



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

(21) **PI0612564-6 A2**

(22) Data de Depósito: 27/06/2006
(43) Data da Publicação: 23/11/2010
(RPI 2081)



(51) *Int.Cl.:*
G10L 19/12

(54) Título: **MÉTODO PARA EXTENSÃO DE LARGURA DE BANDA PARA COMUNICAÇÕES E SISTEMA PARA ESTENDER ARTIFICIALMENTE A LARGURA DE BANDA DE VOZ**

(30) Prioridade Unionista: 30/06/2005 US 11/171.608

(73) Titular(es): MOTOROLA, INC.

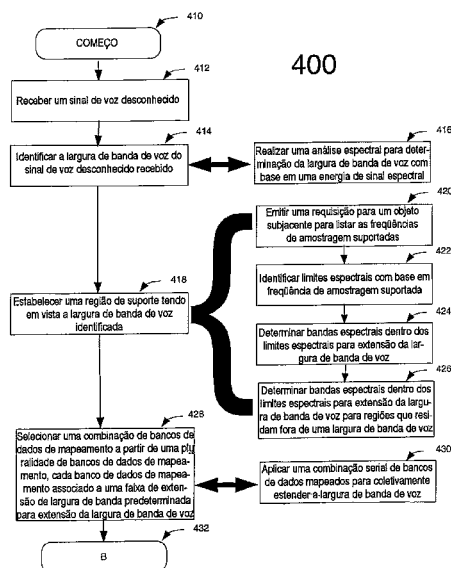
(72) Inventor(es): HARSHA M. SATHYENDRA, ISMAIL UYSAL,
JOHN G. HARRIS, MARC A. BOILLOT

(74) Procurador(es): ORLANDO DE SOUZA

(86) Pedido Internacional: PCT US2006025119 de 27/06/2006

(87) Publicação Internacional: WO 2007/005444 de 11/01/2007

(57) **Resumo:** METODO PARA EXTENSÃO DE LARGURA DE BANDA PARA COMUNICAÇÕES DE VOZ E SISTEMA PARA ESTENDER ARTIFICIALMENTE A LARGURA DE BANDA DE VOZ. A invenção concerne a um método (400) e um sistema (100) para extensão de largura de banda de voz para melhoria da qualidade de voz em um sistema de comunicação. O método pode incluir as etapas de recebimento (412) de um sinal de voz desconhecido (105), identificação (414) da largura de banda de voz (625) do sinal de voz desconhecido recebido e estabelecimento (418) de uma região de suporte (636) tendo em vista o conteúdo espectral do sinal de voz recebido. O método ainda pode incluir a etapa de seleção (428) de uma combinação de bancos de dados de mapeamento (210, 212, 214) a partir de uma pluralidade de bancos de dados de mapeamento. Cada banco de dados de mapeamento pode estar associado a uma faixa de extensão de largura de banda predeterminada para extensão da largura de banda de voz.



MÉTODO PARA EXTENSÃO DE LARGURA DE BANDA PARA COMUNICAÇÕES
DE VOZ E SISTEMA PARA ESTENDER ARTIFICIALMENTE A LARGURA DE
BANDA DE VOZ

ANTECEDENTES DA INVENÇÃO

5 1. Campo da Invenção

Esta invenção se refere em geral à extensão de largura de banda e, mais particularmente, à extensão de sinais de voz de banda estreita para sinais de voz de banda larga.

2. Descrição da Técnica Relacionada

10 O uso de dispositivos eletrônicos portáteis explodiu nos últimos anos. Telefones celulares, em particular, tornaram-se bastante populares com o público. A finalidade primária de telefones celulares é para comunicação de voz. Um telefone celular opera em sinais de voz pela compressão
15 de voz e pelo envio dos sinais de voz por uma rede de comunicações. A compressão reduz a quantidade de dados requeridos para representação do sinal de voz e da largura de banda de voz. Por exemplo, a largura de banda de voz em um telefone celular é uma banda geralmente limitada entre
20 300 Hz e 3,4 kHz, ao passo que a voz falada natural reside principalmente em uma largura de banda entre 20 Hz e 20 kHz. O processo de limitação de banda de voz é uma etapa necessária envolvida na transmissão e na recepção eficiente de sinais digitais em um sistema de comunicação celular.

25 Felizmente, uma voz comprimida preserva suficientemente o caráter e a inteligibilidade da voz natural, embora não inclua todas as componentes de frequência dos dados originais. Em particular, uma compressão de voz remove as regiões de frequência baixa de
30 voz (isto é, abaixo de 300 Hz), bem como as regiões de

freqüência alta de voz (isto é, acima de 3,4 kHz a 10 kHz). Embora uma compressão de voz produza um sinal de voz que é satisfatório para comunicações sem fio, várias técnicas de processamento de fala foram testadas e aplicadas em uma tentativa de restauração das componentes de voz de freqüência baixa e de freqüência alta faltando para a geração de um sinal de qualidade mais alta. Até a presente data, contudo, nenhuma técnica foi desenvolvida que efetivamente recrie as componentes de freqüência removidas. Mais ainda, telefones analógicos convencionais não implementam qualquer conversão. Como tal, eles ainda sofrem de restrições de largura de banda similares devido a padrões de transmissão de décadas de idade.

SUMÁRIO DA INVENÇÃO

A presente invenção concerne a um método para extensão de largura de banda para comunicações de voz. O método pode incluir as etapas de recebimento de um sinal de voz desconhecido, a identificação da largura de banda de voz do sinal de voz desconhecido recebido e o estabelecimento de uma região de suporte tendo em vista o conteúdo espectral do sinal de voz recebido. O método também pode incluir a etapa de seleção de uma combinação de bancos de dados de mapeamento a partir de uma pluralidade de bancos de dados de mapeamento. Cada banco de dados de mapeamento pode estar associado a uma faixa de extensão de largura de banda predeterminada para extensão da largura de banda.

Como um exemplo, a identificação da largura de banda de voz pode incluir a realização de uma análise espectral para a determinação da largura de banda de sinal de voz do sinal de voz desconhecido com base em uma energia espectral

do sinal. Também, o estabelecimento de uma região de suporte pode incluir as etapas de emissão de uma requisição para um objeto subjacente para retornar uma lista de frequências de amostragem para a qual o objeto é capaz de suportar, identificação de limites espectrais com base na frequência de amostragem retornada e a determinação de bandas espectrais nos limites espectrais para extensão da largura de banda de voz para regiões que residem fora da largura de banda de voz. O estabelecimento de uma região de suporte ainda pode incluir a etapa de reamostragem do sinal de voz a uma frequência de amostragem correspondente a pelo menos uma das frequências de amostragem retornadas.

Em um arranjo, a etapa de seleção de uma combinação de bancos de dados de mapeamento pode ser uma operação sequencial. Esta etapa de seleção ainda pode incluir a aplicação de uma combinação serial de bancos de dados mapeados para extensão coletivamente da largura de banda de voz para uma faixa correspondente à adição das faixas de extensão de largura de banda selecionada. Como um exemplo, pode haver um primeiro banco de dados de mapeamento para a faixa de aproximadamente 0 a aproximadamente 8 kHz, um segundo banco de dados de mapeamento para a faixa de aproximadamente 8 kHz a aproximadamente 16 kHz e um terceiro banco de dados de mapeamento para a faixa de aproximadamente 16 kHz a aproximadamente 22 kHz. Os três bancos de dados de mapeamento podem ser Modelos de Mistura Gaussianos.

O método também pode incluir as etapas de aquisição de um conjunto de coeficientes de reflexão de banda estreita que representam a envoltória espectral a partir do sinal de

voz e estendendo o conjunto de coeficientes de reflexão de banda estreita para um conjunto de coeficientes de reflexão de banda larga usando-se os bancos de dados de mapeamento para a geração de uma envoltória espectral de banda larga.

5 Além disso, um conjunto de coeficientes de reflexão pode ser convertido em um conjunto de coeficientes cepstrais para redução de um armazenamento em memória pela compressão de uma matriz de covariância cheia gaussiana em um vetor diagonal de variâncias.

10 Em um outro arranjo, o método ainda pode incluir as etapas de extração de um sinal de excitação de banda estreita a partir do sinal de voz, usando-se um conjunto de coeficientes de reflexão de banda larga e extensão do sinal de excitação de banda estreita para um sinal de excitação
15 banda larga usando-se modulação e filtração. O método ainda pode incluir as etapas de combinação de um sinal de excitação de banda larga com uma envoltória espectral de banda larga para a geração de um sinal de voz de banda larga sintético, extração de um sinal de voz de banda larga
20 suplementar a partir do sinal de voz de banda larga sintético na região de suporte e adição do sinal de voz de banda larga sintético suplementar ao sinal de voz original para a geração de um sinal de voz de banda larga.

A presente invenção também concerne a um método de
25 extensão de um conjunto de coeficientes de reflexão de banda estreita para um conjunto de coeficientes de banda larga para uso em extensão de largura de banda de voz. Este método pode incluir as etapas de geração de uma excitação de banda baixa, geração de uma excitação de banda alta e
30 adição da excitação de banda baixa e da excitação de banda

alta com uma excitação de banda estreita para a criação de uma excitação de meia banda. O método também pode incluir a etapa de geração de uma excitação de banda larga a partir da excitação de meia banda. A etapa de geração da excitação de banda baixa e da excitação de banda alta pode incluir as etapas de modulação da excitação de banda baixa e da excitação de banda alta usando-se uma multiplicação de cosseno e filtrando-se a excitação de banda baixa e a excitação de banda alta.

10 A presente invenção também concerne a um armazenamento que pode ser lido em máquina. O armazenamento que pode ser lido em máquina pode ter armazenado nele um programa de computador que tem uma pluralidade de seções de código executáveis por um dispositivo de computação portátil. As seções de código podem fazer com que o dispositivo de computação portátil realize as etapas de recebimento de um sinal de voz desconhecido, identificação da largura de banda de voz do sinal de voz desconhecido recebido e estabelecimento de uma região de suporte tendo em vista o conteúdo espectral do sinal de voz recebido. As seções de código ainda podem fazer com que o dispositivo de computação portátil realize a etapa de seleção de uma combinação de bancos de dados de mapeamento a partir de uma pluralidade de bancos de dados de mapeamento. Como antes, cada banco de dados de mapeamento pode estar associado a uma faixa de extensão de largura de banda predeterminada para extensão da largura de banda de voz. As seções de código também podem fazer com que o dispositivo de computação portátil realize qualquer uma das outras etapas de método recitadas acima.

15
20
25
30

A presente invenção também concerne a um sistema para a extensão artificial da largura de banda de voz. O sistema pode incluir uma seção de avaliação, um seletor de banco de dados cooperativamente acoplado à seção de avaliação e uma
5 unidade de extensão de largura de banda acoplada de forma cooperativa à seção de avaliação e ao seletor de banco de dados. A seção de avaliação pode receber um sinal de voz desconhecido e pode determinar uma extensão admissível de largura de banda de voz para o sinal de voz desconhecido. O
10 seletor de banco de dados pode escolher uma combinação de bancos de dados de mapeamento de acordo com a extensão admissível de largura de banda de voz. Além disso, a unidade de extensão de largura de banda pode estender a largura de banda de voz do sinal de voz desconhecido para a
15 extensão admissível de largura de banda de voz. A unidade de extensão de largura de banda pode fazer isto pelo uso da combinação de bancos de dados de mapeamento escolhidos pelo seletor de banco de dados. O sistema também pode incluir um circuito adequado e um software para a realização de
20 qualquer uma das etapas de método recitadas acima.

BREVE DESCRIÇÃO DOS DESENHOS

Os recursos da presente invenção, os quais se acredita que sejam novos, são estabelecidos com particularidade nas reivindicações em apenso. A invenção, em conjunto com
25 outros objetivos e vantagens da mesma, pode ser mais bem compreendida por uma referência à descrição a seguir, tomada em conjunto com os desenhos associados, nas várias figuras das quais os números de referência iguais identificam elementos iguais e nas quais:

30 a FIG. 1 ilustra um sistema para a extensão artificial

da largura de banda de voz de acordo com uma modalidade dos arranjos inventivos;

a FIG. 2 ilustra alguns dos componentes da FIG. 1 em maiores detalhes de acordo com uma modalidade dos arranjos
5 inventivos;

a FIG. 3 ilustra um exemplo de um estágio de excitação de percurso múltiplo de acordo com uma modalidade dos arranjos inventivos;

a FIG. 4 ilustra uma porção de um método para extensão
10 de largura de banda de voz de acordo com uma modalidade dos arranjos inventivos;

a FIG. 5 ilustra uma outra porção de um método para extensão de largura de banda de voz de acordo com uma modalidade dos arranjos inventivos;

15 a FIG. 6 ilustra vários gráficos associados à extensão de largura de banda de um sinal de voz de acordo com uma modalidade dos arranjos inventivos; e

a FIG. 7 ilustra um sistema para conversão de um conjunto de coeficientes de banda estreita para um conjunto
20 de coeficientes de banda larga de acordo com uma modalidade dos arranjos inventivos.

