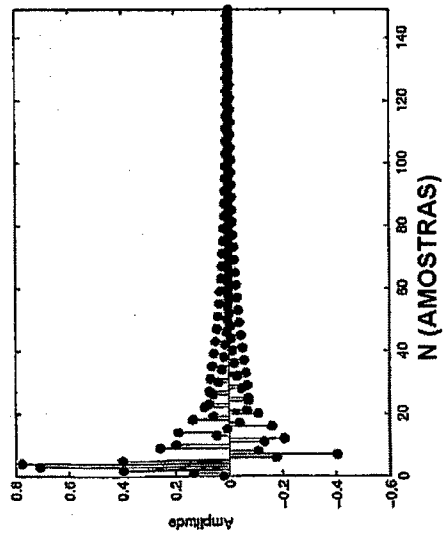
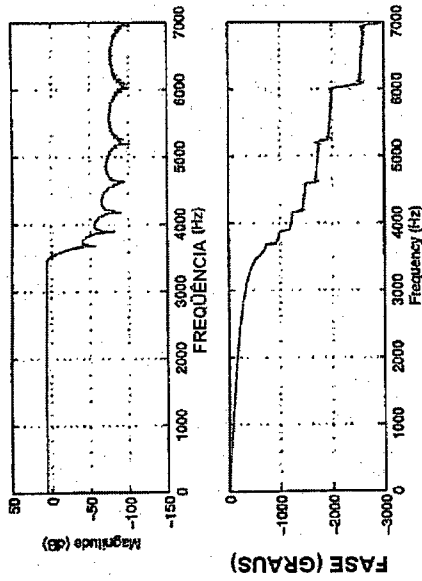


FIG. 38a

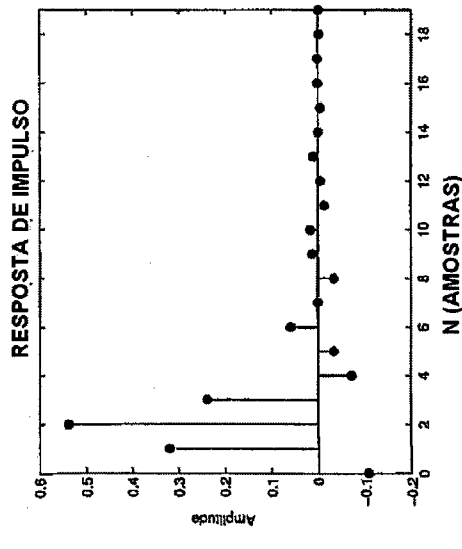
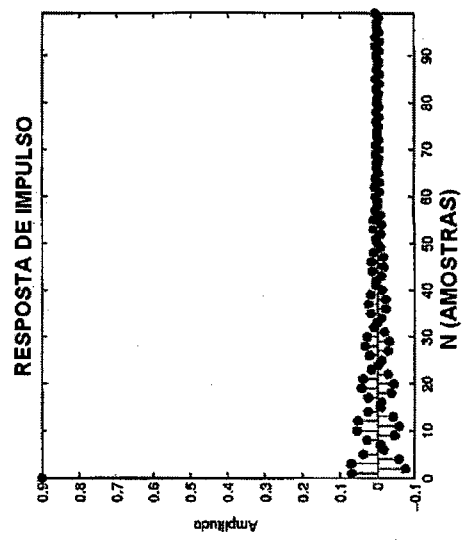
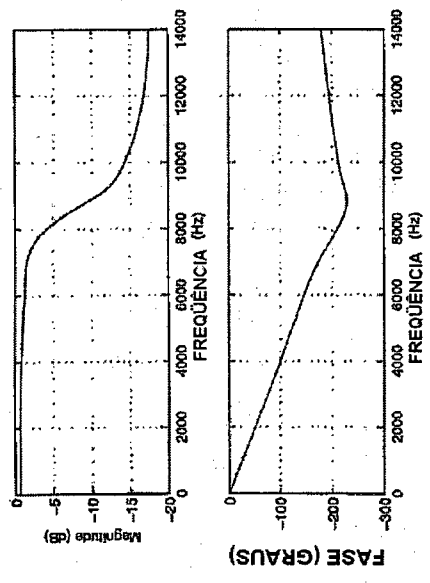
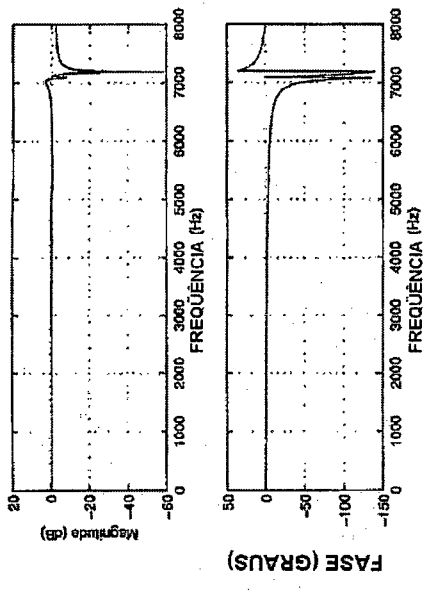


FIG. 39b

FIG. 39a