DESCRIÇÃO DETALHADA DA INVENÇÃO

Embora o relatório descritivo seja concluído com reivindicações definindo os recursos da invenção que são
25 considerados novidade, acredita-se que a invenção será mais bem compreendida a partir de uma consideração da descrição a seguir em conjunto com os desenhos, nos quais os números de referência são transportados.

Conforme requerido, as modalidades detalhadas da
30 presente invenção são mostradas aqui; contudo, é para ser

compreendido que as modalidades mostradas são meramente de exemplo da invenção, a qual pode ser concretizada de várias formas. Portanto, detalhes estruturais específicos e funcionais mostrados aqui não devem ser interpretados como
5 limitantes, mas meramente como uma base para as reivindicações e como uma base representativa para se ensinar alguém versado na técnica a empregar variadamente a presente invenção em virtualmente qualquer estrutura apropriadamente detalhada. Ainda, não se pretende que os
10 termos e as frases usados aqui sejam limitativos, mas, ao invés disso, provejam uma descrição compreensível da invenção.

Os termos "um" ou "uma", conforme usado aqui, são definidos como um ou mais de um. O termo "pluralidade",
15 conforme usado aqui, é definido como dois ou mais de dois. O termo "um outro", conforme usado aqui, é definido como pelo menos um segundo ou mais. Os termos "incluindo" e/ou "tendo", conforme usados aqui, são definidos como compreendendo (isto é, de linguagem aberta). O termo
20 "acoplado", conforme usado aqui, é definido como conectado, embora não necessariamente de forma direta, e não necessariamente de forma mecânica. Os termos "programa", "aplicativo de software" e similares, conforme usados aqui, são definidos como uma sequência de instruções designadas
25 para execução em um sistema de computador. Um programa, um programa de computador ou um aplicativo de software pode incluir uma sub-rotina, uma função, um procedimento, um método de objeto, uma implementação de objeto, um aplicativo executável, um miniaplicativo, um miniaplicativo
30 de servidor, um código fonte, um código de objeto, uma

biblioteca compartilhada/biblioteca de carga dinâmica e/ou uma outra seqüência de instruções designadas para execução em um sistema de computador.

Um objetivo de extensão de largura de banda de voz é restaurar a qualidade de voz comprimida para um nível que combinem com o nível de qualidade subjetiva da voz original. A invenção concerne a um método e um sistema para extensão de largura de banda de voz para melhoria da qualidade de voz em um sistema de comunicação. O método pode incluir as etapas de recebimento de um sinal de voz desconhecido, identificação da largura de banda de voz do sinal de voz desconhecido recebido e estabelecimento de uma região de suporte tendo em vista o conteúdo espectral do sinal de voz recebido. O método também pode incluir a etapa de seleção de uma combinação de bancos de dados de mapeamento a partir de uma pluralidade de bancos de dados de mapeamento, em que cada banco de dados de mapeamento pode estar associado a uma faixa de extensão de largura de banda predeterminada para extensão da largura de banda de voz para a região de suporte. Através destas etapas e outros processos que serão descritos abaixo, a largura de banda do sinal de voz desconhecido pode ser estendida.

Com referência à FIG. 1, um exemplo de um sistema para extensão artificial da largura de banda de voz é mostrado. Em um arranjo, o sistema 100 pode incluir uma seção de avaliação 110, um seletor de banco de dados 120, o qual pode ser cooperativamente acoplado à seção de avaliação 110, e uma unidade de extensão de largura de banda 130. A unidade de extensão de largura de banda 130 pode ser acoplada de forma cooperativa à seção de avaliação

110 e ao seletor de banco de dados 120. Em uma modalidade, a seção de avaliação 110, o seletor de banco de dados 120 e a unidade de extensão de largura de banda 130 podem ser parte de uma unidade de comunicações móveis 140, como um
5 telefone celular. Em um caso como esse, a unidade de comunicações móveis 140 pode incluir um receptor 150 e/ou um transmissor 160 para recepção e/ou transmissão de sinais de voz ou de dados.

A seção de avaliação 110 pode receber um sinal de voz
10 desconhecido 105 e pode determinar uma extensão admissível de largura de banda de voz para o sinal de voz desconhecido 105. Este sinal de voz desconhecido 105, tendo em vista um processamento subsequente realizado nele, também pode ser referido simplesmente como o sinal de voz 105 ou o sinal de
15 voz reamostrado 105. A extensão admissível de largura de banda de voz pode corresponder a uma região de suporte. Como um exemplo, o seletor de banco de dados 120 pode escolher uma combinação de bancos de dados de mapeamento (não mostrada aqui) de acordo com a extensão admissível de
20 largura de banda de voz. Também, a unidade de extensão de largura de banda 130 pode estender a largura de banda de voz do sinal de voz desconhecido 105 para a extensão admissível de largura de banda de voz. Por exemplo, a unidade de extensão de largura de banda 130 pode estender a
25 largura de banda de voz do sinal de voz desconhecido 105 usando a combinação de bancos de dados de mapeamento escolhida pelo seletor de banco de dados 120.

Com referência à FIG. 2, um diagrama de blocos mais detalhado da seção de avaliação 110, do seletor de banco de
30 dados 120 e da unidade de extensão de largura de banda 130

é mostrado. Em um arranjo, a seção de avaliação 110 pode incluir um módulo de análise 202, um módulo de investigação 204 e um módulo de amostragem 206. O módulo de análise 202 pode ser acoplado ao módulo de investigação 204, o qual
5 pode ser acoplado ao módulo de amostragem 206. Adicionalmente, o módulo de amostragem 206 pode ser acoplado ao módulo de análise 202.

Brevemente, o módulo de análise 202 é capaz de identificar a largura de banda de voz do sinal de voz
10 desconhecido recebido 105. O módulo de investigação 204 é capaz de identificar uma lista de taxas de amostragem suportadas associadas ao sistema 100, onde cada taxa de amostragem suportada pode revelar a extensão até a qual a largura de banda de voz pode ser estendida. Como um
15 exemplo, as taxas de amostragem suportadas podem estar associadas à unidade móvel 140. O módulo de amostragem 206 pode reamostrar o sinal de voz desconhecido 105 a uma taxa de amostragem identificada pelo módulo de investigação 204, o que pode produzir um sinal de voz reamostrado 105. Assim,
20 a seção de avaliação 110 pode efetivamente 1) analisar o sinal de voz desconhecido 105 para determinar a largura de banda de voz; 2) identificar as taxas de amostragem que o sistema 100 pode suportar; 3) determinar uma extensão admissível de largura de banda de voz; e 4) reamostrar o
25 sinal de voz 105 a uma das taxas de amostragem identificadas.

Em um arranjo, o seletor de banco de dados 120 pode incluir uma pluralidade de bancos de dados de mapeamento 210, 212 e 214, na qual cada banco de dados de mapeamento
30 210, 212 e 214 pode estar associado a uma faixa de extensão

de largura de banda predeterminada para extensão da largura de banda de voz. O seletor de banco de dados 120 pode escolher os bancos de dados de mapeamento 210, 212 e 214 para seletivamente estenderem a largura de banda do sinal de voz 105 até a largura de banda suportada pelo sistema. Em particular, os bancos de dados de mapeamento 210, 212 e 214 podem prover capacidades incrementais para extensão da largura de banda de voz com base nas frequências de amostragem suportadas pelo sistema. Este processo será explicado em maiores detalhes abaixo.

Em um arranjo, a unidade de extensão de largura de banda 130 pode incluir um processador de envoltória 220, um processador de excitação 240 e um processador de mistura 260. O processador de envoltória 220 pode ser acoplado de forma comunicativa à seção de avaliação 110 e ao seletor de banco de dados 120. O processador de excitação 240 pode ser acoplado de forma comunicativa à seção de avaliação 110 e ao processador de envoltória 220. Além disso, o processador de mistura 260 pode ser acoplado de forma comunicativa à seção de avaliação 110, ao processador de envoltória 220 e ao processador de excitação 240.

Brevemente, o processador de envoltória 220 pode determinar uma envoltória de banda estreita a partir do sinal de voz 105 e subsequente uma envoltória espectral de banda larga. Como um exemplo, e sem limitação, o processador de envoltória 220 pode prover um conjunto de coeficientes de banda larga representando uma envoltória espectral de banda larga. Usando a envoltória espectral de banda larga (por exemplo, o conjunto de coeficientes de banda larga) provida pelo processador de envoltória 220, o

processador de excitação 240 pode determinar um sinal de excitação de banda estreita a partir do sinal de voz 105 para subseqüentemente criar um sinal de excitação de banda larga. O processador de mistura 260 pode criar um sinal de
5 banda larga suplementar a partir do sinal de excitação de banda larga e da envoltória espectral de banda larga, o qual então pode ser combinado com o sinal de voz 105 para a criação de um sinal de voz de banda larga.

Como um exemplo, o processador de envoltória 220 pode
10 incluir um extrator de recurso 222, um conversor de banda estreita 223, um estimador de envoltória 224 e um conversor de banda larga 225. O extrator de recurso 222 pode ser acoplado de forma comunicativa ao módulo de amostragem 206 para o recebimento do sinal de voz reamostrado 105 e para a
15 aquisição de um conjunto de coeficientes de análise de predição linear (LPC) representando uma envoltória espectral de banda estreita do sinal de voz reamostrado 105. Ainda, o conversor de banda estreita 223, o qual pode ser acoplado de forma comunicativa ao extrator de recurso
20 222, pode converter o conjunto de coeficientes de LPC em um conjunto de coeficientes de reflexão de banda estreita.

O estimador de envoltória 224 pode ser acoplado de forma comunicativa ao conversor de banda estreita 223 e pode receber o conjunto de coeficientes de reflexão de
25 banda estreita representando a envoltória espectral de banda estreita. Usando os bancos de dados de mapeamento 210, 212 e 214, o estimador de envoltória 224, em conjunto com o seletor de banco de dados 120, pode estender o conjunto de coeficientes de reflexão de banda estreita para
30 um conjunto de coeficientes de reflexão de banda larga, o

que pode permitir que o estimador de envoltória 224 (e o seletor de banco de dados 120) estime uma envoltória espectral de banda larga a partir de uma envoltória espectral de banda estreita. Acoplado de forma comunicativa
5 ao estimador de envoltória 224, um conversor de banda larga 225 pode converter os coeficientes de reflexão de banda larga em um conjunto de coeficientes de LPC de banda larga.

O processador de excitação 240 pode incluir uma seção de análise de banda larga 242 e um estágio de excitação de percurso múltiplo 244, ambos os quais podendo ser acoplados
10 de forma comunicativa um ao outro. A seção de análise de banda larga 242 pode ser acoplada ao módulo de amostragem 206 para o recebimento do sinal de voz reamostrado 105. Uma vez recebido, a seção de análise de banda larga 242 pode
15 extrair um sinal de excitação de banda estreita a partir do sinal de voz reamostrado 105 usando a envoltória espectral de banda larga produzida pelo estimador de envoltória 224. Conforme será discutido mais tarde, uma outra abordagem é
20 de um sinal de excitação de banda estreita a partir do sinal de voz reamostrado 105. O estágio de excitação de percurso múltiplo 244 pode gerar um sinal de excitação de banda larga a partir do sinal de excitação de banda estreita extraído pela seção de análise de banda larga 242.

25 O processador de mistura 260 pode incluir uma seção de síntese de banda larga 262, um filtro de pára banda 264 e um adicionador 266. A seção de síntese de banda larga 262 pode combinar o sinal de excitação de banda larga provido pelo processador de excitação 240 em conjunto com a
30 envoltória de banda larga provida pelo processador de

envoltória 220 para a geração de um sinal de voz de banda larga sintético. O filtro de pára banda 264 pode suprimir o conteúdo espectral do sinal de voz de banda larga sintético dentro das regiões de freqüência já ocupadas pelo sinal de voz 105. Como resultado, o filtro de pára banda 264 pode prover um sinal de voz de banda larga suplementar que inclui uma informação de freqüência dentro da extensão admissível de largura de banda de voz. O adicionador 266 pode combinar o sinal de banda larga suplementar recebido a partir do filtro de pára banda 264 com o sinal de voz a partir do módulo de amostragem 206, para a criação de um sinal de voz de banda larga.

Embora as FIG. 1 e 2 representem exemplos de sistemas e componentes (de hardware e software) que permitiriam que se praticasse o método inventivo, é entendido que a invenção não está limitada dessa forma. O método pode ser praticado em qualquer sistema de processamento de voz adequado usando qualquer combinação adequada de componentes, de software e hardware.

Com referência à FIG. 3, um exemplo de um diagrama de blocos mais detalhado do estágio de excitação de percurso múltiplo 244 é mostrado. É compreendido, contudo, que esta representação em particular do estágio de excitação de percurso múltiplo 244 é meramente um exemplo de um componente como esse. Aqueles de conhecimento na técnica apreciarão que outros layouts adequados podem ser empregados na invenção.

Em um arranjo, o estágio de excitação de percurso múltiplo 244 pode incluir um estágio de excitação de banda baixa 310, um estágio de excitação de banda alta 320 e um

estágio de excitação de passa banda 330, cuja combinação é capaz de processar o sinal de excitação de banda estreita recebido a partir da seção de análise de banda larga 242 (veja a FIG. 2).

5 O estágio de excitação de banda baixa 310 pode incluir um modulador 312 e um filtro de passa baixa 314. O estágio de excitação de banda alta 320 pode incluir um modulador 322 e um filtro de passa banda 324. O estágio de excitação de passa banda 330 pode passar o sinal de excitação de
10 banda estreita não processado. Uma finalidade do estágio de excitação de banda baixa 310, do estágio de excitação de banda alta 320 e do estágio de excitação de passa banda 330 é estender artificialmente o sinal de excitação para uma faixa de frequência identificada pelo módulo de
15 investigação 204.

O estágio de excitação de percurso múltiplo 244 também pode incluir um adicionador 340 para somar os sinais expirados de banda baixa, de banda alta e de passa banda em um sinal de excitação de meia banda compósito. O estágio de
20 excitação de percurso múltiplo 244 também pode ter um modulador 350 para estender artificialmente a excitação de meia banda para uma excitação de banda larga, a qual pode ser considerada uma excitação de banda plena ou de banda larga. Conforme citado anteriormente, o sinal de excitação
25 de banda larga gerado pelo estágio de excitação de percurso múltiplo 244 pode ser combinado com uma envoltória de banda larga para a geração de um sinal de voz de banda larga sintético.

Com referência às FIG. 4 a 5, um método 400 será usado
30 para explicação da extensão da largura de banda de voz.

Embora as FIG. 1 a 3 sejam usadas para ajudarem na descrição do método 400, deve ser compreendido que o método 400 pode ser implementado em qualquer outro dispositivo ou sistema adequado usando-se quaisquer componentes adequados.

5 Mais ainda, a invenção não está limitada à ordem na qual as etapas são listadas no método 400. Além disso, o método 400 pode conter um número maior ou menor de etapas do que aquelas mostradas nas FIG. 4 a 5.

Na etapa 410, o método 400 pode começar. Na etapa 412,
10 um sinal de voz desconhecido pode ser recebido. O termo "desconhecido" neste contexto pode significar que a taxa de amostragem ou largura de banda do sinal de voz recebido é desconhecida. Na etapa 414, a largura de banda de voz do sinal de voz desconhecido recebido pode ser identificada.
15 Como um exemplo, na etapa 416, uma análise espectral pode ser realizada no sinal de voz desconhecido para a determinação de uma largura de banda de sinal de voz com base na energia espectral.

Por exemplo, com referência à FIG. 2, o módulo de
20 análise 202 pode receber o sinal de voz desconhecido 105 e pode determinar a largura de banda de voz desconhecida, de acordo com as etapas 412 e 414. Aqueles de conhecimento na técnica apreciarão que há muitas formas diferentes de determinação da largura de banda de um sinal de voz, e a
25 invenção não está limitada a qualquer técnica em particular. Com referência à FIG. 6, um exemplo de uma resposta de frequência 620 do sinal de voz desconhecido é mostrado. O módulo de análise 202 da FIG. 2 pode gerar a resposta de frequência 620 e pode identificar a largura de
30 banda de voz com base na distribuição de energia espectral.

Por exemplo, uma largura de banda de voz 625 da resposta de frequência 620 pode ocupar uma região entre aproximadamente 300 Hz e aproximadamente 3,4 kHz, embora outros valores adequados possam ser facilmente substituídos na invenção.

- 5 Esta largura de banda de voz pode representar a largura de banda de pós-compressão do sinal de voz 105 (isto é, um sinal de voz de banda estreita).

O sinal de voz 105 aqui pode ter uma taxa de amostragem de 8 kHz, o que significa que o conteúdo
10 espectral não estará presente de 4 kHz a 8 kHz, tendo em vista o teorema de Nyquist. Embora não restrito pelo teorema de Nyquist, o conteúdo espectral pode não estar presente de 0 Hz a 300 Hz ou de 3,4 kHz a 4 kHz para o sinal de voz 105, o que é comum em muitos sistemas de
15 comunicações sem fio.

Com referência de volta ao método 400 das FIG. 4 e 5, na etapa 418, uma região de suporte tendo em vista a largura de banda de voz pode ser estabelecida. Como um exemplo, a região de suporte pode descrever regiões de
20 frequência de fala em que um conteúdo espectral pode estar ausente e em que uma extensão de largura de banda de voz pode ser aplicada. As etapas 420 a 426 descrevem um exemplo de como uma região de suporte pode ser estabelecida. Em particular, na etapa 420, uma requisição pode ser emitida
25 para um objeto subjacente para listar frequências de amostragem que o objeto é capaz de suportar. Um conhecimento das frequências de amostragem, conforme determinado acima, pode ser requerido, porque as taxas de amostragem revelam a extensão até a qual a largura de banda
30 de voz pode ser estendida. Limites espectrais baseados nas

taxas de amostragem suportadas podem ser identificados, conforme mostrado na etapa 422. Os limites espectrais podem definir as delimitações de frequência em que o sistema pode adicionar um conteúdo espectral ao sinal de voz.

5 Na etapa 424, as bandas espectrais podem ser determinadas nos limites espectrais para extensão de largura de banda de voz para regiões que residam fora da largura de banda de voz do sinal de voz. Na etapa 426, o sinal de voz pode ser reamostrado a uma taxa de amostragem
10 selecionada correspondente a pelo menos uma das frequências de amostragem retornadas. Este processo pode preparar a faixa de frequência para extensão do conteúdo espectral dentro do sinal de voz de banda estreita.

 Por exemplo, com referência às FIG. 2 e 6, o módulo de
15 investigação 204 pode emitir uma requisição para um objeto subjacente para listar as frequências de amostragem suportadas. O objeto subjacente pode ser um dispositivo físico ou uma interface de software que proveja uma capacidade de realizar um processamento de sinal, e pode
20 estar ciente das taxas de amostragem que pode suportar. Por exemplo, um dispositivo tocador de áudio pode prover numerosas taxas de amostragem, tais como 8 kHz para voz, 22,5 kHz para MP3 e 44,1 kHz para um disco compacto. Conforme é conhecido na técnica, a largura de banda do
25 sistema então pode ser determinada a partir da frequência de amostragem, usando-se o critério de Nyquist. Como tal, uma frequência de amostragem de 8 kHz pode prover uma largura de banda de voz de metade da frequência de amostragem, a qual é de 4 kHz.

30 Dado um conhecimento da largura de banda de voz do

5 sinal de voz desconhecido 105 e da largura de banda de sistema disponível, a seção de avaliação 110 pode determinar regiões em que um conteúdo espectral está ausente no sinal de voz 105. Especificamente, a seção de avaliação 110 pode definir limites espectrais das delimitações de frequência em que um conteúdo espectral pode ser adicionado ao sinal de voz 105, de acordo com a etapa 422 do método 400. Por exemplo, os limites espectrais para a resposta de frequência 625 do sinal de voz desconhecido 105 são demarcados pelos limites 623 e 627. Neste exemplo, isto corresponde aos limites espectrais inferiores de 0 a 300 Hz (limite 623) e limites espectrais mais altos de 3,4 kHz a 8 kHz (limite 627).

15 A unidade de avaliação 110 também pode determinar bandas espectrais nos limites espectrais identificados para a determinação da extensão de largura de banda de voz com base na largura de banda de sistema, de acordo com a etapa 424. Em um arranjo, as bandas espectrais podem definir uma região de suporte 636. A região de suporte 636 pode descrever as regiões de frequência em que um conteúdo espectral pode ser adicionado à largura de banda de voz, para o que há atualmente pouco ou nenhum conteúdo de frequência de voz. Como tal, a região de suporte 636 inerentemente descreve a extensão admissível de largura de banda de voz.

25 Por exemplo, o módulo de análise 202 pode realizar uma análise espectral do sinal de voz desconhecido 105, o que pode revelar que a largura de banda de voz está entre 300 Hz e 3,4 kHz, conforme visto na largura de banda de voz 625. Conforme é conhecido na técnica, o teorema de Nyquist

estabelece que a taxa de amostragem associada ao sinal de voz desconhecido deve ser de pelo menos duas vezes a largura de banda de sinal, o que é uma taxa de amostragem de 8 kHz em nosso exemplo. Uma investigação do objeto subjacente pode revelar que as taxas de amostragem de 8 kHz, 16 kHz, 22 kHz e 44 kHz são suportadas. Como um exemplo, a uma taxa de amostragem de 8 kHz, nem toda a região superior de suporte (4 kHz ou 8 kHz) pode estar disponível (embora possa haver uma região inferior de suporte (0 Hz ou 300 Hz) e parte de uma região superior de suporte (3,4 kHz a 4 kHz)).

Caso o módulo de investigação 204 identifique uma frequência de amostragem mais alta suportada de 16 kHz, contudo, uma região superior de suporte é possível. Uma taxa de amostragem suportada pelo sistema de 16 kHz sugere que pelo menos uma porção da região superior admissível de suporte 637 é de 4 kHz, ou a largura de banda de sinal para uma frequência de amostragem de 16 kHz menos o limite de banda estreita superior da largura de banda de voz (8 kHz menos 4 kHz). Neste exemplo, a amostragem do sinal de voz a 16 kHz pode permitir a adição de conteúdo espectral superior na região superior de suporte 637 entre 4 kHz e 8 kHz. Este conteúdo espectral superior adicional pode complementar um conteúdo espectral inferior que pode ser adicionado a uma região inferior de suporte 633 de 0 a 300 Hz e o conteúdo espectral na região superior de suporte 637 de 3,4 kHz a 4 kHz.

Neste exemplo, a região de suporte 636 pode incluir a região superior de suporte 637 e a região inferior de suporte 633. Aqueles de conhecimento comum na técnica

apreciarão, contudo, que a invenção não está limitada a este exemplo. Em particular, a região de suporte 636 pode não incluir uma região de suporte superior e inferior. Além disso, a região de suporte 636 não necessariamente tem que
5 cobrir a extensão plena dos limites espectrais identificados.

Conforme citado anteriormente, o módulo de amostragem 206 pode reamostrar o sinal de voz 105. A seção de avaliação 110 pode selecionar a taxa de reamostragem que
10 corresponde a uma das taxas de amostragem suportadas pelo sistema identificadas. Em um arranjo, a seção de avaliação 110 pode prover uma seleção automática ou manual. Em uma configuração de seleção manual, o usuário usando o sistema 100 pode selecionar a taxa de amostragem de sua escolha
15 através, por exemplo, de uma interface gráfica de usuário ou de qualquer outra interface adequada. Por exemplo, o usuário pode querer uma fala de alta qualidade e pode escolher a taxa de amostragem disponível mais alta. Alternativamente, na configuração de seleção automática, um
20 provedor de sistema, tal como uma concessionária sem fio, pode controlar a taxa de amostragem. Por exemplo, o provedor de sistema pode querer limitar a taxa de amostragem com base em uma medição de qualidade de serviço ou em uma estrutura de custo, onde o provedor de sistema
25 pode cobrar do usuário um honorário de serviço mais alto para uma fala de qualidade mais alta.

A reamostragem pelo módulo de amostragem 206, com efeito, estabelece a largura de banda de sistema disponível e prepara o sinal de voz 105 para uma extensão de largura
30 de banda. A reamostragem efetivamente permite a extensão da

largura de banda de voz para a região de suporte 636. Em
resumo, se a frequência de amostragem suportada pelo
sistema for mais alta do que a frequência de amostragem de
voz desconhecida, então, a largura de banda de sinal
5 ocupada pela voz desconhecida poderá ser considerada uma
banda estreita. Se o sinal de banda estreita puder ser
estendido em qualquer região até uma largura de banda
suportada pelo sistema, o sinal será considerado um sinal
de banda larga. A diferença no conteúdo de frequência entre
10 um sinal de banda estreita e um sinal de banda larga pode
ser a região de suporte. É entendido, contudo, que a
invenção não está limitada de forma alguma a qualquer um
dos exemplos recitados acima com respeito a uma banda
estreita ou a sinais de banda larga ou a uma região de
15 suporte.

Com referência de volta à FIG. 4, na etapa 428, uma
combinação de bancos de dados de mapeamento pode ser
selecionada a partir de uma pluralidade de bancos de dados
de mapeamento nos quais cada banco de dados de mapeamento
20 pode estar associado a uma faixa de extensão de largura de
banda predeterminada para extensão da largura de banda de
voz. Esta seleção pode ser considerada tendo em vista a
região de suporte. Conforme explicado anteriormente, a
região de suporte pode refletir a extensão admissível até a
25 qual a largura de banda de voz pode ser estendida. A
combinação de bancos de dados de mapeamento pode ser
selecionada para efetivamente adicionar conteúdo espectral
à região de suporte.

Os bancos de dados de mapeamento podem ser criados de
30 modo que um primeiro banco de dados de mapeamento possa

prover uma primeira faixa, um segundo banco de dados de mapeamento possa prover uma segunda faixa começando a partir do fim da primeira faixa, e um terceiro banco de dados de mapeamento possa prover uma terceira faixa começando a partir do fim da segunda faixa. Desta maneira, na etapa 430, os bancos de dados podem ser combinados serialmente para estenderem coletivamente a largura de banda de voz para a provisão de conteúdo espectral dentro da região de suporte.

10 Para ilustração, com referência às FIG. 2 e 6 e conforme explicado anteriormente, uma análise espectral pode revelar que a largura de banda de voz para um sinal a uma frequência de amostragem de 8 kHz está entre 500 e 3,4 kHz (veja a largura de banda de voz 625). As frequências
15 entre 4 kHz e 8 kHz são frequências em que a voz não pode estar presente devido ao teorema de amostragem de Nyquist. Daí, a largura de banda de voz, tendo em vista a frequência de amostragem de 8 kHz, pode ser estendida apenas para as frequências mais baixas, de 0 Hz a 300 Hz, e uma porção das
20 frequências superiores, de 3,4 kHz a 4 kHz. Se o sinal de voz 105 for reamostrado a uma taxa mais alta de 16 kHz, por exemplo, a largura de banda de voz pode ser estendida a partir de 4 kHz a 8 kHz. Em nosso exemplo, a região hachurada 639 denota uma região (de 8 kHz a 16 kHz) em que
25 a voz pode não estar presente devido ao teorema de amostragem de Nyquist, com base em uma taxa de amostragem de 16 kHz.

Um ou mais dos bancos de dados de mapeamento 210, 212 e 214 podem ser selecionados para preencherem a região
30 inferior de suporte 633 e a região superior de suporte 637.

Por exemplo, o primeiro banco de dados de mapeamento 210 pode permitir a extensão de largura de banda até 8 kHz, o que pode ser suficiente para uma voz amostrada a 16 kHz. Como um outro exemplo, para uma taxa de amostragem de 22 kHz, o banco de dados de mapeamento 210 e o banco de dados de mapeamento 212 podem ser combinados para a obtenção de uma extensão de banda e voz para até 11 kHz, o que pode ajudar a preencher uma porção da região hachurada 639. Isto é, o banco de dados de mapeamento 210 pode ser selecionado para ajudar na provisão de conteúdo espectral a partir de 0 Hz a 300 Hz e de 3,4 kHz a 8 kHz, enquanto o banco de dados de mapeamento 212 pode ajudar a preencher a faixa de 8 kHz a 11 kHz para uma frequência de amostragem de 22 kHz. Tendo em vista a taxa de amostragem mais alta de 22 kHz, uma porção da região hachurada 639 agora pode ser parte da região de suporte 636. Conforme se pode ver, a seleção de uma combinação de bancos de dados de mapeamento pode ser uma operação seqüencial, embora a invenção não esteja necessariamente limitada a um arranjo como esse.

Em um arranjo, o primeiro banco de dados de mapeamento 210 pode estar associado a uma faixa de extensão de largura de banda predeterminada de aproximadamente 0 Hz a aproximadamente 8 kHz, e o segundo banco de dados de mapeamento 212 pode estar associado a uma faixa de extensão de largura de banda predeterminada de aproximadamente 8 kHz a aproximadamente 16 kHz. Adicionalmente, o terceiro banco de dados de mapeamento 214 pode estar associado a uma faixa de extensão de largura de banda predeterminada de aproximadamente 16 kHz a aproximadamente 22 kHz.

Obviamente, aqueles de conhecimento na técnica

apreciarão que a invenção não está limitada a estes bancos de dados de mapeamento 210, 212 e 214. A invenção pode incluir qualquer número adequado de bancos de dados de mapeamento que estejam associados a quaisquer frequências 5 adequadas. Também, a invenção não está limitada a bancos de dados de mapeamento com base em faixas de extensão de frequência estendidas linearmente. Por exemplo, os bancos de dados de mapeamento poderiam todos suportar a mesma faixa de frequência, mas prover vários graus de 10 amplificação e supressão através da faixa de frequência comum.

Com referência de volta à FIG. 4, o método 400 pode continuar para a FIG. 5 pela etapa 432. Na etapa 434, a extensão de largura de banda pode ser aplicada na região de 15 suporte. As etapas 436 a 456 provêem um exemplo de como este processo pode ser realizado.

Na etapa 436, uma envoltória espectral de banda larga pode ser criada a partir do sinal de voz. Em particular, a envoltória espectral de banda larga pode ser determinada 20 pela estimativa da envoltória espectral de banda estreita que pode ser adquirida através de uma extração de recurso. Por exemplo, na etapa 438, um conjunto de coeficientes de reflexão de banda estreita que representa a envoltória espectral de banda estreita pode ser adquirido a partir do 25 sinal de voz. Na etapa 440, o conjunto de coeficientes de reflexão de banda estreita pode ser estendido para um conjunto de coeficientes de reflexão de banda larga usando-se os bancos de dados de mapeamento.

Como um exemplo, com referência à FIG. 2, o extrator 30 de recurso 222 pode receber o sinal de voz reamostrado 105

e pode realizar uma análise de predição linear (LPC) de banda estreita. De acordo com os princípios bem conhecidos de LPC, o extrator de recurso 222 pode extrair uma envoltória a partir do sinal de voz reamostrado 105. Devido
5 ao fato de o sinal de voz reamostrado 105 ser de banda estreita, a envoltória, em geral, é de banda estreita. A envoltória de banda estreita pode ser representada por um conjunto de coeficientes de LPC que descreve uma aproximação de modelo todo de pólos ("all-pole") da
10 envoltória de voz de banda estreita.

O extrator de recurso 222 pode gerar um conjunto de coeficientes de LPC, denotado por $A(z)$. O conversor de banda estreita 223 pode converter o conjunto de coeficientes de LPC em um conjunto de coeficientes de
15 reflexão. Os coeficientes de reflexão podem ser úteis no método inventivo, porque eles podem ser mais adequados para implementação de filtros digitais. Os coeficientes de reflexão podem ser mais robustos quanto a ruído, em comparação com os coeficientes de LPC, também. Aqueles
20 versados na técnica apreciarão, contudo, que a invenção não está limitada dessa forma, já que uma transformação como essa pode não ser necessária e que outras representações de coeficiente podem ser empregadas. Em qualquer caso, o conjunto de coeficientes de reflexão de banda estreita pode
25 representar, de forma análoga, a envoltória espectral, embora em uma forma matemática diferente.

Além disso, os coeficientes de reflexão podem ser convertidos para um conjunto de coeficientes cepstrais, os quais também são robustos quanto a um ruído numérico. Os
30 coeficientes de reflexão são estatisticamente dependentes

uns dos outros, significando que uma informação mútua está contida nos coeficientes individuais do conjunto de coeficientes de reflexão. Inversamente, os coeficientes cepstrais são estatisticamente independentes uns dos outros com uma informação mútua mínima entre os coeficientes. Esta independência é um atributo importante para fins de armazenamento em memória e pode ser relevante com respeito à discussão abaixo sobre os bancos de dados de mapeamento 210, 212 e 214. Como tal, o banco de dados de mapeamento 210, 212 e 24 pode ser treinado para suportar coeficientes de reflexão ou coeficientes cepstrais.

O estimador de envoltória 224 pode realizar a ampla tarefa de estimar uma envoltória espectral de banda larga a partir de uma envoltória espectral de banda estreita. O estimador de envoltória 224 pode receber como entrada, a partir do conversor de banda estreita 223, um conjunto de coeficientes de reflexão de banda estreita que o estimador de envoltória 224 pode apresentar para o seletor de banco de dados 120. O seletor de banco de dados 120 pode converter o conjunto de coeficientes de reflexão de banda estreita em um conjunto de coeficientes de reflexão de banda larga. Assim, o estimador de envoltória 224, através do seletor de banco de dados 120, pode estimar uma envoltória espectral de banda larga a partir de uma envoltória de banda estreita com base em uma transformação não linear dos coeficientes de reflexão de banda estreita usando os bancos de dados de mapeamento selecionados 210, 212 e 214.

Por exemplo, o seletor de banco de dados 120 pode receber como entrada um conjunto de coeficientes de

reflexão de banda estreita gerado pelo conversor de banda estreita 223. Através de uma modelagem estatística, o seletor de banco de dados 120 pode converter o conjunto de coeficientes de reflexão de banda estreita em um conjunto
5 de coeficientes de reflexão de banda larga. O estimador de envoltória 224 então pode passar os coeficientes de reflexão de banda larga para o conversor de banda larga 225, o qual pode convertê-los em um conjunto de coeficientes de LPC de banda larga. Os coeficientes de LPC
10 podem ser denotados por $B(z)$, o que pode representar uma aproximação all-pole para uma envoltória espectral de banda larga.

Conforme citado anteriormente, o seletor de banco de dados 120 pode receber a informação de taxa de amostragem
15 selecionada a partir da seção de avaliação 110. A seção de avaliação 110 pode identificar uma região de suporte com base em taxas de amostragem suportadas pelo sistema. A taxa de amostragem selecionada pode determinar quais bancos de dados de mapeamento 210, 212 e 214 são selecionados pelo
20 seletor de banco de dados 120. Como um exemplo, os bancos de dados de mapeamento 210, 212 e 214 podem ser Modelos de Mistura Gaussianos. Deve ser notado, contudo, que os bancos de dados de mapeamento 210, 212 e 214 não estão limitados a esta configuração em particular. Por exemplo, aqueles de
25 conhecimento na técnica apreciarão que há formas diferentes de implementação de funções de mapeamento, tais como Quantificação de Vetor ou Modelos de Markov Ocultos.

GMMS podem ser úteis em aplicações de modelagem estatística nas quais uma informação que representa
30 características gerais ou tendências deve ser extraída a

partir de uma quantidade grande de dados. Funções de mapeamento, tais como GMMs são úteis ao proporcionarem um insight estatístico de grandes quantidades de dados e para aplicação da informação estatística. Os GMMs são conhecidos
5 na técnica, embora uma breve descrição seja de utilidade para ilustração da maneira pela qual os GMMs são aplicados para a conversão de um conjunto de coeficientes de banda estreita em um conjunto de coeficientes de banda larga.

Com referência às FIG. 2 e 7, um conjunto de
10 coeficientes de banda estreita provido pelo extrator de recurso 222 pode ser submetido como uma entrada 702 para um GMM 700 através do seletor de banco de dados 120. O GMM 700 pode representar um dos bancos de dados de mapeamento 210, 212 e 214, por exemplo. Pode haver quatorze coeficientes de
15 entrada, denotados como X_1 a X_{14} , e quatorze coeficientes de saída correspondentes, denotados como X_{est1} a X_{est14} , na ilustração da FIG. 7, embora o GMM 700 possa receber como entrada e saída qualquer número adequado de coeficientes. O seletor de banco de dados 120 pode decidir qual combinação
20 de GMMs 700 é para ser usada para mapeamento do conjunto de coeficientes de reflexão. A saída do GMM 700 será um conjunto de coeficientes de banda larga 704, os quais representam a envoltória espectral de banda larga. O GMM 700 pode determinar estatisticamente um conjunto de
25 coeficientes de banda larga que mais bem representem as características de uma envoltória de banda larga, dado o conjunto submetido de coeficientes de banda estreita.

Conforme é conhecido na técnica, um GMM tenta determinar uma transformação ótima, conhecida como
30 mapeamento, a qual pode ser aplicada a um sinal de entrada

para conversão dele em um sinal de saída de acordo com a informação estatística provida pelo GMM. Deve ser notado que o GMM pode prover capacidades de modelagem estatística com base em um procedimento de aprendizado denominado

5 treinamento, um processo que é conhecido na técnica. Em resumo, um GMM é originalmente apresentado off-line como dados de treinamento de entrada e de saída para aprender as estatísticas associadas às transformações de dados de entrada para saída. O GMM pode empregar um algoritmo de

10 Maximização de Expectativa (EM) para aprender o mapeamento entre o conjunto de entrada e de saída de coeficientes.

Com referência à FIG. 7, o GMM 700 pode suportar um conjunto de 128 gaussianos 706, onde cada gaussiano é representado por um conjunto de parâmetros μ , Σ , ω

15 descrevendo as estatísticas de um gaussiano único 706. Um gaussiano único 706 pode representar uma função de probabilidade que pode ser descrita pela equação abaixo:

$$p(x) = \frac{1}{(2\pi)^{D/2} |\Sigma|^{1/2}} \exp \left\{ -\frac{1}{2} (x - \mu)' (\Sigma)^{-1} (x - \mu) \right\}$$

20 onde x pode ser o vetor de coeficiente de reflexão de comprimento 14×1 , μ é o vetor de coeficiente de reflexão médio de comprimento, Σ é a matriz de covariância de tamanho 14×14 para os quatorze coeficientes de reflexão, e D pode ser a dimensão do gaussiano 706, o qual é igual ao

25 comprimento do vetor x , o qual é de 14.

Cada gaussiano 706 pode capturar uma porção da informação estatística total contida nos mapeamentos treinados entre coeficientes de reflexão de banda estreita e de banda larga. Por exemplo, a distribuição de

30 probabilidade de um gaussiano único 706 com dimensão $D = 2$

pode ser vista como a curva em formato de sino 740. O gaussiano 706 pode ser uma função de distribuição de probabilidade que descreve uma probabilidade de observação de um coeficiente de reflexão de entrada no gaussiano associado 706. Cada gaussiano 706 pode prover um valor de probabilidade para cada coeficiente de reflexão na entrada representado como uma medida de probabilidade para o gaussiano 706. Em resumo, cada conjunto de entrada de coeficientes será comparado a cada gaussiano 706, e cada gaussiano 706 pode prover alguma porção da informação de mapeamento estatística 708.

A informação de probabilidade a partir de cada gaussiano 706 pode ser ponderada 710 e adicionada em conjunto 712 para instanciação do mapeamento de banda estreita para banda larga. O termo ponderação neste contexto pode significar que a informação de probabilidade provida por cada gaussiano 706 é multiplicada por um valor ponderado. O vetor de média, μ , e a matriz de covariância, Σ , representam as estatísticas associadas a cada gaussiano 706.

Um GMM 700 pode suportar qualquer número de gaussianos 706, embora um 700 que incluía 128 gaussianos possa prover as capacidades de mapeamento adequadas para o conjunto de coeficientes de reflexão, quando uma informação estatística suficiente for adquirida a partir de um conjunto grande de dados de treinamento. Deve ser notado, também, que o conjunto de coeficientes de reflexão pode ser convertido em um conjunto de coeficientes cepstrais, o qual pode ser usado com o mapeamento de GMM. Esta conversão pode reduzir a quantidade de memória requerida pelo GMM 700 porque pode

comprimir uma matriz de covariância cheia gaussiana em um vetor diagonal de variâncias.

Por exemplo, a conversão pode consistir em uma transformação matemática linear que pode converter um conjunto de coeficientes de reflexão dependentes estatisticamente em um conjunto de coeficientes cepstrais estatisticamente independente. Um conjunto estatisticamente dependente de coeficientes geralmente requer uma matriz de covariância cheia 750. Uma matriz cheia significa que todos os termos na matriz são usados no GMM 700. Um conjunto estatisticamente independente de coeficientes apenas geralmente requer o vetor diagonal de uma matriz de covariância 760. Um vetor diagonal significa que é necessário que apenas os termos da diagonal da matriz de covariância sejam armazenados no GMM 700. Por exemplo, uma matriz de covariância $N \times N$ pode ser reduzida para um vetor $N \times 1$, o que pode reduzir as exigências de armazenamento em memória do GMM 700 por um fator de N .

Cada um dos quatorze coeficientes de reflexão da entrada 702 pode ser apresentado para cada um dos 128 gaussianos 706. Cada gaussiano 706, por exemplo, o 128º gaussiano 706 pode ser caracterizado por sua média μ 744 e sua covariância Σ 750, as quais em conjunto podem descrever o formato da função de probabilidade gaussiana 740. Um GMM 700 pode ser um grupo de 128 gaussianos que são misturados em conjunto com base nas características do sinal de entrada. Os 128 gaussianos 706 podem ser misturados em conjunto usando-se um conjunto de pesos ω 710 e uma operação de adição 712. Os pesos ω 710 podem ser determinados durante um treinamento de um algoritmo de EM.

Para um vetor de recurso de dimensão 14 (isto é, 14 coeficientes de reflexão), a operação de mistura 712 usada para a função de probabilidade pode ser:

$$p(x) = \sum_{i=1}^M w_i p_i(x)$$

5

a qual é uma combinação linear ponderada de $M=128$ gaussianos 706 com um vetor de média μ e uma matriz de covariância Σ_i . Os pesos de mistura podem ser restritos a $\sum_{i=1}^M w_i = 1$. Os parâmetros do modelo de densidade podem ser $\lambda =$
 10 $\{w_i, \mu_i, \Sigma_i\}$, onde $i = 1, \dots, M$.

Uma vez que $p(x)$ seja encontrada, a estimativa para o conjunto de coeficientes de reflexão de banda larga pode ser determinada conforme se segue:

$$\rho(x) = \frac{w_i \cdot \rho_i(x)}{\rho(x/\lambda)}$$

15

$$x_{est} = \sum_j \rho(x) \cdot \left((\mu_j - (x' - \mu_i))' \cdot (\sum_{ij})^{-1} (\sum_{ij}') \right)$$

A equação acima revela as propriedades de mapeamento do GMM 700 expressas como uma equação e se refere ao conjunto de
 20 coeficientes de reflexão de banda estreita como uma entrada 702 para o GMM 700 para uma saída 704 representando o conjunto de coeficientes de reflexão de banda larga. O termo $p(x)$ pode ser determinado pelo GMM 700 (isto é, μ_i é o iésimo vetor de média para o iésimo gaussiano 706), e x
 25 (por exemplo, X_1 a X_{14}) representa o conjunto de entrada de coeficientes de reflexão de banda estreita. Também, x_{est} (por exemplo, X_{est1} a X_{est14}) reflete o conjunto estimado de coeficientes de reflexão de banda larga avaliados para o conjunto de entrada de coeficientes de reflexão de banda
 30 estreita. As operações matemáticas do mapeamento de GMM

descritas acima podem ser realizadas pelo estimador de envoltória 224 e pelo seletor de banco de dados 120 da FIG. 2, de acordo com a etapa 440 da FIG. 4.

Com referência de volta à FIG. 5, na etapa 442, uma
5 excitação espectral de banda larga pode ser criada a partir da envoltória espectral de banda larga e do sinal de voz. Um exemplo deste processo é apresentado nas etapas 444 a 448. Na etapa 444, uma excitação espectral de banda estreita pode ser extraída a partir do sinal de voz usando-
10 se o conjunto de coeficientes de reflexão de banda larga ou um conjunto de coeficientes de LPC de banda estreita, conforme provido na etapa 440. Na etapa 446, o sinal de excitação de banda estreita pode ser estendido para um sinal de excitação de banda larga. Um exemplo de como esse
15 processo pode ser realizado é mostrado nas etapas 448A a 448F.

Especificamente, na etapa 448A, uma excitação de banda baixa pode ser gerada e, na etapa 448B, uma excitação de banda alta pode ser gerada. Por exemplo, na etapa de opção
20 448C, a excitação de banda baixa e a excitação de banda alta podem ser moduladas usando-se uma multiplicação de cosseno. Na etapa de opção 448D, a excitação de banda baixa e a excitação de banda alta podem ser adicionadas com a excitação de banda estreita (ou excitação de passa banda)
25 para a criação de uma excitação de meia banda. Na etapa 448F, uma excitação de banda larga pode ser gerada a partir da excitação de meia banda.

Por exemplo, com referência à FIG. 2, a seção de análise de banda larga 242 pode gerar a excitação de banda
30 estreita pela filtração inversa do sinal de voz reamostrado

105 com um conjunto de coeficientes de reflexão. A
filtração inversa pode requerer o conjunto de coeficientes
de banda larga apresentado pelo estimador de envoltória 224
ou, alternativamente, ele pode usar os coeficientes de LPC
5 de banda estreita gerados no extrator de recurso 222. O
conjunto de coeficientes de banda estreita ou de banda
larga pode ser usado na seção de análise de banda larga 242
para a geração da excitação de banda estreita. Uma
filtração inversa do sinal de voz reamostrado 105 com
10 qualquer um dos conjuntos de coeficientes pode gerar um
sinal de excitação de banda estreita, porque o sinal de voz
reamostrado 105 em si é de banda estreita.

A excitação de banda estreita pode ser passada através
do estágio de excitação de percurso múltiplo 244 para a
15 criação de uma excitação de banda larga. A finalidade do
estágio de excitação de percurso múltiplo 244 é criar um
sinal de excitação artificial na região de suporte 636
(veja a FIG. 6). Pode ser considerado artificial no sentido
que a excitação suplementar pode ser gerada pela replicação
20 e pelo deslocamento do sinal de excitação de banda estreita
reamostrado.

Com referência, agora, às FIG. 2, 3 e 6, o estágio de
excitação de percurso múltiplo 244 pode receber a excitação
de banda estreita a partir da seção de análise de banda
25 larga 242. A excitação de banda estreita pode divergir
através de vários percursos que podem acumular ou estender
a excitação de banda estreita recebida. Por exemplo, a
excitação de banda estreita pode passar através do estágio
de excitação de banda baixa 310, do estágio de excitação de
30 banda alta 320 e do estágio de excitação de passa banda

330.

O modulador 312 do estágio de excitação de banda baixa 310 pode modular a excitação de banda estreita para, por exemplo, uma região que ocorre na região de frequência inferior de suporte 633 (por exemplo, de 0 Hz a 300 Hz). O modulador 322 do estágio de excitação de banda alta 320 pode modular a excitação de banda estreita para uma região ocorrendo em uma porção da região superior de frequência mais alta de suporte 637 (por exemplo, de 3,4 kHz a 4 kHz). Como um exemplo, uma multiplicação de co-seno pode ser usada para a modulação do sinal de excitação de banda estreita para regiões de suporte 633, 637 descrita acima.

O filtro de passa baixa 314 do estágio de excitação de banda baixa 310 pode remover as componentes entrelaçadas ("aliased") causadas pela modulação. De forma similar, o filtro de passa banda 324 do estágio de excitação de banda alta 320 pode remover as componentes entrelaçadas feitas pela modulação. O estágio de excitação de passa banda 330 pode permitir que a excitação de banda estreita passe não processada, o que pode permitir que ela permaneça em sua largura de banda original (por exemplo, de 300 Hz a 3,4 kHz).

O adicionador 340 pode somar em conjunto as excitações de banda baixa, de banda alta e de passa banda para a geração de uma excitação de meia banda, a qual pode se estender a partir de 0 Hz a 4 kHz, com base em nosso exemplo. Em seguida, o modulador 350, usando uma multiplicação de co-seno, por exemplo, pode modular a excitação de meia banda para a criação de uma excitação de banda plena ou de banda larga. A modulação da excitação de

meia banda para uma excitação de banda larga pode corresponder às frequências de 4 kHz a 8 kHz. Mediante a conclusão do estágio de excitação de percurso múltiplo 244, o sinal de excitação de banda estreita pode ser estendido
5 para um sinal de excitação de banda larga.

Deve ser notado que o modulador 312, o modulador 322 e o modulador de meia banda 350 não estão restritos à modulação de dados apenas na região de suporte 636. Por exemplo, pode ser necessário ter alguma superposição no
10 deslocamento nas fronteiras da região de suporte 636. Através desta superposição, a resposta de frequência do sinal de excitação de banda larga pode ser espectralmente plana, uma característica desejada, conforme é conhecido na técnica.

15 Com referência de volta ao método 400 da FIG. 5, na etapa 450, um sinal de voz de banda larga pode ser gerado pela combinação da envoltória espectral de banda larga criada em conjunto com a excitação de banda larga criada e o sinal de voz. As etapas 452 a 456 apresentam um exemplo
20 de como este processo pode ser feito. Em particular, a envoltória de banda larga provida pela etapa 436 pode ser combinada com a excitação de banda larga provida pela etapa 442 para a geração de um sinal de voz de banda larga sintético, conforme mostrado na etapa 452. O sinal de voz
25 de banda larga sintético pode conter um conteúdo espectral dentro da região de suporte e também a largura de banda de voz desconhecida original.

Na etapa 454, um sinal de voz de banda larga suplementar pode ser extraído a partir do sinal de voz de
30 banda larga sintético na região de suporte. O conteúdo

espectral no sinal de voz de banda larga sintético que representa a mesma região de frequência da largura de banda de voz desconhecida original pode ser removido, se o sinal de voz desconhecido original for para ser combinado com o
5 sinal de voz de banda larga suplementar. Esta etapa pode ser executada porque não é necessário duplicar o conteúdo espectral original do sinal de voz. Na etapa 456, o sinal de voz de banda larga suplementar pode ser adicionado ao sinal de voz para a geração de um sinal de voz de banda
10 larga. O método 400 pode terminar na etapa 458.

Como um exemplo e com referência às FIG. 2 e 6, o processador de mistura 260 pode misturar um sinal de voz de banda larga suplementar com o sinal de voz reamostrado 105 para a geração de um sinal de voz de banda larga. O sinal
15 de voz de banda larga suplementar pode ser extraído a partir de um sinal de voz de banda larga sintético. Por exemplo, a seção de síntese de banda larga 262 pode usar os coeficientes de LPC de banda larga providos pelo conversor de banda larga 225 como coeficientes de filtro de síntese.
20 A seção de síntese de banda larga 262 também pode receber como entrada o sinal de excitação de banda larga provido pelo estágio de excitação de percurso múltiplo 244. A seção de síntese de banda larga 262 pode gerar um sinal de voz de banda larga sintético pela filtração do sinal de excitação
25 de banda larga com coeficientes de filtro de LPC de banda larga. O sinal de voz resultante é um sinal de voz de banda larga sintético. Em nosso exemplo, o sinal de voz de banda larga sintético pode ser estendido de 0 Hz para 8 kHz.

Conforme mencionado previamente, o conteúdo espectral
30 pode ser seletivamente removido do sinal de voz de banda

larga sintético para a geração de um sinal de voz de banda larga complementar. O sinal de voz de banda larga complementar pode ser gerado pela passagem de um sinal de voz de banda larga sintético através do filtro de pára banda 264. O filtro de pára banda 264 pode suprimir o conteúdo espectral fora ou dentro da região de suporte 636.

Especificamente, o sinal de voz desconhecido original já provê um conteúdo espectral na largura de banda de voz 625 (por exemplo, de 300 Hz a 3,4 kHz). Devido ao fato de o sinal de voz de banda larga sintético também conter um conteúdo espectral que corresponde ao conteúdo espectral contido na largura de banda de voz 625, o filtro de pára banda 264 pode suprimir o conteúdo espectral do sinal de voz de banda larga sintético que se sobrepõe ao conteúdo espectral do sinal de voz reamostrado 105. Assim, o sinal de voz desconhecido pode precisar apenas de um conteúdo espectral suplementar fora de sua própria largura de banda (por exemplo, de 0 a 300 Hz e de 3,4 kHz a 8 kHz). O adicionador 266 pode adicionar o sinal de voz de banda larga suplementar ao sinal de voz reamostrado 105 para a geração do sinal de voz de banda larga.

Quando aplicável, a presente invenção pode ser realizada em hardware, subsequente ou em uma combinação de hardware e de software. Qualquer tipo de sistema de computador ou um outro aparelho adaptado para a realização dos métodos descritos aqui é adequado. Uma combinação típica de hardware e de software pode ser um dispositivo de comunicações móveis com um programa de computador que, quando for carregado e executado, pode controlar o dispositivo de comunicações móveis de modo que ele realize

os métodos descritos aqui. Porções da presente invenção também podem ser embutidas em um produto de programa de computador, o qual compreende todos os recursos permitindo a implementação dos métodos descritos aqui e os qual,
5 quando carregado em um sistema de computador, é capaz de realizar estes métodos.

Embora as modalidades preferidas da invenção tenham sido ilustradas e descritas, será claro que a invenção não está limitada dessa forma. Numerosas modificações, mudanças
10 e variações, substituições e equivalentes ocorrerão àqueles versados na técnica, sem se desviar do conceito inventivo e do escopo da presente invenção, conforme definido pelas reivindicações em apenso.



PI0612564-6

REIVINDICAÇÕES

1. Método para extensão de largura de banda para comunicações de voz, caracterizado por compreender:

o recebimento de um sinal de voz desconhecido;

5 a identificação da largura de banda de voz do sinal de voz desconhecido recebido;

o estabelecimento de uma região de suporte tendo em vista o conteúdo espectral do sinal de voz recebido; e

a seleção de uma combinação de bancos de dados de mapeamento a partir de uma pluralidade de bancos de dados de mapeamento, cada banco de dados de mapeamento associado a uma faixa de extensão de largura de banda predeterminada para estender a largura de banda de voz.

2. Método, de acordo com a reivindicação 1, caracterizado pelo fato do estabelecimento de uma região de suporte compreender:

a emissão de uma requisição para um objeto subjacente para retorno de uma lista de frequências de amostragem para as quais um objeto é capaz de suportar;

20 a identificação de limites espectrais com base na frequência de amostragem retornada; e

a determinação de bandas espectrais nos limites espectrais para extensão da largura de banda de voz para regiões que residem fora da largura de banda de voz.

25 3. Método, de acordo com a reivindicação 1, caracterizado pelo fato da seleção de uma combinação de bancos de dados de mapeamento ser uma operação sequencial e compreender ainda a aplicação de uma combinação serial de bancos de dados mapeados para se estender coletivamente a largura de banda de voz para uma faixa correspondente à

adição das faixas de extensão de largura de banda selecionadas.

4. Método, de acordo com a reivindicação 1, caracterizado por compreender ainda:

5 a aquisição de um conjunto de coeficientes de reflexão de banda estreita que representam a envoltória espectral a partir do sinal de voz; e

a extensão do conjunto de coeficientes de reflexão de banda estreita para um conjunto de coeficientes de reflexão de banda larga usando-se bancos de dados de mapeamento para a geração de uma envoltória espectral de banda larga.

5. Método, de acordo com a reivindicação 1, caracterizado por compreender ainda:

15 a extração de um sinal de excitação de banda estreita a partir do sinal de voz usando-se um conjunto de coeficientes de reflexão de banda larga ou um conjunto de coeficientes de análise de predição linear de banda estreita; e

20 a extensão do sinal de excitação de banda estreita para um sinal de excitação de banda larga usando-se modulação e filtração.

6. Método, de acordo com a reivindicação 1, caracterizado por compreender ainda:

25 a combinação de um sinal de excitação de banda larga com uma envoltória espectral de banda larga para a geração de um sinal de voz de banda larga sintética;

a extração de um sinal de voz de banda larga suplementar a partir do sinal de voz de banda larga sintético na região de suporte; e

30 a adição do sinal de voz de banda larga sintético

suplementar com o sinal de voz para a geração de um sinal de voz de banda larga.

7. Sistema para estender artificialmente a largura de banda de voz, caracterizado por compreender:

5 uma seção de avaliação, que recebe um sinal de voz desconhecido e determina uma extensão admissível de largura de banda de voz para o sinal de voz desconhecido;

10 um seletor de banco de dados cooperativamente acoplado à seção de avaliação, onde o seletor de banco de dados escolhe uma combinação de bancos de dados de mapeamento de acordo com a extensão admissível de largura de banda de voz; e

15 uma unidade de extensão de largura de banda acoplada de forma cooperativa à seção de avaliação e ao seletor de banco de dados, onde a unidade de extensão de largura de banda estende a largura de banda de voz do sinal de voz desconhecido para a extensão admissível de largura de banda de voz usando-se a combinação de bancos de dados de mapeamento escolhida pelo seletor de banco de dados.

20 8. Sistema, de acordo com a reivindicação 7, caracterizado pelo fato da seção de avaliação compreender:

 um módulo de análise que identifica uma largura de banda de voz associada ao sinal de voz desconhecido;

25 um módulo de investigação acoplado de forma cooperativa ao módulo de análise, onde o módulo de investigação identifica taxas de amostragem suportadas, onde as taxas de amostragem suportadas revelam a extensão até a qual a largura de banda de voz pode ser estendida; e

30 um módulo de amostragem acoplado de forma cooperativa ao módulo de análise e ao módulo de investigação, onde o

módulo de amostragem reamostra o sinal de voz desconhecido em uma das taxas de amostragem suportadas identificadas pelo módulo de investigação, onde a reamostragem prepara o sinal de voz para extensão de largura de banda.

5 9. Sistema, de acordo com a reivindicação 7, caracterizado pelo fato da unidade de extensão de largura de banda compreender:

um processador de envoltória cooperativamente acoplado à seção de avaliação e ao seletor de banco de dados, onde o
10 processador de envoltória determina uma envoltória espectral de banda estreita a partir do sinal de voz e subsequente provê um conjunto de coeficientes de banda larga representando uma envoltória espectral de banda larga;

15 um processador de excitação acoplado de forma cooperativa à seção de avaliação e ao processador de envoltória, onde o processador de excitação determina um sinal de excitação de banda estreita a partir do sinal de voz usando-se um conjunto de coeficientes de reflexão de
20 banda larga ou um conjunto de coeficientes de análise de predição linear de banda estreita e subsequente cria um sinal de excitação de banda larga; e

um processador de mistura acoplado de forma cooperativa à seção de avaliação, o processador de
25 envoltória e o processador de excitação, onde o processador de mistura combina o sinal de voz em conjunto com o sinal de excitação de banda larga e a envoltória espectral de banda larga para a criação de um sinal de voz de banda larga.

30 10. Sistema, de acordo com a reivindicação 9,

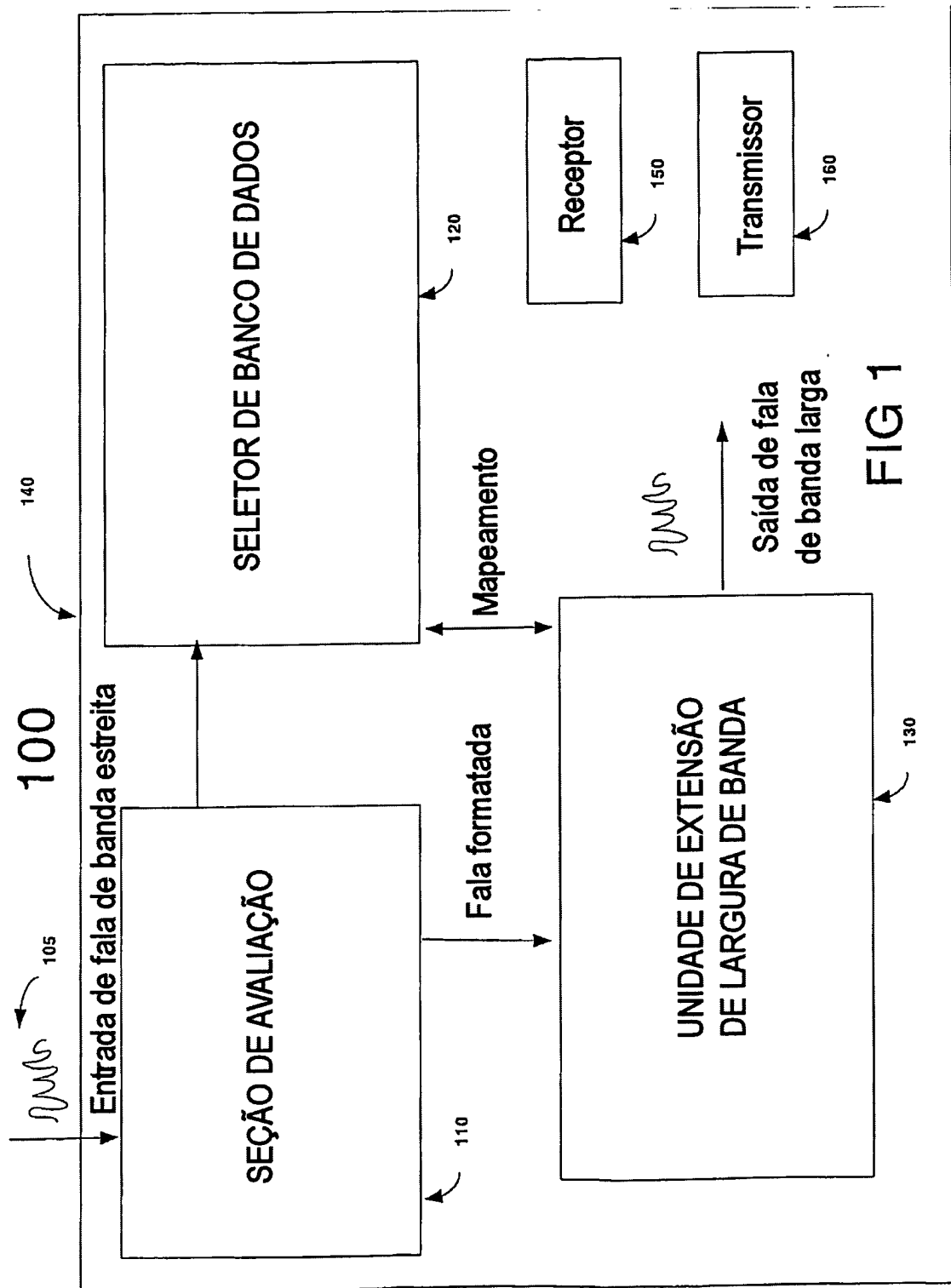
caracterizado pelo fato do processador de envoltória compreender:

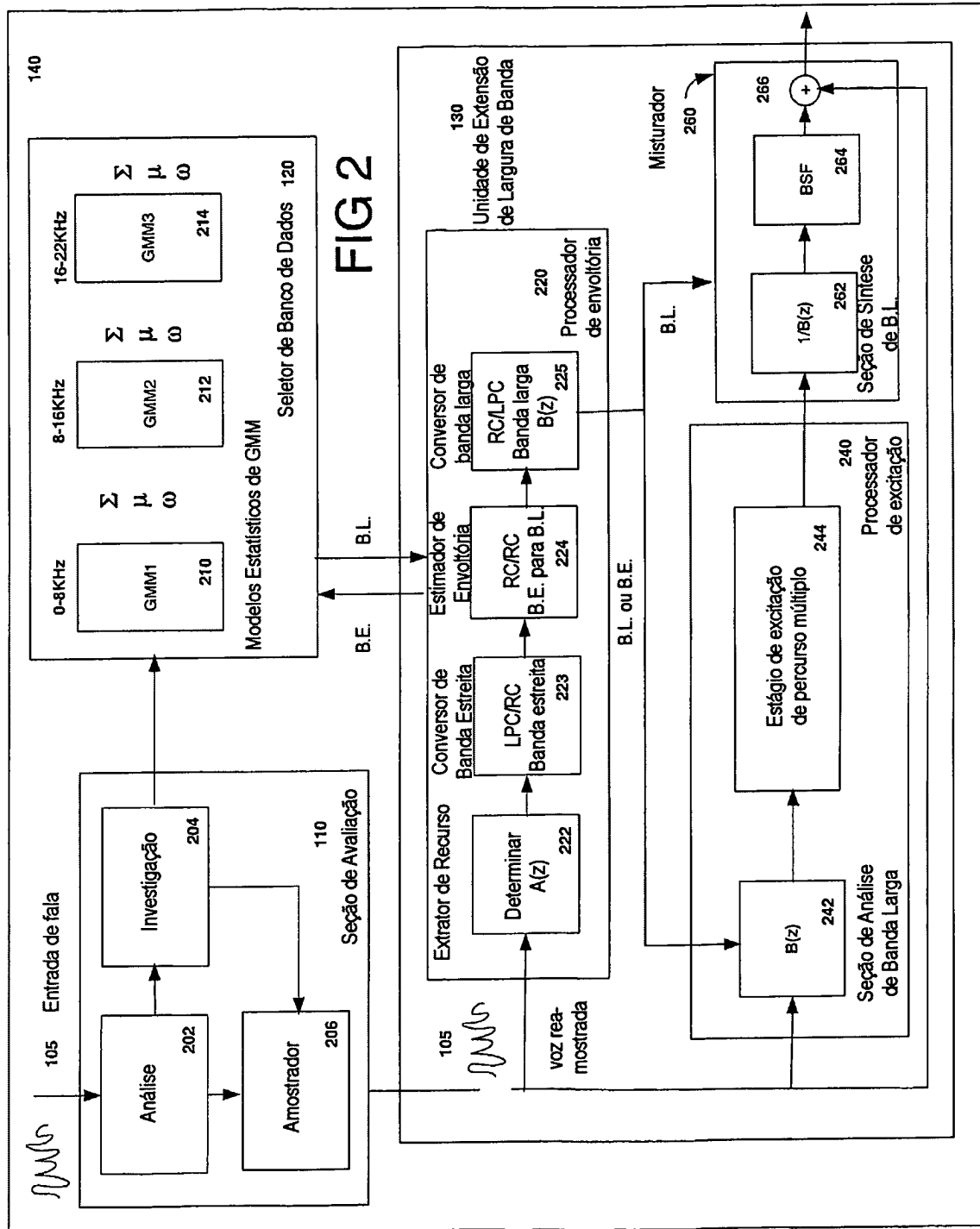
um extrator de recurso que adquire um conjunto de coeficientes de análise de predição linear que representam
5 a envoltória espectral do sinal de voz;

um conversor de banda estreita acoplado de forma comunicativa ao extrator de recurso, onde o conversor de banda estreita converte o conjunto de coeficientes de análise de predição linear em um conjunto de coeficientes
10 de reflexão de banda estreita;

um estimador acoplado de forma comunicativa ao conversor de banda estreita, onde o estimador, em conjunto com o seletor de banco de dados, estende o conjunto de coeficientes de reflexão de banda estreita para um conjunto
15 de coeficientes de reflexão de banda larga usando os bancos de dados de mapeamento; e

um conversor de banda larga acoplado de forma comunicativa ao estimador, onde o conversor de banda larga converte os coeficientes de reflexão de banda larga em um
20 conjunto de coeficientes de análise de predição linear de banda larga.





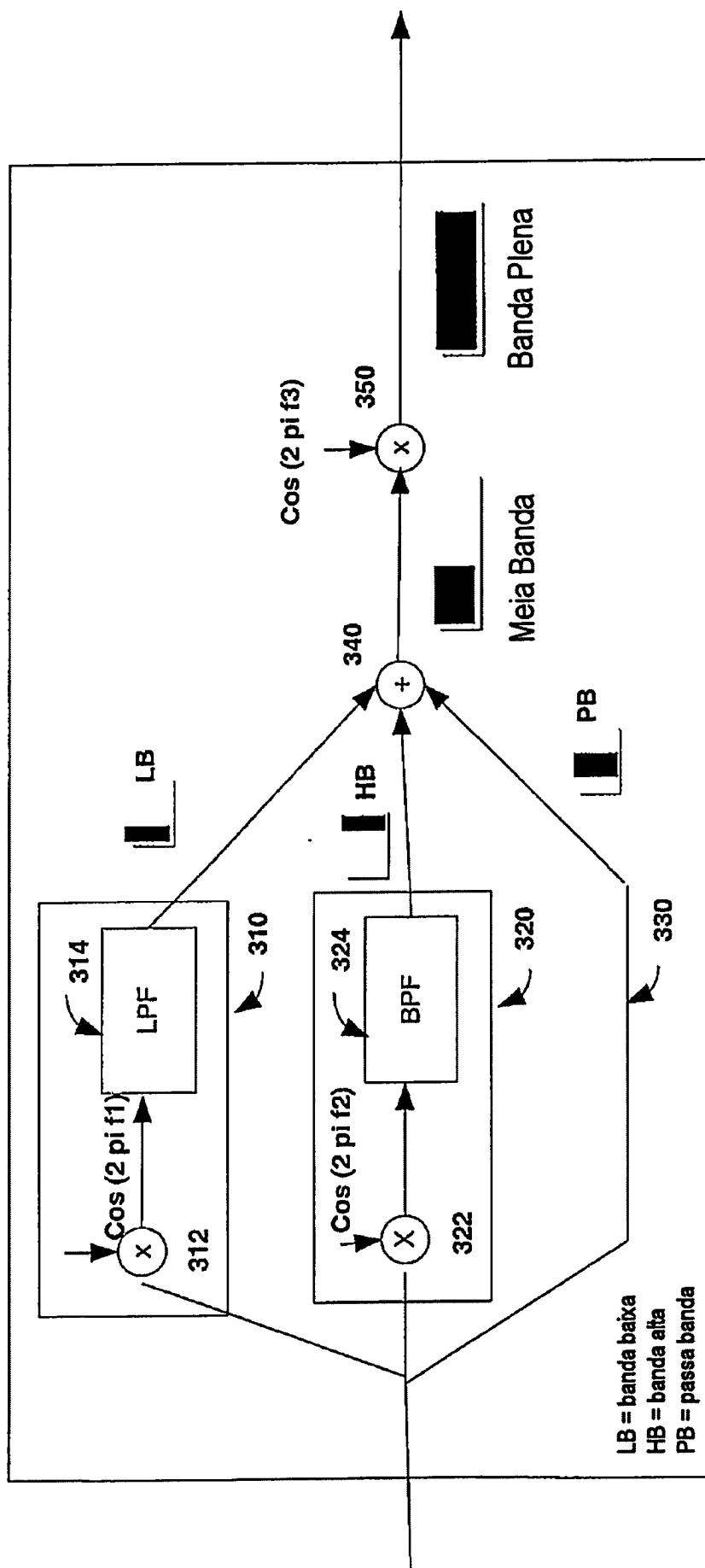
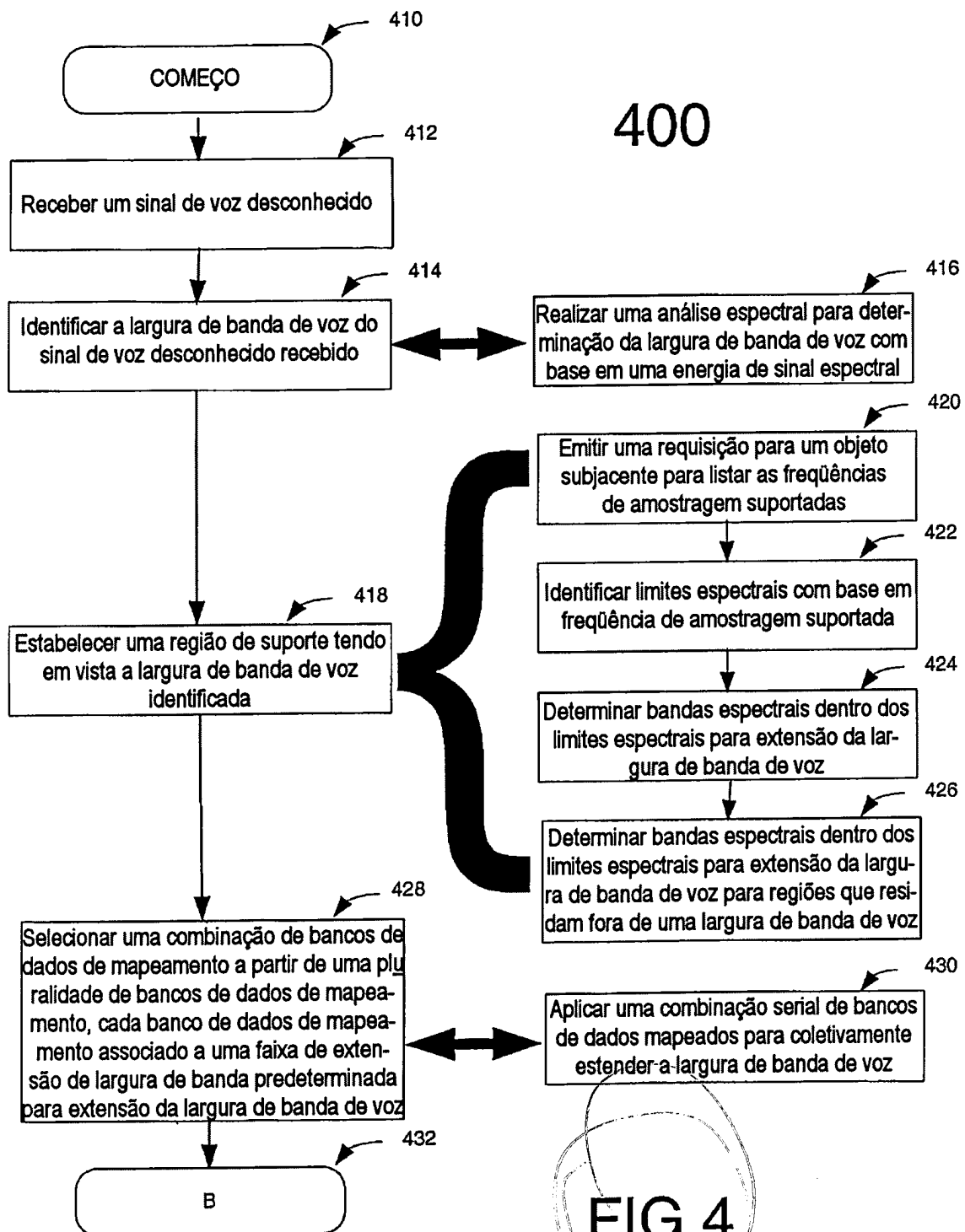


FIG 3



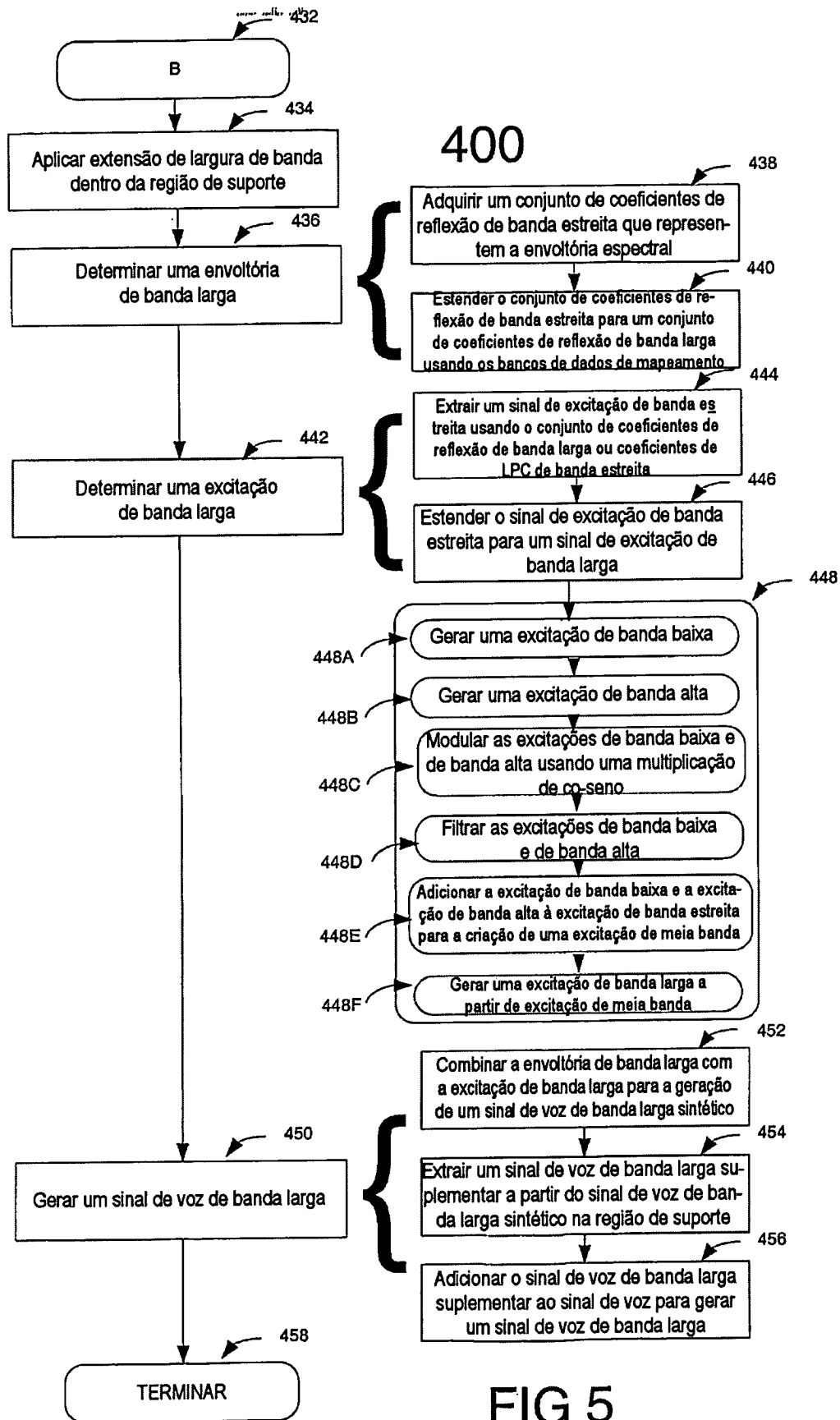


FIG 5

FIG 6

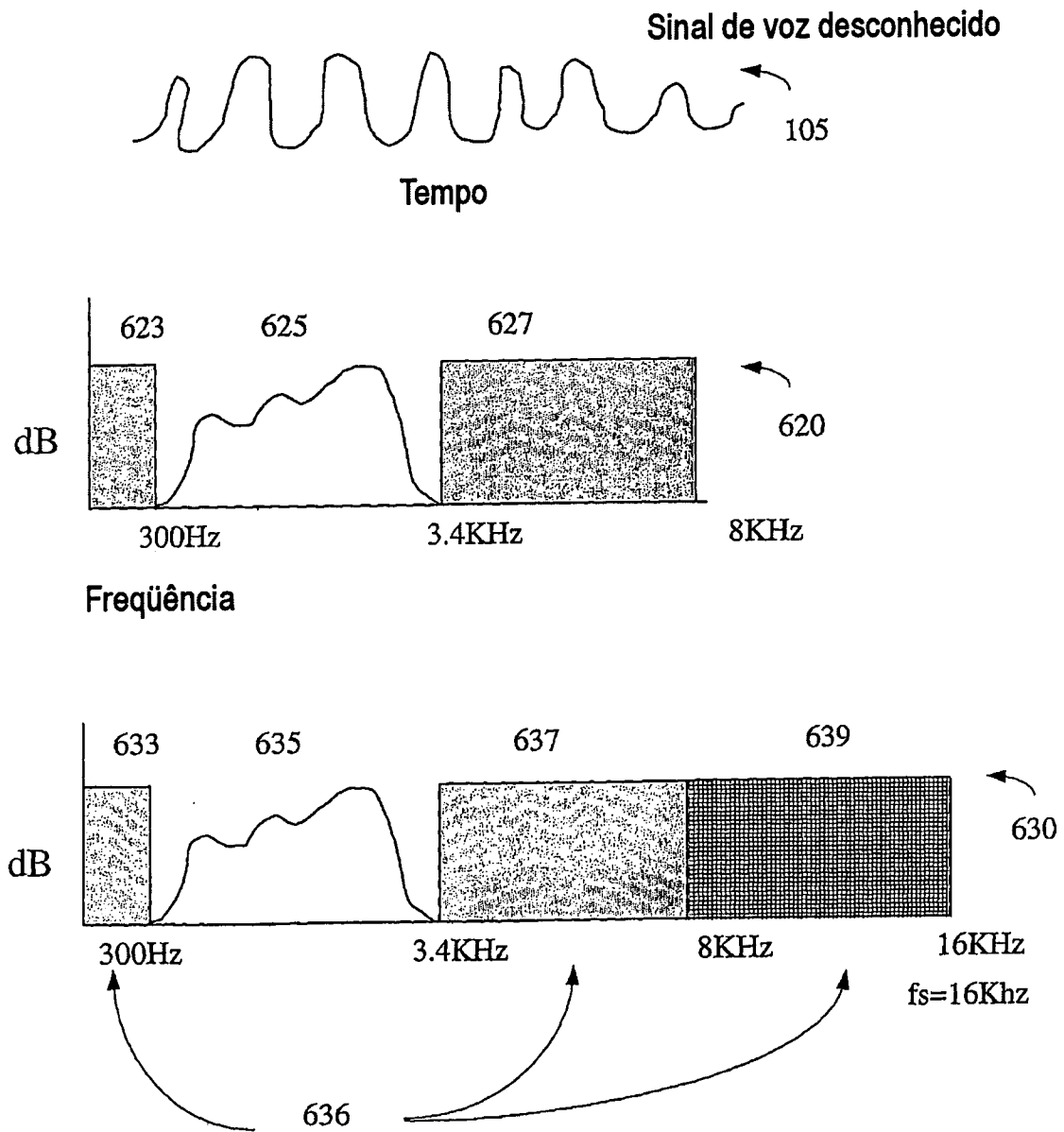
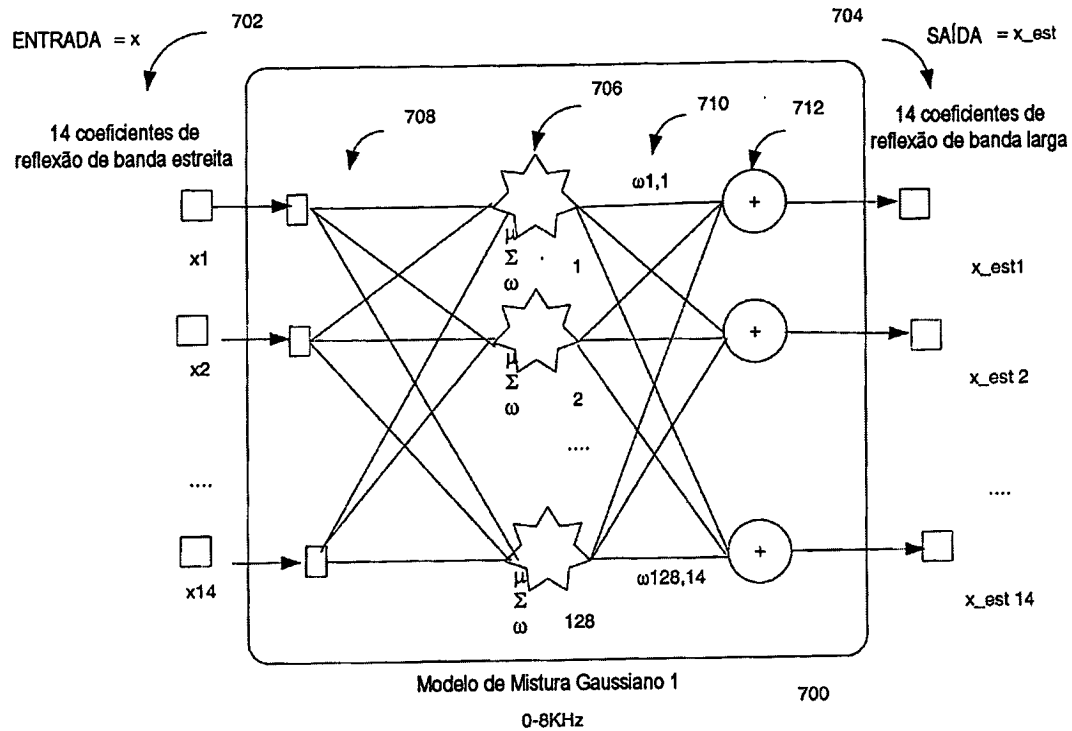


FIG 7



750

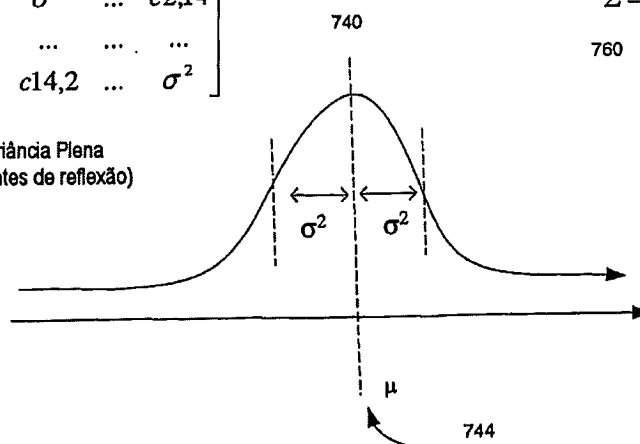
$$\Sigma = \begin{bmatrix} \sigma^2 & c_{1,2} & \dots & c_{1,14} \\ c_{2,1} & \sigma^2 & \dots & c_{2,14} \\ \dots & \dots & \dots & \dots \\ c_{14,1} & c_{14,2} & \dots & \sigma^2 \end{bmatrix}$$

Covariância Plena
(coeficientes de reflexão)

760

$$\Sigma = \begin{bmatrix} \sigma^2 & \dots & \dots \\ \dots & \sigma^2 & \dots \\ \dots & \dots & \dots & \sigma^2 \end{bmatrix}$$

Covariância Diagonal
(coeficientes cepstrais)



**MÉTODO E SISTEMA PARA EXPANSÃO DE BANDA LARGA PARA
COMUNICAÇÕES DE VOZ**

A invenção concerne a um método (400) e um sistema (100) para extensão de largura de banda de voz para
5 melhoria da qualidade de voz em um sistema de comunicação. O método pode incluir as etapas de recebimento (412) de um sinal de voz desconhecido (105), identificação (414) da largura de banda de voz (625) do sinal de voz desconhecido recebido e estabelecimento (418) de uma região de suporte
10 (636) tendo em vista o conteúdo espectral do sinal de voz recebido. O método ainda pode incluir a etapa de seleção (428) de uma combinação de bancos de dados de mapeamento (210, 212, 214) a partir de uma pluralidade de bancos de dados de mapeamento. Cada banco de dados de mapeamento pode
15 estar associado a uma faixa de extensão de largura de banda predeterminada para extensão da largura de banda de voz.